



Discussion Paper Series

2003 – 09

Department of Economics
Royal Holloway College
University of London
Egham TW20 0EX

©2003 Roberto Ricciuti. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit including © notice, is given to the source.

Bringing Macroeconomics into the Lab

Roberto Ricciuti

Economics Department, Royal Holloway College, University of London

Abstract

This paper reviews experiments in macroeconomics, pointing out the theoretical justifications, the strengths and weaknesses of this approach. We identify two broad classes of experiments: general equilibrium and single-issue experiments, and emphasize the idea of theory testing that is behind these. A large number of macroeconomic issues have been analyzed in the laboratory spanning from monetary economics to fiscal policy, from international trade and finance, to growth and macroeconomic imperfections. In a large number of cases results give support to the theories tested. We also highlight that experimental macroeconomics has increased the number of tools available to experimentalists.

Keywords: Macroeconomics, experiments.

“When an engineer wants to find out how the temperature affects material’s conductivity, she builds an experiment in which she changes the temperature, makes sure that everything else remains the same, and looks at the changes in conductivity. But macroeconomists who want to find out, for example, how changes in the money supply affect aggregate activity cannot perform such controlled experiments; they cannot make the world stop while they ask the central bank to change the money supply”.

Olivier Blanchard (1997)

1. Introduction

Although the experimental method has gained considerable standing in the study of many areas of economics, experimental macroeconomics has still a long way to go, as exemplified by the above quotation. Probably no experimental macroeconomics would have been possible without the microfoundation revolution that hit macroeconomics in the 1970s. Since then macroeconomists have started obtaining macroeconomic propositions originating from the maximizing behavior of consumers, given some constraints. This revolution gave birth to the New Classical Economics in which general equilibrium propositions were extended to macroeconomics. Subsequently, in response to the puzzles opened by this theory, New Keynesian Macroeconomics has originated to look at the frictions that prevent markets to clear. Therefore, the work surveyed here - in which motivated agents interact in the laboratory to give rise both to individual and aggregate outcomes - can be seen as a partial assessment of the power of micro-models in macroeconomics. For this reason the reader may not find a distinctive feature of macroeconomic experiments over microeconomic ones: the working of the economy is based on agents and the only flavor of macroeconomics is the explicit analysis of aggregates such as inflation, investments, and employment.

There are two general classes of experiments in macroeconomics. The first is concerned with general equilibrium, studying equilibration and spillovers between markets. This approach is maintained by Charles Plott in several of his experiments. The second

approach concentrates on a specific issue and looks at experimentation as a way to test a specific theory. In principle, only the first kind of experiments have a proper macroeconomic content, since these are interested in the inter-relations between several markets and the spill-over between them. The economy at the aggregate level is a complex system in which consumption decisions affect supply, which in turn have an effect on investments, which determine the employment level, and changes in the latter impinge on aggregate demand. Laboratory economies are of course much simpler than the real economy, and the implicit message of this work is: if a simplified version of the economy rejects a model of macroeconomic behavior, this model cannot be applied to the more complex real world. Therefore, non-rejection provides first evidence of the plausibility of a model. The other view is more widespread: most of the experiments reported here fall within this category. It is usually centered around a single market, and there are no feedbacks to the rest of the economy. In fact, these experiments are akin to microeconomic experiments, where the *ceteris paribus* condition can be easily applied. This view is consistent with most current macroeconomic modeling, in which a single market is considered at a time. Since the “testing theory” approach is dominant within experimental macroeconomics, there is no surprise that these kind of experiments have such a prominence.

There are mainly three specific reasons to look at experiments in macroeconomics. Let us draw some examples from the papers we will consider below. The first relies on theoretical models that are primarily concerned with steady-state equilibrium and are based on rational expectations, therefore neglecting any out of equilibrium dynamics. This is the case of the existence of stationary solutions in inflationary processes and reputational models of monetary policy. Therefore, running experiments may shed some light on issues on which models are silent: dynamics towards the equilibrium, the process of expectations formation and coordination is left undefined because expectations are assumed to be fulfilled, and the

complexity of particular market structures and policies is not taken into account. Moreover, some of these models predict multiple equilibria, and experiments become a natural way to look at which one is most likely to occur when individuals are placed in the conditions of the theory. In addition, many models are based on expectations formation, which are typically unobservable and therefore cannot be operationalized in econometric analysis. In contrast, within an experiment subjects may be requested to make conjectures about the future values of a variable, therefore one can follow and evaluate the expectations path.

The second argument is related to empirical work with deficiencies in field data. To see whether a theory is confirmed by data one needs to impose some restrictions. For example, tests on demand of money would require identification of money demand, separation of the transaction demand from other motives, and inclusion of all important determinants of money demand as explanatory variables. It is difficult to get all this information to run an econometric analysis with field data: the identification problem is virtually unsolvable, whereas experiments allow constructing economies where only the transaction motive exists. When it comes to international trade, a direct test of the law of competitive advantages is not possible because autarky is never observed in the real world. To test for theories on exchange rate, one needs to unambiguously measure the price level of the economy, and in the field data price indices may not be uniformly accurate measures of the short-term buying power of a good. The experimentalist supplies consistent inflation data that have a known generating process.

Another argument in favor of experiments is related to the endogeneity of policy in real-world economies that makes it difficult to analyze data and formulate correct inferences on changes that have occurred, and to the reactions that policy makers put in place to contrast economic fluctuations. In the lab the experimenter has full control over the parameters, one can make “what if” experimentation perfectly knowing which parameters

have changed and in which direction. Therefore, modeling is flexible since one can implement as many changes in parameters and rules of the game as time and budget constraints allow. Furthermore, results may be replicated by other scholars.

Three main problems can be detected in the use of the experimental method in macroeconomics. Firstly, there is a motivation issue: individuals in macroeconomic systems take several roles: they are workers paid a salary that is necessary to fulfil basic needs, consumers using their rewards, and savers that make both short- and long-term decisions choosing between alternative assets. At the same time, firms co-exist and interact with them with their own features. Students in a laboratory are just motivated by a small amount of money that cannot be representative of the needs and aspirations of an agent interacting in a large economy.¹ A second and related problem is concerned with the number of people involved, which in macroeconomics is huge and in a laboratory economy is rather small. Taken together, these issues put into doubt the realism of experiments. The position we take is that models themselves are simplifications of the working of an economy, and these experiments have their scientific strength in building on economic theory. Furthermore, we share the view that “Laboratory methods allow a dramatic reduction in the number of auxiliary hypotheses involved in examining a primary hypothesis” (Davis and Holt, 1993: 16). The issue of inadequacy of rewards has been extensively analyzed, and there is evidence maintaining that payoffs may matter – higher payoffs reduce “noise” and also bring down, but not eliminate, deviations from theoretical behavior - but it is very difficult to ascertain in which situations this can occur, and in which direction (Smith and Walker, 1993; Camerer, 1995).

A third problem is related with learning. Macroeconomic experiments are often complex, matching and exchanging rules may interact in a rather artificial way, and therefore it is necessary to repeat the experiment many times to get subjects acquainted with the rules

of the game. Referring to their overlapping generations experiment, Marimon and Sunder (1995: 125) point out that “Assets cannot be carried from one ‘life’ to the next but memory and experience obviously are.” In some situations learning processes are the objective of the study, as in their experiments where they aimed at comparing rational expectations and adaptive learning in equilibrium selection. In others we cannot distinguish between learning the rules of the game and just learning how to play with a given set of parameters.

The reader will notice that almost a single exchange institution is used in the following experiments: the double auction, i.e., a market in which buyers and sellers can make asks (offers to sell) and bids (offers to buy) for standardized units of well-defined commodities and securities. This institution is used at the Chicago Board of Trade. It has been popularized by Vernon Smith in market experiments since the birth of experimental economics, and there is a significant body of literature (see Friedman and Rust, 1993 for a comprehensive review) that maintains the superiority of this institution in terms of efficiency and ability to simulate a competitive market. Therefore it is not surprising that it has also been extensively used in macroeconomics. Given the properties of this market institution in the domain of partial (static) equilibrium notion of induced demand and supply, is its overwhelming use in general equilibrium experiments justified? Macroeconomists used to think in terms of short-run disequilibrium and dynamic feedback between different markets would not recognize this as an approximation of how the economy works. We think that there are several reasons that made this exchange institution so popular in experimental macroeconomics. First, its widespread use in experimental markets has made this tool easily available to experimenters that have moved to macroeconomics. Second, and more important, many of the macro models are inherently neoclassical, therefore they rely on the notion of perfect competition, and the experimental market that performs in this way is the double auction one. In this sense the experimenter attempts to mimic one of the conditions of

the model, therefore this choice is coherent with the theory testing approach. For example this is particularly clear in the experiment reported in Section 2, when a general equilibrium model is put under experiment, and one of the key theoretical features of these models is perfect competition. At the same time the reader must bear in mind that the result may be institutionally-sensitive. In this respect, comparisons of the same model with different trading institutions would be an interesting task for the future, since it has been substantially overlooked until now. This also represents a different undertaking compared to model testing, being much closer to the approach put forward by Sunder that experiments are useful to find which hypotheses are crucial in obtaining a result and which other mainly represent nuisances of the model (Sunder, 2001).²

In this paper we do not consider the experiments on co-ordination games with strategic complementarities (Cooper, 1999: 1-17). New Keynesian Macroeconomics attaches great importance to this feature to explain macroeconomic imperfections. Strategic complementarities may result in multiple equilibria with self-fulfilling beliefs. Given our previous discussion, this topic should be one in which experimentation gives the opportunity to see which equilibrium prevails and under which conditions. Unfortunately, this area is rarely investigated in the laboratory, with the notable exception of Van Huyck et al. (1990) who propose an experimental test of the Bryant (1983) model. In this game the best choice for each player would be to choose the maximum effort level, but if a certain agent select another level, the best action for the other players is to choose this other (suboptimal) effort level (in other words there are several rankable Nash equilibria). Results put forward that the chosen effort is usually lower than the optimal level, and therefore coordination failures may take an important role in explaining departures from standard theory.³

Experimental macroeconomics has also been the source of new methods in experiments to deal with issues that are typical of the discipline. One refers to the

implementation of overlapping generations models, which will be discussed in detail in Sections 3 and 4, the other is concerned with procedures that mimic the infinite horizon of many theoretical models in a typically time-constrained situation such as a laboratory experiment.

The paper is organized as follows. Section 2 reviews a general equilibrium macroeconomic experiment that relates to several of the subsequent topics that we analyze. In Section 3 and Section 4, respectively, monetary and fiscal policy experiments are surveyed. Section 5 reports on laboratory methodology and evidence on international economics. In Section 6 a growth model experiment is reviewed. Section 7 looks at experiments exploring macroeconomic imperfections. The last Section concludes.

2. A general equilibrium macroeconomic experiment

We start our review of experiments with the paper by Lian and Plott (1998). It studies the general equilibrium properties of an economy in which consumers sell labor and buy goods to maximize their utility, competitive firms hire workers and sell goods to achieve maximum profits, a financial market is existent, and money supply may be changed. This paper tries to answer the question on the technical possibility of running experiments characterized by such complexity and then, provided that this answer is positive, to qualify the results in light of the neoclassical model.⁴

In Figure 1 the circular flow model of the economy is represented. There are two groups of agents (consumers and producers) in equal number and two goods (X and Y). Consumers desire both goods and in each period they have an endowment of ten units of Y (labor) and none of X (output). Therefore they have to sell labor in exchange for fiat money (francs) to buy X . Producers use labor as an input for X , have a desire to consume Y , and no

preferences over X . They also have a non-linear production function that enables them to transform labor into output. Producers buy Y on the market, use some of it to produce X , and sell it into the market for fiat money. The money could be used to purchase additional labor or to obtain rewards for the period. Only in the first period producers had an endowment of three units of output and some cash (together with consumers) to let the experiment start. There is a financial sector: agents can borrow money by selling bonds. Two types of bonds exist paying a fixed number of francs with certainty at specific dates, and return the capital at different dates after selling. The financial sector is aimed at reducing possible binding effects deriving from the cash in advance nature of the economy. The experiment consists of a series of trading periods; all markets are organized as double auctions. At the end of the experiment, all fiat money was converted into real values (X and Y) using the average price during the final period. This procedure was implemented to avoid the backward induction argument that makes the value of money equal to zero in every period, therefore refraining from trading.

[Figure 1]

This economy is a quite complex system, nonetheless it does not show disequilibria, cycles, instability and coordination failures. Efficiency measures reveal that the median efficiency is equal to 88.9% per period, and production efficiency is virtually equal to 100%. This means that substantial gains from trade are achieved. Furthermore, the variability of allocations (consumption, production and inventories) decreases over time, together with the variability of the price ratio. This finding does not mean that the competitive equilibrium model is accepted, although convergence toward its predictions is uncovered in a weak sense. When moving the analysis towards the partial equilibrium models, both the input and the output markets are in disequilibrium, with volumes that are lower than the theoretical demand and supply quantities. However, a movement towards the partial equilibrium is

found in both cases. In addition, given the price ratio, too much Y is consumed by consumers, producers under-produce X and over-consume Y .

So far we have underlined the microeconomic properties of the economy. Macroeconomic variables can of course be analyzed too, in particular those which are concerned with inflation and unemployment. With a constant money supply the price level increases and approaches asymptotes, while the inflation rate approaches zero. The constant level of money supply has a positive effect on nominal variables but no effect on real variables. The Okun's law – a negative relationship between changes in unemployment and the percentage change in real GNP - is strongly observed in the data, while the same support is not found for the Philips curve. The authors claim that an overall confirmation of the neoclassical model is found, and poignantly ask skeptics to be able to analytically compute the competitive equilibrium solution from the set of given parameters. They observe that this task is rather difficult without the help of the invisible hand of the market that was in place during the experiment.⁵

3. Monetary economics

Monetary economics has certainly been the most studied field of macroeconomics in experimental economics, and besides interesting results it has also been one of the most successful in terms of methodological contributions.

3.1 The role of money

Money is valuable because it is a “medium of exchange”, and people should hold it because of this reason. However, people will hold it only if it has value, i.e., money as a “store of value”. McCabe (1989) provides the first experiment on this topic, in which money acts in

these two ways. Players trade fiat experimental money for non-durable goods over six periods. They can propose to buy or to sell these goods at a unit price, and the market institution in this experiment is a clearing house that randomly assigns buyers to sellers. This institution with fixed prices relates this experiment to macroeconomic models of disequilibrium in which the adjustment process occurs through quantity rather than prices (Benassy, 1982). Usually not all proposed trades are satisfied. At the end of the six repetitions money is valueless, therefore nobody would like to hold it, but if anybody wants to hold it in the sixth period, this will also be true in the fifth. A backward induction argument applies here, and money will not be accepted in any period.

Figure 2 shows that the Nash equilibrium of no trade is rejected: in two out of nine experiments trading ceased completely whilst in the others it continued at a reduced level. Trade breakout is reached only when players with previous experience of this experiment were recruited. The conclusion drawn on the dual nature of money is that individuals use it as a medium of exchange even though it is an unstable store of value.

[Figure 2]

Whilst the previous experiment does not aim to test any particular theory, in a pair of experiments the Kiyotaki and Wright (1989) model is tested. In this model money arises as a medium of exchange because of trading frictions existing in the economy. The framework is the following: there are equal numbers of three types of agents (type 1, 2, and 3); each of them produces a good that he does not want to consume (agents of type 1 want good A but produce good B , agents of type 2 want good B but produce good C , and agents of type 3 want good C but produce good A). In every period agents are randomly paired and have to decide whether to trade their good or to keep it, entering in the next period with the same good. In this case they have to pay a storage cost, which is different across goods (in particular, $C_A < C_B < C_C$). In case of trading, the agent will consume that good at the

beginning of the next period, receiving the related utility, and produce the next good. Individuals are infinitely lived. Individual payoffs depend on the good held at the end of the period and on the incurred storage cost. In the experiment there is not a double coincidence of wants, and some agents have to act as a middleman, buying a good that they do not want to consume in hope of forming a double coincidence in the future. Thus, at least one commodity has to emerge as a medium of exchange. Two pure strategy equilibria can arise: a fundamental equilibrium where agents always prefer lower storage cost good to one with higher storage cost. This equilibrium is characterized by type 1 and 3 never trading for any good other than their own consumption good. The type 2 agents trade with type 3 for good *A* and then trade it with type 3 for good *B*. Good *A* is the medium of exchange because of its low storage cost. The other equilibrium is speculative and occurs when the storage cost of good *C* is not “that much” higher than the cost of storing good *B*. Besides the previous medium of exchange, another one emerges (good *C*): type 1 trades with type 2 for good *C* in order to exchange it with type 3 for good *A*.

The main reason to run experiments on this model is that it is interested only in the equilibrium predictions, it is silent on the path towards this equilibrium. Brown (1996) implements an experimental version of the model that has speculative equilibria only. Subjects are randomly assigned their types with the respective production abilities and consumption needs. At the beginning of each period players are randomly assigned a trading partner and have the opportunity to exchange their goods according to their decisions, which are taken before being paired. Decisions consistent with the speculative strategy are taken by type 2 and 3 subjects in 99% of cases, and in roughly 30% of cases by type 1. These choices are fairly consistent with profit maximization given that only 4 out of 36 players would have earned higher incomes using the speculative strategy subject to their partners choices. However, the average income was only 80% of the income resulting from the speculative

strategy. This effect was mainly due to individual parameters. Furthermore, average losses tended to increase over time. Overall results support the idea that subjects trade for goods that do not give them immediate utility, but the support to the speculative strategy is mixed given the failure of type 1 players to implement it.

Duffy and Ochs (1999) improve on Brown's experiment in three ways in order to stay closer to the Kiyotaki-Wright environment. Firstly, they induce risk-neutrality paying subjects according to a binary lottery.⁶ Secondly, an infinite horizon environment with a discount factor is implemented (at the end of each round there is a 10% probability of stopping the game). Thirdly, the common knowledge assumption is operationalized, informing subjects of the historical average proportions of goods held by each player type in the population. The last two changes are meant to improve coordination in the speculative strategies.⁷

*3.2 Inflation, expectations and coordination*⁸

This is a series of experiments that started from the problem of finding stationary solutions from overlapping generation models with fiat money. Follow-up experiments have enriched the environment and asked different and related questions regarding expectations formation, inflation processes, and coordination amongst agents. It has also led to the birth of a technology to implement overlapping generation experiments and has contributed to the debate on the performance of different market institutions. We begin by reviewing the problems and related experimental results, then look at the procedures.

The first experiment in this wave (Lim *et al.*, 1994) is linked to the topic of the previous Subsection and is based on the notion that in the overlapping generations (OLG) model with fiat money there is an infinite number of competitive equilibria, only one of which is stationary. In this economy young members are endowed with seven units of

“chips” and a given amount of money (“francs”), when old they receive a further unit of chips. Members of the two generations are allowed to trade francs for chips, and at the end of each period old individuals calculate their reward by multiplying the number of chips held at the end of the first period and at the end of the second period.⁹ In all the economies the price of chips in terms of francs is substantially different from the competitive stationary equilibrium in the first period, but as long as repetitions take place a marked convergence towards the stationary solution is found (except in one out of the four economies). Furthermore, even though there is not a complete emergence of the stationary equilibrium, in no cases are there any patterns consistent with demonetization of the economy, that is, where the trading volume is zero. The observed trade volume is over time closer to the range defined by the competitive stationary equilibrium and the Nash stationary equilibrium (Figure 3). The efficiency of these economies is remarkable: it is very close to 100% of the competitive equilibrium, compared with roughly 70% of the demonetized economy.

[Figure 3]

Following experiments kept the same basic structure and compared the behavior of different policy regimes, with a focus on the emergence of rational expectations equilibria. Marimon and Sunder (1993) examined a monetary regime characterized by a constant level of deficit financed through seignorage. In each period the experimenter buys a number of chips (equal to the number of playing agents multiplied by per-capita government deficit) at the market-clearing price thereby injecting money into the economy.⁶ The model predicts both a low and a high inflation steady state (ISS), and any initial inflation above the low ISS brings the system towards the high ISS. This pattern is not detected, and economies tend to converge to the low ISS. In terms of macroeconomic theory, this result may show that high inflation patterns that are found in many models may not be an accurate finding in economies with real agents. Furthermore, this is consistent with the idea that it is possible to

reduce the inflation level by reducing the seignorage. The rational expectations path has very little power in explaining the data, and the clustering of data around this equilibrium is consistent with adaptive learning as a behavioral assumption, because prediction errors show a positive dependence on past inflation. However, the authors do not investigate which learning rule seems to fit the data better.

This issue is subsequently studied by Marimon and Sunder (1994) together with the analysis of other policy regimes: an inflation targeting regime in which the deficit level is adjusted, and announcements of changes in the deficit level, which mainly consists in an increase in government expenditure. If all agents share the same belief about a constant rate of inflation, the inflation target is achieved in one period, and nonstationary rational expectations equilibria tend to the solution with no value of money and zero deficit. Experimental results differ from the rational expectations path and show more volatility than the simple learning algorithms. Inflation rates tend to their targets, but deficit levels are larger than predicted. A possible explanation hinges upon the endogenous uncertainty that is part of these economies due to the interaction of a small number of agents, which is not part of the deterministic theoretical model. Savings are increased to counteract this uncertainty, and because deficits are proportional to savings, they are higher because of increased uncertainty. In an economy in which a pre-announced change in government expenditure occurs, under the rational expectation hypothesis some form of anticipation should be observed along the equilibrium path. The unique change was announced in period zero. A change in both prices and volumes takes place before and after this change, and the least square pattern appears to be a fair explanation of the data, whilst the long anticipation effects postulated in nonstationary rational expectations equilibria path are never observed.

Marimon and Sunder (1995) aim at testing whether simple policy rules promote economic stability because they can be easily learnt. They compare the behavior of an OLG

economy in which the government finances a fixed level of real deficit through seignorage (deficit rule) with an economy that allows money supply to grow at a predetermined rate whilst the level of government expenditure changes accordingly (money rule). The second regime is very similar to the Friedman's prescription to stabilize inflation rates. Once a k -percent money rule is announced, for a broad class of learning rules, agents' beliefs on inflation are expected to converge to the announced rate. If agents believe that the announced money growth rate will be the realized inflation rate, then the realized inflation rate is the announced money growth rate: the long-run policy objective is achieved through short-run aggregation of agents' beliefs. Results do not support the superiority of the money rule over the deficit rule: under both regimes persistent fluctuations occur, and much of the price volatility is correlated with the parameter instability of the economy, regardless of the more stable policy rule. In an economy in which the deficit rule is changed in favor of the money rule, a dump in inflation level is observed.

When implementing an experimental version of the OLG model, one faces two main problems. First, in such models there is an infinite number of agents, and due to laboratory size and money constraints this condition cannot be met. In this environment, N agents are recruited and n of them played the role of the young generation, n played the role of the old generation, while the remaining $(N - 2n > n)$ were outside the game. At the beginning of each period, n of the $(N - 2n)$ players that were outside in the previous one are randomly selected to enter into the game. This procedure makes sure that each subject stays out of the game after exit and before re-entering, and avoids the possibility of playing a supergame. A long series of trades is played, and when parents die in the n period, they are reborn as children in the $n + 1$ period. The game is stopped just once, after a long series of games. Another procedure has been independently developed by Cadsby and Frank (1991) and it is discussed in Section 4. Second, the infinite horizon of the model needed to not end up in the

outcome of no-trade when fiat money is involved cannot be implemented in the laboratory; therefore it is necessary to construct a terminal condition that does not affect the strategies and the outcomes available to agents. A solution is conceived by Lim *et al.* (1994) and subsequently applied with little modifications:¹⁰ subjects temporarily outside the market play a forecasting game. At the beginning of each period they conjecture the market-clearing price for the current period, and the best forecaster(s) is (are) rewarded accordingly. Without any previous announcement, and after forecasts for period $T + 1$ have been submitted, the experimenter announces that period T is the last of the game. At this point money holdings of agents who entered in the economy in period T are converted into chips using the average predicted market price for period $T + 1$ by outside market participants.

In Lim *et al.* two trading rules are implemented. Firstly, a single unit double auction with the provision that the last transaction of an old subject in any period could be exchanged for a fractional unit that enables him to use all his cash. This institution was time consuming, led to many computational mistakes by subjects that caused unintended inflows and outflows of money from the economy, and also, led some agents not to exchange all their cash with consumption, causing inefficient outcomes. Under the second rule, market clearing prices and allocations were computed from the supply function that was solicited from the members of the “young” generation, and then summed up. A different scheme has been applied by Marimon and Sunder (1995): “young” agents were asked to submit an inflation forecast at the beginning of period t , and this forecast is used by the computer to calculate the corresponding optimal consumption. Individual supplies are summed up across young agents to obtain the aggregate supply for that period. In both cases buyers were passive.¹¹

3.3 Dynamic inconsistency of monetary policy

A new area of research is concerned with reputational models of monetary policy. They show that when the policy maker cannot commit to an inflation policy, optimal equilibria (Ramsey) are time-inconsistent, while sub-optimal ones (Nash) are time-consistent (Kydland and Prescott, 1977; Barro and Gordon, 1983). Therefore there is an inflation bias. Arifovic and Sargent (2001) provide an experimental framework that links expectations and decisions made by policy makers and the public through a Phillips Curve that relates actual and expected inflation with unemployment, in a world where monetary policy authorities have an incomplete ability to set the inflation rate.

A group of $N + 1$ subjects make up the economy. N forms the public; they have to forecast the inflation rate for each period of the experiment. Call agent i 's forecast x'_{it} and let \bar{x}'_t be the average of the citizens' forecasts. Citizens receive payoffs that rise as their session-average squared forecast errors fall. Agent i 's payoff at the end of time period t is given by $-0.5 (y_t - x_{i,t})^2$. The remaining agent is the policy maker. In each period he sets a target inflation rate, x_t , then a random number generator sets v_{2t} and the actual inflation rate equals $y_t = x_t + v_{2t}$. Unemployment is then generated by the Phillips curve. Agent $N + 1$'s payoff is given by $-0.5 (U^2_t + y^2_t)$. A random stopping rule is used to implement an infinite horizon and to discount future payoffs with a discount factor d between zero and one. At the end of each period, the computer program draw a random number from a uniform distribution over $(0,1)$. If this random number were less than d , the experimental session would continue for one more period, otherwise the session would be terminated. An upper bound on the duration of a session is set at 100 time periods. Each subject received a \$10 payment for completing a two-hour experiment. They could also earn a prize of an additional \$10, according to a variation of the Roth and Malouf (1979) binary lottery to control for risk attitude.

Subjects show a remarkable attitude in forecasting actual inflation (Figure 4). Results are fairly consistent with the theory. In nine out of twelve experiments the outcome is close to the Ramsey equilibrium, although four economies experience backsliding to the non-optimal equilibrium after having achieved the optimal one. Transitions from Nash to Ramsey equilibria are slow and do not cause dramatic increases in unemployment. On average, inflation estimates by the public are good and do not contain systematic errors. Individuals put more weight on the recent past when forming expectations concerning the inflation rate.

[Figure 4]

4. Fiscal policy

The second largest area in experimental macroeconomics is represented by fiscal policy. In particular, Ricardian Equivalence has received considerable attention. According to Barro (1974), given a pattern of government expenditure, a reduction in taxes today implies their rise in the future. Therefore, provided that parents care for the well-being of their descendants, they will save the tax cut to enable them to repay the debt the government has to incur. Cadsby and Frank (1991) developed an overlapping generations model independently of Lim *et al.* (1991) designing an environment in which two groups play the game for eight years, each year is made of three periods, and the two groups overlap in the medium period. In contrast to the previous setting the game is stopped and restarted every time a pair of generations has played, and in each game subjects enter only once. In the first period, the current generation has to allocate a given endowment between certificates and savings. In the second period a further endowment, which represents government deficit, may be given to them. Then they decide the allocation between certificates and savings. In this case savings represent the bequest left to the future generation. The future generation receives an endowment and the bequest, and allocates this sum between certificates and

savings. An amount equal to the second endowment given to the first generation is then subtracted from their resources. In the third period the second generation may only buy certificates. The experiments examined both expansionary and contractionary fiscal policies. Parents' utility function includes first and second period consumption, together with the utility level attained by descendants, which in turn is given by their second and third period consumption. The utility function is multiplicative in the arguments, which entails a high degree of consumption smoothing. If a descendant is left with a poor bequest, he may not be able to repay the debt, and therefore, by convention consume zero, which becomes the utility level of the parent.¹² Whenever the theory predicted a positive bequest, outcomes close to those predicted by Ricardian equivalence occurred, with some allowance for learning. Moreover, deviations are not unbiased as they display a tendency towards Keynesian behavior.

The Ricardian theorem requires some strong assumptions; subsequent experiments have focused on their violations to test for the robustness of the theory. Slate *et al.* (1995) build their experiment on Cadsby and Frank and aim at testing whether uncertainty of debt retirement, that is, the probability of repaying the debt is less than one, affects bequest decisions. When this probability is low, consumption by the current generation increases, as predicted by the Keynesian theory. However, when the probability of debt retirement increases, bequests rise to offset the future generations' expected repayment liability and deficit spending becomes much less expansionary.

Ricciuti and Di Laurea (2003) look at two different possible departures from Ricardian equivalence: liquidity constraints and parents' uncertainty of their future income. In both cases the theory predicts that agents engage in too much early consumption and bequeath an amount of money that is unable to offset the debt repayment that descendants have to make. The authors construct three treatments: the first one resembles Cadsby and

Frank except for the fact that they use a matching rule similar to the one used by Sunder and his associates to avoid a supergame effect, and therefore, possible backward induction. In the liquidity constraint treatment parents face increasing income in the two periods in which they live; whilst in the uncertainty treatment they do not know at the beginning of each game the income they will have in the second period. To implement this difference between the two periods the disposable and extra income (which Cadsby and Frank and Slate *et al.* gave respectively at the beginning of the first period) is given in two installments. For the sake of comparison this procedure is also applied in the baseline. Results for the baseline treatment are substantially in line with the two previous experiments although there are a larger number of parameters, and therefore, more computational difficulties. Results for the two departures do not completely confirm the theoretical claims: in the credit-constraint case subjects do not equate consumption over time and do leave positive bequests. This does not happen in the uncertainty case.

An interesting experiment founded by the Dutch Ministry of Social Affairs and Employment under the auspices of the Dutch Second Chamber of Parliament on the effects of the wage tax on budget deficits and unemployment in an open economy is made by Riedl and Van Winden (2001). Unemployment benefits are usually funded via a tax on labor, which may have a negative effect on the working of the economy. The economic system consists of two countries, a small “home” country and a large “foreign” one. Four goods are traded: two input goods (capital and labor) and two consumption goods called X and Y (output). In each country there are two types of agents, consumers and producers.¹³ Consumers wish to consume the two outputs and enjoy leisure. In each trading period they are endowed with some units of labor and capital but none of X and Y . They can sell their endowments on the input market to producers for fiat money and buy the outputs from producers with the proceeds. Consumers have an additional source of money through

unemployment benefits. For each unsold unit of labor (which counts as leisure) they receive a given benefit. Consumers' real money payoffs are determined by their consumption of X , Y , and leisure. On the input market producers buy capital and labor to produce output, therefore X and Y are produced and sold on the market. Producers' payoffs are determined by the profit they earn. The labor market is local, consumers can sell their labor only to producers in their own country. In contrast, the capital market is international; the Y -market is local, and the X -market is international. The economy consists of a sequence of trading periods: at the beginning of phase one consumers receive their endowments and, together with producers, they receive some cash to allow for trading. After labor and capital are traded, production takes place automatically, and agents are allowed to trade X and Y via a double auction, then the period ends. In this stylized economy government expenditure is represented by the unemployment benefit, and the revenue side is represented by a tax on employed labor levied on producers needed to fund the latter. There are two treatments concerning the government budget. In the first, a "constant tax regime" is obtained fixing the wage tax at a rate that ensures a balanced budget, according to the general equilibrium solution. In the second, called "dynamic tax regime", taxes can be adjusted in period $t + 1$ after a deficit in period t . The first treatment allows the economy to stabilize and to see whether this happens running a deficit or a budget surplus. The dynamic treatment allows for an assessment of what happens to the economy when all parameters but the tax wage are held constant.

For the constant tax regime in both countries a large and persistent budget deficit is observed, which does not vanish over time, whilst unemployment converges to its equilibrium value from above. Nominal wages appear to be too low for a balanced budget. The reason for this is twofold. On the one hand consumers tend to supply too much labor, therefore reducing its nominal value. On the other hand producers are reluctant to employ

labor. The authors believe that the former effect has its roots in the risk-compensated price-mechanism: producers face uncertainty about output prices and consequently their revenues. Together with risk neutrality, this can explain why they use fewer resources (in particular labor) than the optimal level. This also brings unemployment to a higher level. In the dynamic tax regime there is still a tendency to run a budget deficit, even though they are small and tend to vanish over time in both countries. The cost associated with a close to balance government budget is high in terms of unemployment: this rate increases from 6% to 12% in the small economy and from 4% to 18% in the large economy. There is also a sharp decrease in the long-term real GDP. The negative effect of a wage tax is therefore confirmed by this experiment.¹⁴

Bernasconi et al. (2003) provide an interesting method to test for anti-Keynesian effects of fiscal policy. According to the “expectation view” on fiscal policy, government expenditure cuts may have expansionary effects on private consumption in contrast with the standard Keynesian view. The mechanism at work here is based on the expectations of the future tax burden. Bertola and Drazen (1993) and Sutherland (1997) develop two models in which, with rational expectations, a simple optimizing consumers’ behavior may give rise to some anti-Keynesian effects when government expenditure and debt reach some unspecified critical levels. A restrictive policy at this point may have an expansionary effect.

Once again the problem with these models is related with the unobservability of expectations. The lab is therefore well suited to induce agents to express their views about future values of the variables of interest. However, generating *ad hoc* variables, and telling them that fiscal series follow a Brownian motion would be too demanding. The problem is circumvent exposing players to real data on public debt; changes in public debt, taxes and government expenditure from fifteen European countries without giving them references on the countries and the period considered. Then players are asked to submit forecasts of the

subsequent values of taxes and government expenditure. Subjects derive their utility from consumption of two subsequent periods, and optimal consumption is determined by the computer from the above forecasts. The attained utility level is converted in per minute wage for each participant. Saliency is warranted by the circumstance that the wage is strongly increasing as long as the forecast is closer to the real values.

Agents' expectations are found neither to be consistent with rational nor with purely adaptive expectations. Expectations follow an augmented-adaptive scheme, which embodies the 'spend and tax hypothesis' on the relationship between taxes and expenditure. This is a short-run causal relationship that holds regardless of the actual causal relationship between taxes and expenditure in the field. These results are consistent with the anti-Keynesian view of fiscal policy, but data reject possible non-linearities of its effects.

5. International trade and foreign exchange rates

A small area of research is concerned with international economics, in particular international trade and foreign exchange rates. In this Section we return to a group of general equilibrium experiments. Noussair *et al.* (1995) provide the first experimental framework to test for competing theories on trade, in particular they look at patterns of trade and output predicted by the law of comparative advantages. According to this theory, in a two-country, two-good model with different production functions, if a country has an advantage in the production of a certain good it will specialize on its production, and will export it. Similarly, the other country will specialize in the other good, and the prices of the final goods will be equalized. In contrast, the autarchy model predicts no specialization and trade, with a convergence to the domestic competitive equilibrium levels of the economic variables.

The authors construct two environments, for the sake of brevity we only consider the first one, which is a version of the Ricardian model. There are two output goods, Y and Z , and one input L , and two types of agents, consumers and producers. Consumers own factors of production and have induced preferences over the two consumption goods. Producers have initial endowments of input, and earn profits buying L and selling Y and Z . Half of the agents belong to a country, half to another, with an equal number of consumers and producers in both countries. The two countries differ in their production functions, but have the same currency. Consumers sell their endowment of L to producers of their own country (the factor of production is not mobile between countries), and then buy production goods from either country. Consumers get utility from consumption and any profit made by price speculation. Producers buy L in their own country and produce Y and Z for consumers in either country, getting utility from profits attained in the production and market activity. In some experiments a tariff on Z is imposed in order to mimic the effect of transportation costs. Therefore the experiment includes six markets that are organized according to a double auction.

An important contribution of this paper is on the empirical side of its analysis. To test whether the data converges over time to the equilibrium predicted by the theory, usual regression models are not good because data exhibits serial correlation and heteroscedasticity. The estimated model is the following:

$$z_{it} = B_{11}D_1\left(\frac{1}{t}\right) + \dots + B_{1k}D_k\left(\frac{1}{t}\right) + B_2\left(\frac{t-1}{t}\right) + u_{it}, \quad (1)$$

where i indicates the particular experiment, t represents time measured by the number of market periods in the experiment, D_i is a dummy variable that takes value 1 for i and 0

otherwise, and B_{1i} is the origin of the possible convergence process. B_2 is the asymptote of the dependent variable, as t gets larger the weight of B_{1i} becomes smaller because $1/t$ approaches 0 while B_2 is larger $(t - 1)/t$ approaches 1. The model is used to test the hypothesis that the data converges to the predictions of various models by testing whether or not the estimates of B_2 are significantly different from the predictions of the models (strong convergence). In contrast, if the B_2 term is closer to the model's prediction than B_{1i} is, data is partially converging.

Data generated by the experiment generally supports neither the competitive equilibrium nor the autarchy model in the point estimates (which include prices, production, and net exports), whilst at the same time, the law of comparative advantages predicts trade patterns (Figure 5 reports data for output price). However, convergence towards the free-trade competitive model predictions is found for aggregate production and individual consumption patterns, as well as, output and factor prices. As in the competitive model, tariffs reduce international trade and market efficiency. Overall the results lend support to the competitive equilibrium model.

[Figure 5]

In a companion paper Noussair *et al.* (1997) aim at testing the law of one price and the flow of funds theory. The framework closely resembles the previous one with the addition of two currencies that enter into the utility function of the agents of the home country. In the absence of tariffs, taxes, transportation costs, and other frictions, the law of one price maintains that the prices of the two goods will be the same in both countries after prices are factored by exchange rates.

The results support the competitive model over the autarchy one in a number of ways. Data on exchange rates strongly supports convergence to the competitive model predictions, whilst those on the volume of exchange show a bias towards a level lower than

the one predicted by this model. Prices in both countries partially converge to those predicted by the competitive model, as well as prices for X and Y . The flow of funds are moving towards the competitive prediction, the law of one price is only supported in the market for Y , whilst the purchasing power parity theory is not supported by the data. Failure of the law of one price is attributed to the asymmetry in the adjustment speed of the price discovery process of local markets. In addition, the flow of funds theory determines the supply and demand of currency and therefore the behavior of the exchange rate from one period to another.¹⁵

Arifovic (1996), in turn, provides an experiment exclusively designed to address exchange rate fluctuations. An interesting feature of this work is that it combines results from experiments with those obtained from simulations. Subjects were divided in two equal groups, one of which was young in the odd period and old in the even, and the other vice versa. They received a high endowment (consumption good) when young and a low endowment when old. The experimenter at his will decided the termination of the game, unknown in advance to the subjects. In the high endowment period agents have to decide the quantity to sell in the market, and each young person was asked how many units of the total offer he wanted to sell in the market for currency one. The remaining part was sold in currency two. The difference between the agents' endowment and the quantity offered for sale represented consumption for that period. In the low endowment periods subjects have to use all their inventories of currency one and two to purchase the good, such that, this quantity plus the low endowment represent the consumption for that period.

The exchange rate fluctuated between 0.5 and 2 over time. The fluctuation was driven by changes in agents' portfolio decisions on how much to save in each currency. Fluctuations showed a unit root, a clear sign of shocks persistence. The median of first-period consumption over time goes to the perfect-forecast stationary equilibrium value.

6. A growth experiment

A new research topic is represented by experimental research on economic growth models. The idea of these experiments is not to replicate any real economy, but to compare the numerical predictions of the models with the observed data. Lei and Noussair (2002) analyze the exogenous optimal growth model based on Cass (1965) and Koopmans (1965), in which the level of investment is endogenized in an economy where a representative agent makes optimal consumption and investment decisions over time for a given technology. If production and utility functions are concave, there is a unique optimal steady-state level of consumption and capital stock.

Two main different treatments are implemented. In the *social planner* treatment, each agent represents a single economy, which has to choose between consumption today and investment for future consumption in the future. This treatment is closer to the literal formulation of the model. In the *market treatment* each economy includes five heterogeneous agents that are allowed to trade their capital good through a continuous double auction. This second treatment has been added because of the properties of this market institution to achieve efficiency. Under each treatment two cases are considered: the low and high endowment, that are situations in which the endowment is lower or higher than the equilibrium level of capital. The model predicts that in the first case convergence occurs from below, whilst in the second it is achieved from above. Under the market treatment each agent has his own production function and an individual utility function, which indicates the number of experimental currency units the agent can get when he consumes the good. The overall amount of experimental money is converted into dollars at a given exchange rate at the end of the experiment. The individual and aggregate production and the utility functions

are concave. In each period a market for capital takes place: agents can make ask or bids for multiple units of capital at a named per-unit price. At any time buyers or sellers may accept offers made by another agent, or a part of an offer. To achieve aggregate efficiency capital must go from low- to high-productivity agents. To allow trading, each agent has an endowment of capital and another endowment of money that decreases as long as units of capital are bought, and increases when they are sold. The infinite horizon of the model is obtained rolling a 20-sided dice at the end of each period. If the dice shows numbers 1 or 2 the experiment is stopped, therefore there is a 10% probability in each period to stop playing. All the features of the market treatment but trading are saved in the social planner treatment. Each subject is an independent economy.

Using the linear regression method outlined in the previous Section, in both treatments consumption, capital stock, the price of capital and the realized levels of consumption converge to the optimal steady-state levels predicted by the theory, after a few initial periods. Convergence to the equilibrium is faster and stronger under the market treatment than in the social planner treatment, showing that the price mechanism helps agents at making intertemporal choices. There are no significant differences between the low- and high- endowment treatments.¹⁶

7. Macroeconomic imperfections

In this Section we discuss three experiments that focus on a specific issue. The first one concerns credit constraints in a general equilibrium model (in a sense it belongs to both types of macro experiments since it analyzes a single issue within a simple system of related markets), the other two fit in the second category of experiments and analyze a single market/issue.

In the Bosch-Domenech and Silvestre (1997) experiments, consumers operate in a single period with three goods: labor (input), output, and a non-productive good. There are two sources of income: current income obtained by selling labor, and future income obtained from selling some of the non-productive good. In each period consumers sell some units of labor and buy some units of output. This purchase may be partially financed by borrowing against future wealth. Any credit must be repaid at the end of the period out of non-labor wealth. There is a limit (a “credit line”) to the amount that can be borrowed, which is a fraction of the non-labor wealth. This parameter is varied across experiments to create environments with high- and low-credit.¹⁷

The results show that prices are rather sensitive to credit constraints: as long as the credit limit is close to zero, input and output quantities and prices are close to zero and sharply increase when the limit is increased up to a certain level. After this level prices become independent of the credit limit. Transactions are close to productive efficiency, and the average number of transactions vary significantly at lower levels of the credit limit, while is not sensitive to it when large enough. These results lead us to the conclusion that when credit constraints are in place they have strong effects on the economy, and a recession that reduces the amount of credit in the economy determines a fall in economic activity. These results are also consistent with a modified version of the Arrow-Debreu general equilibrium model.

Hey and Di Cagno (1998) build an experiment on the dual-decision hypothesis developed by Clower (1965). In the economy workers firstly sell their labor to firms in markets, then sequentially buy goods from firms with the money earned in the first transaction. In this dual decision problem unemployment may occur even in competitive markets. The experiment considers a pure exchange economy with two groups of agents of equal size, workers and firms. The latter are endowed with labor (X) and not goods (Y), the

former otherwise. Both agents derive their utility from a Cobb-Douglas function in both arguments. The markets for X and Y open sequentially, first the labor, then the goods; both markets are organized as double auctions. The first result to point out is that the volume of trading is lower than the expected competitive equilibrium: subjects in group one tend to hold more X and less Y than predicted (the reverse is true for subjects in group two, in contrast with the theory). Furthermore, there is no evidence of convergence to the competitive equilibrium over time. This result is seen as a confirmation of the model where competitive outcomes do not arise and resources are under-employed. Individual performance in terms of actual to theoretical earnings appears to be low, at a 75% level. The model does not make predictions on absolute prices but on relative ones. The result is also disappointing in that: relative prices tend to stay at unity, which is lower than the optimal level, because of a focal point effect. The stock of money does not seem to affect absolute prices. The message from this experiment is the possible inefficiency of the double auction. However, in this experiment the two double auctions do not resemble the typical features of this institution: there is neither a redemption- nor a cost-schedule for buyers and sellers in the two markets, therefore the informational content that agents have to manage is extremely limited to device a strategy. It appears fruitful to address the same issue with a more common double auction structure to assess the robustness of this result, which is of some relevance given the observed widespread use of the double auction mechanism in macroeconomic experiments.

The effects of a negative shock in monetary policy on the real economy constitute a long-standing debate in macroeconomics. Several explanations have been offered from menu costs to staggered contracts, from near rationality to informational frictions. So far, two experiments have addressed this issue in quite different ways. Wilson (1998) provides an experimental analysis of the Mankiw (1985) model of menu costs and business cycles. After

a tightening in monetary policy, a monopolist has to incur a small fixed cost to adjust his price. Whilst it may be individually rational not to change the price if the loss from doing so is smaller than the fixed cost, this decision may have large effects on social welfare. In the experiment a posted offer institution is used, each monopolist has a constant marginal cost of production, faces simulated buyers, and only knows the demand function.

The experiment has three treatments, which include a baseline without any menu cost, a static menu cost, and a dynamic menu cost decision problem. Each treatment has been run under two information settings: in the first, subjects were not informed about the direction of the demand shock, whilst they do know it in the second setting. The market institution is a posted offer. In the absence of menu cost the theoretical level of output remains unchanged with a full adjustment of the real wage to the shock. When static menu costs are present, each monopolist can either choose the theoretical monopoly price from the first twelve periods, or change the price and incur the cost. Each period decision to change the price or not only affects the current period profit and not future profits. In the dynamic cost treatment each decision taken in one period affects future profits. Statistical analysis of the data reveals that variables indicating menu costs treatments are significantly different from zero, therefore prices are not perfectly flexible. At the same time one can reject the hypothesis that they are equal to the difference between the post-shock price and the initial monopoly price, indicating a certain adjustment. A further analysis carried out constructing 95% confidence intervals for mean response shows that they collapse around the monopoly price as time passes in each demand regime and treatment.

The second experiment in this field considers a source of non-neutrality of money that has been overlooked, and in fact is not considered in standard analysis of business cycles: money illusion. It is defined as a violation of the homogeneity postulate, i.e. supply and demand functions are homogeneous of degree zero in all nominal prices, so they only

depend on relative and not on absolute prices. The typical first-year textbook example is that when prices and income double, the optimal choice is unchanged. Fehr and Tyran (2001) provide an experimental framework to test for money illusion. They consider an n -player pricing game with strategic complementarity and a unique equilibrium. There are two types of players: x -type have to choose lower prices with respect to y -type. After T periods a fully anticipated negative shock on money supply occurs and the game continues for another T periods. Subjects are endowed with payoff matrices that are changed after the shock hit the economy. To see whether an exogenous and fully anticipated monetary shock creates money illusion, the experiment should be framed in nominal and real terms. However, there are two ways in which individuals can make this kind of mistake. A direct way is that the individual decision-maker is unable to fully adjust for the shock. An indirect way occurs when the decision-maker understands that his optimal choice should not change, but believes that other agents are unable to completely adjust, therefore he can change his choice to take advantage of this situation. To cope with these two sources and disentangle their effects, the two framed problems receive two different treatments. In the first, all subjects are human and therefore both sources are possible and may reinforce themselves, in particular the coordination problem based on expectations about other players' behavior is in place. In the second treatment $n - 1$ players are computerized, and the human agent knows that they only play optimal responses, so there is no coordination problem.

Results show the relevance of money illusion in explaining nominal inertia. The nominal treatment with human subjects clearly shows sizable price inertia, whilst in the other treatments this is much lower. In the computerized treatments all subjects in the real frame instantaneously adjusted to the post shock equilibrium, nominal inertia was completely absent. In the nominal frame not all subjects (about 20%) adjusted to the new equilibrium, and though a tendency over time towards equilibrium prices is observed, they are never

actually observed, showing long-lasting nominal inertia. In the all-human treatments, the real framework shows an initial amount of nominal inertia that leads prices above the equilibrium, but it vanishes relatively quickly bringing prices quite close to the equilibrium. Price stickiness in the nominal case is very large: only 11.5% of agents play the equilibrium price after shock, whilst they amounted to 80% in the last period before the shock. Furthermore, 73% of subjects named prices three or more price units above the equilibrium. Compare these figures with 35% and 93% respectively and note that 23% named two or less unit prices above the equilibrium. A formal test rejects the null hypothesis that prices are equal in the two treatments at the 2% level for the first nine periods after the shock, and at the 10% level in the three remaining periods. Taken together the results of these experiments suggest that the individual money illusion plays a small role in explaining nominal inertia, and that the coordination problem exacerbates its relevance. Furthermore, a human treatment in the real and nominal frame with a positive monetary shock is ran. It shows that price stickiness is much lower, perhaps bringing evidence to the Keynes' idea that price flexibility is asymmetric: it works mainly upwards.

8. Conclusions

In this paper we have extensively reviewed the applications of the experimental method to macroeconomics. Although it is based on a relatively limited number of papers, with respect to other more established fields in experimental economics, experimental macroeconomics covers a large number of macroeconomic issues and it is characterized by an increasing number of papers.

Ten years ago Friedman and Sunder (1994) wished that 'perhaps macroeconomics too, like meteorology and astronomy, will become an indirectly experimental science, one

that relies on experimentally verified results in constructing his central theories, although the central theories themselves are not amendable to direct experimentation.” We think that the works discussed in this paper show that ten years later macroeconomics has become a field in which experimentation is viable and able to offer specific insights. However, to some extent their hope has not been fulfilled: although experiments are able to test specific theories, the results of these experiments have not brought new ideas and hypotheses in the making of macro models. Several factors may underlie these circumstances. First, experimental macroeconomics has not yet reached a critical mass that makes it a reliable source of inspiration for macroeconomic modelers. Second, most of the experimental work surveyed here offers a substantial confirmation of the theoretical models, therefore, apparently there is no need to build new models based on experimental results. Third, the feedback from experiments to theory - even in a success field such as microeconomics - has proven to be difficult: for example, poor experimental evidence of the expected utility model has not been matched by a new convincing model of decision under uncertainty that has been able to take over the incumbent.

An interesting area for future work is the use of laboratory experiments for policy experiments, as it has already been done in designing markets for regulated sectors (e.g., water and electricity) and for comparing auctions rules (see the design of auctions for allocating the spectrum between mobile phone companies). In some sense we are all unintentional subjects of real world macroeconomic experiments that may cause economic disasters because of their unforeseen consequences (for example on work incentives, or on the price level). The experiment of Riedl and van Winden (2001) designed to advise the Dutch government on taxation (the so-called van Elswijk plan) opens the way to explore in the lab the effects of alternative policies and to have an approximation of their possible effects.

In his Nobel lecture Akerlof (2002) argues that reciprocity, fairness, identity, money illusion, loss aversion, herding, and procrastination are behavioral reasons that help to explain departures from the general equilibrium model. He advocates for a “behavioral macroeconomics” that builds on them. Although some of these issues come from the experimental literature he does not make any reference to the experiments reviewed in this paper. We have already observed the tendency of experiments to replicate the comparative statics results of general equilibrium models, but experiments also show interesting departures that may be attributed to behavioral aspects. We believe that experiments can start a positive feedback with theoretical macroeconomics in designing models that take into account behavioristic issues.

Acknowledgements

A previous version of this paper was presented at the workshop “Laboratory methods in economics”, held at the University of Siena. I wish to acknowledge the influence exerted by Shyam Sunder, comments received by three anonymous referees and the Editor, encouragement by Alessandro Innocenti, and assistance by Sukhvinder Kaur. The usual disclaimer applies.

Notes

¹ However, it is a common opinion among experimentalists that professional players do not perform better than students, because they may try to show their skills, and therefore, their behavior is biased by other considerations rather than induced value.

² For example, according to Sunder experimental results on double auction markets have shown that to obtain a perfect competition allocation the assumption of a large number of traders is not necessary. In the remainder of the paper we often refer to the limitations of the “theory testing” approach to macro experiments. However, we do not believe that this method is inappropriate or wrong. We share the view that experiments are an easy way

to obtain accurate and cheap data, that otherwise, may not be available to test a theory. Of course this line of interpreting experiments has its drawbacks that we highlight in the rest of the paper.

³ Some experiments on monetary economics in Section 3.2 follow this coordination approach, together with Lei and Noussair (2003). Heinemann (2002) compares theoretical predictions for a coordination game used to explain the onset of a currency crisis with observations obtained in the laboratory, making particular reference to situations of private vs. public information.

⁴ Other experiments concerned with simpler general equilibrium models include Goodfellow and Plott (1990) and Aliprantis and Plott (1992).

⁵ Tietz (1972) proposed an early macroeconomic experiment. The framework was built around five sectors (industry, households, credit banks, the central bank, and the government). Sectors are connected through five markets. The key market is the labour market where the employers association and the labour union bargain over the wage. Bargains occur on a table between the two counterparts with the assistance of the central bank. Bargainers have information on about 200 variables computed by means of software. Individuals are assumed to be bounded rational, taking an aspiration level approach. The software produces new values of the economic variables that are consistent with the decisions taken by players. With respect to the Lian and Plott experiment it is easy to note the differences in these approaches: formal bargaining instead of anonymous trading; central role given to the labour market; computer simulations to obtain economic values as a cascade from the labour market opposed to different trading in different atomistic not explicitly linked markets.

⁶ In experiments using a binary lottery, subjects are paid an equal amount of money. In addition to this, one subject may earn an additional prize with probability given by the points earned during the experiment. Roth and Malauf (1979) show that this procedure induces risk-neutrality.

⁷ On the same issue see also Duffy (2001) and Duffy and Ochs (2002). Hens et al. (2002) provide a microfoundation of money in which markets are well-organized (i.e., every potential buyer meets every potential seller of a given service) but individual preferences are stochastic. In this model there exists a stationary solution with an optimal quantity of money. Experimental data matches fairly well the theoretical properties of the model. We argue that a monetary system is more stable than one would expect from models based solely on individual rationality. Hens and Vogt (2002) show that positive reciprocity stabilizes the monetary system, provided that every participant considers accepting money as a reasonable option. If, however, some participants notoriously refuse to accept money then due to negative reciprocity their behavior will eventually induce a break down of the monetary system.

⁸ Most of these experiments have been already surveyed by Ochs (1995) and Duffy (1998). I refer the interested reader to these papers for a detailed analysis.

⁹ Probabilistic and deterministic reward schemes are used in different economies. The former is used to induce risk aversion in consumption.

¹⁰ Marimon and Sunder (1993) show that this procedure does not affect the set of equilibria of standard OLG models, as long as they behave competitively. In turn, a random termination date without the forecasting game distorts the model introducing an implicit discount rate.

¹¹ On this experimental design, Hazlett and Kernén (2002) steadily increased the value of government expenditure for the first ten periods of the session, and then held it constant. The effect is to increase the low

stationary equilibrium level of inflation, whilst at the same time decreasing the high one. The result is an explosive inflation path and eventually a currency collapse, because the expectations of the young generation on prices were so high that they ceased trading. It should be noted that this happened only with inexperienced subjects. Indeed they were able to coordinate to the low equilibrium level. See also Bernasconi and Kirchkamp (2000) for a treatment in which both forecasts and actual saving decisions are made. Convergence towards the low inflation stationary state is found, although equilibria are more complex under both regimes and for all experimental economies than what can be approximated by a first order or by any other simple adaptive scheme. Inflation rates under the real deficit regime are on average lower than the equilibrium levels, and than the inflation rates observed on average under the (revenue equivalent) money growth rule. This comes at the price of greater inflation volatility. A significant amount of over-saving at both the individual and aggregate level is found, which is interpreted as precautionary savings due to increased uncertainty.

¹² Some different specifications of parents' utility function are also studied using an additive form or a function in which the utility of the descendants counts less than the own consumption of the parents.

¹³ In each economy the number of agents is the same, the scale effect is obtained giving the large economy seven times the resources of the small one.

¹⁴ With the same framework Riedl and van Vinden (2003) analyse the economic effects for a small open economy to switch from a wage-tax system to a sales-tax-cum-labor-subsidy system while the rest of the world uses the latter system. They found that the sales tax system outperforms the wage tax one.

¹⁵ Results found in the previous experiment concerning the real economy (comparative advantage, production and consumption patterns) are confirmed in this experiment, which can also be seen as a robustness check for the experiment without exchange rates. In a more theory-friendly experimental setup, Fisher (2001) finds strong evidence for the (absolute and relative) purchasing power parity, covered interest parity, and uncovered interest parity.

¹⁶ Lei and Noussair (2003) build an economy with two Pareto-rankable locally stable equilibria and find that without specific reasons the economy may end up in the lower equilibrium, which they interpret as a poverty trap. This occurs more likely under the low endowment treatment, and affects both the market and the central planner environments. This experiment opens the way to future investigations on which institutions and policies may help agents to coordinate to the Pareto-superior equilibrium.

¹⁷ In this economy there are also producers, who only buy input, transform it into output, and sell it. Each individual has assigned redemption values from buying and selling input and output, respectively, to control for their preferences. Agents exchange input and output through a single unit double auction.

References

- Akerlof, G.E. (2002) Behavioral macroeconomics and macroeconomic behavior. *American Economic Review* 92, 411-3433
- Aliprantis, C.D. and Plott, C.R. (1992) Competitive equilibria in overlapping generations experiments. *Economic Theory* 2, 389-426
- Arifovic, J. (1996) The behavior of the exchange rate in the genetic algorithm and experimental economies. *Journal of Political Economy* 104, 510-541
- Arifovic, J. and Sargent, T.J. (2001) Laboratory Experiments with an Expectational Phillips Curve. Stanford University, mimeo
- Barro, R.J. (1974) Are government bonds net wealth? *Journal of Political Economy* 82, 1095-1117
- Barro, R.J. and Gordon, D.B. (1983) A positive theory of monetary policy in a natural rate model. *Journal of Political Economy* 12, 101-121
- Benassy, J.P. (1982) *The Economics of Market Disequilibrium*, New York: Academic Press
- Bernasconi, M. and Kirchkamp, O. (2000) Why monetary policy matters - An experimental study of saving, inflation and monetary policies in an overlapping generations model. *Journal of Monetary Economics* 46, 315-343
- Bernasconi, M., Kirchkamp, O. and P. Paruolo (2003) Expectations and perceived causality in fiscal policy: An experimental analysis using real world data. SFB 504 Discussion Paper n. 03-03, University of Mannheim
- Bertola, G. and Drazen, A. (1993) Trigger points and budget cuts: explaining the effects of fiscal austerity. *American Economic Review* 83, 11-26
- Blanchard, O.J. (1997) *Macroeconomics*. New Jersey: Prentice Hall
- Bosch-Domenech, A. and Silvestre, J. (1997) Credit constraints in general equilibrium: experimental results. *Economic Journal* 107, 1445-1465
- Brown, P.M. (1996) Experimental evidence on money as a medium of exchange. *Journal of Economic Dynamics and Control* 20, 583-600
- Bryant, J. (1983) A simple rational expectations Keynes-type model. *Quarterly Journal of Economics* 98, 525-538
- Cadsby, C.B. and Frank, M. (1991) Experimental tests of Ricardian equivalence. *Economic Inquiry* 29, 645-664
- Camerer, C. (1995) Individual decision making. In Kagel, J. and Roth, A. (eds.) *The Handbook of Experimental Economics*, Princeton, Princeton University Press
- Cass, D. (1965) Optimum growth in an aggregative model of capital accumulation. *Review of Economic Studies* 32, 233-240
- Clower, R.W. (1965) The Keynesian counter-revolution: a theoretical appraisal. In F.H. Hahn and F. Brechling (eds.), *Theory of Interest Rates*. New York: McMillan
- Cooper, R.W. (1999) *Coordination Games. Complementarities and Macroeconomics*. Cambridge: Cambridge University Press
- Davis, D. and Holt, C. (1993) *Experimental Economics*. Princeton: Princeton University Press
- Duffy, J. (1998). Monetary theory in the laboratory. *Federal Reserve Bank of St. Louis Review* September/October, 9-26
- Duffy, J. (2001) Learning to speculate: experiments with artificial and real agents. *Journal of Economic Dynamics and Control* 25, 295-319
- Duffy, J. and Ochs, J. (1999) Emergence of money as a medium of exchange. *American Economic Review* 89, 847-877
- Duffy, J. and Ochs, J. (2002) Intrinsically worthless objects as media of exchange: experimental evidence. *International Economic Review* 43, 637-674

- Fehr, E. and Tyran, J.R. (2001) Does money illusion matter? *American Economic Review* 91, 1239-1262
- Fisher, E. (2001) Purchasing power parity and interest parity in the laboratory. *Australian Economic Papers* 40, 586-603
- Friedman, D. and Rust, J. (1993) The double auction market. Institutions, theories and evidence. Reading: Addison-Wesley
- Goodfellow, G. and Plott, C.R. (1990) An experimental examination of the simultaneous determination of input prices and output prices. *Southern Economic Journal* 56, 969-983
- Hazlett, D. and Kernen, A. (2002) Hyperinflation and seigniorage in an experimental overlapping generations economy. Whitman College, mimeo
- Hens, T., Shenk-Hoppé, R. and Vogt, B. (2002) On the micro-foundation of money: the Capitol Hill baby-sitting co-op. Working Papers Institute for Empirical Research in Economics n. 108, University of Zürich
- Hens, T. and Vogt, B. (2002). Money and reciprocity. Working Papers Institute for Empirical Research in Economics n. 138, University of Zürich
- Hey, J.D. and Di Cagno, D. (1998) Sequential markets: an experimental investigation of Clower's dual-decision hypothesis. *Experimental Economics* 1, 63-85
- Heinemann, F. (2002) Exchange-rate attack as a coordination game: theory and Experimental Evidence. *Oxford Review of Economic Policy* 18, 462-478
- Koopmans, T.C. (1965) On the concept of optimal economic growth. In *The Econometric Approach to Development Planning*, Amsterdam: North Holland
- Lei, V. and Noussair, C. (2002) An experimental test of an optimal growth model. *American Economic Review* 92, 549-570
- Lei, V. and Noussair, C. (2003) Equilibrium selection in an experimental macroeconomy. Emory, University mimeo
- Lian, P., and Plott, C.R. (1998) General equilibrium, markets, macroeconomics and money in a laboratory experimental environment. *Economic Theory* 12, 21-75
- Lim, S., Prescott, E.C. and Sunder, S. (1994) Stationary solution to the overlapping generations model of fiat money: experimental evidence. *Empirical Economics* 19, 1005-1014
- Kiyotaki, N., and Wright, R. (1989) On money as a medium of exchange. *Journal of Political Economy* 97, 927-954
- Kydland, E.C. and Prescott, F. (1977) Rules rather than discretion: the inconsistency of optimal plans. *Journal of Political Economy* 85, 473-491
- Mankiw, N.G. (1985) Small menu costs and large business cycles: a macroeconomic model of monopoly. *Quarterly Journal of Economics* 100, 529-539
- Marimon, R., Sunder, S. (1993) Indeterminacy of equilibria in a hyperinflationary world: experimental evidence. *Econometrica* 61, 1073-1107
- Marimon, R. and Sunder, S. (1994) Expectations and learning under alternative monetary regimes. *Economic Theory* 4, 131-162
- Marimon, R. and Sunder, S. (1995) Does a constant money growth rule help stabilize inflation? Experimental evidence. *Carnegie Rochester Conference Series on Public Policy* 43, 111-156
- McCabe, K.A. (1989) Fiat money as a store of value in an experimental market. *Journal of Economic Behavior and Organization* 12, 215-231
- Noussair, C.N., Plott, C.R. and Riezmann, R. (1995) An experimental investigation of the patterns of international trade. *American Economic Review* 85, 462-491

- Noussair, C.N., Plott, C.R. and Riezman R. (1997) The principles of exchange rate determination in an international finance experiment. *Journal of Political Economy* 105, 822-861
- Ochs, J. (1995) Coordination problems. In: J.H. Kagel and A.E. Roth (eds.) *Handbook of Experimental Economics*. Princeton: Princeton University Press, 195-251
- Ricciuti, R. and Di Laurea, D. (2003) An experimental analysis of two departures from Ricardian equivalence. *Economics Bulletin*, 8, 1-11
- Riedl, A. and Van Winden, F. (2001) Does the wage tax system cause budget deficits? A macro-experiment. *Public Choice*, 109, 371-394
- Riedl, A. and Van Winden, F. (2003) Input versus output taxation in an experimental international economy. University of Amsterdam, mimeo
- Roth, A. and Malouf, M.W.K. (1979) Game theoretic models and the role of information in bargaining. *Psychological Review* 86, 574-94
- Slate, S., McKee, M., Beck, W. and Alm, J. (1995) Testing Ricardian equivalence under uncertainty. *Public Choice* 85, 11-29
- Smith, V. and Walker, J. (1993) Monetary rewards and decision cost in experimental economics. *Economic Inquiry* 31, 237-244
- Sunder, S. (2001) Real phenomena, theory and design of laboratory experiments in economics. Mimeo, Yale University
- Sutherland, A. (1997) Fiscal crises and aggregate demand: can high public debt reverse the effects of fiscal policy. *Journal of Public Economics* 65, 147-162
- Tietz, R. (1972) The macroeconomic experimental game KRESKO. Experimental design and the influence of economic knowledge on decision behavior. In: H. Sauermann (Ed.) *Beiträge zur experimentellen Wirtschaftsforschung*, Vol. 3, Tübingen, 267-288
- Van Huyck, J.B, Battailo, R.C. and Beil, R.O. (1990) Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review* 80, 234-248
- Wilson, B.J. (1998) Menu costs and nominal price frictions: an experimental examination. *Journal of Economic Behavior and Organization* 35, 371-388

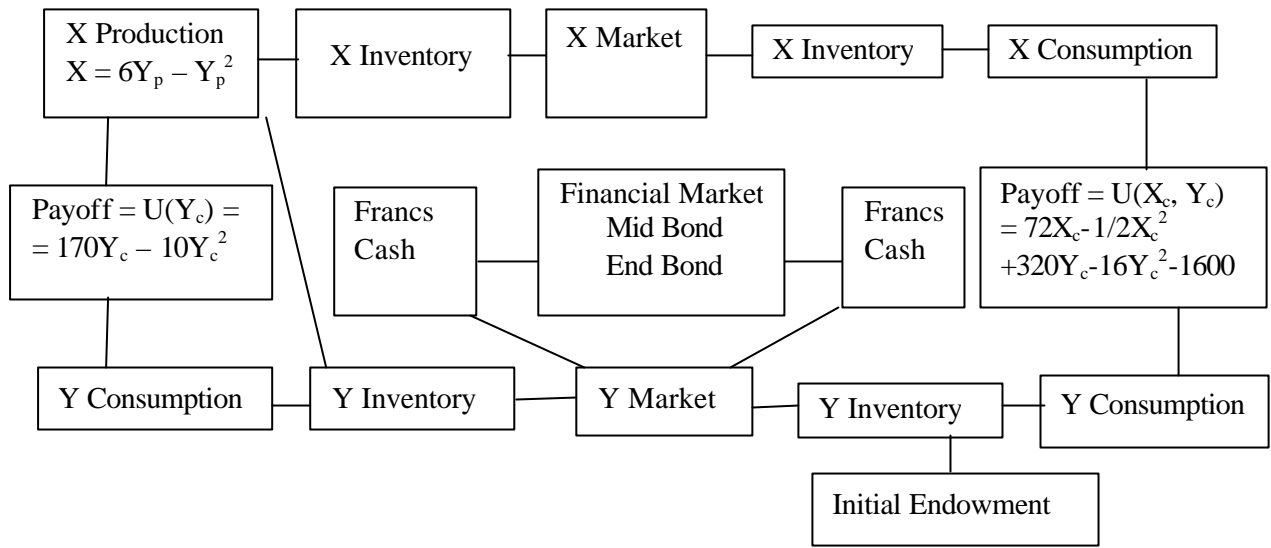


Figure 1 – Circular flow of the economy
Source: Lian and Plott (1998)

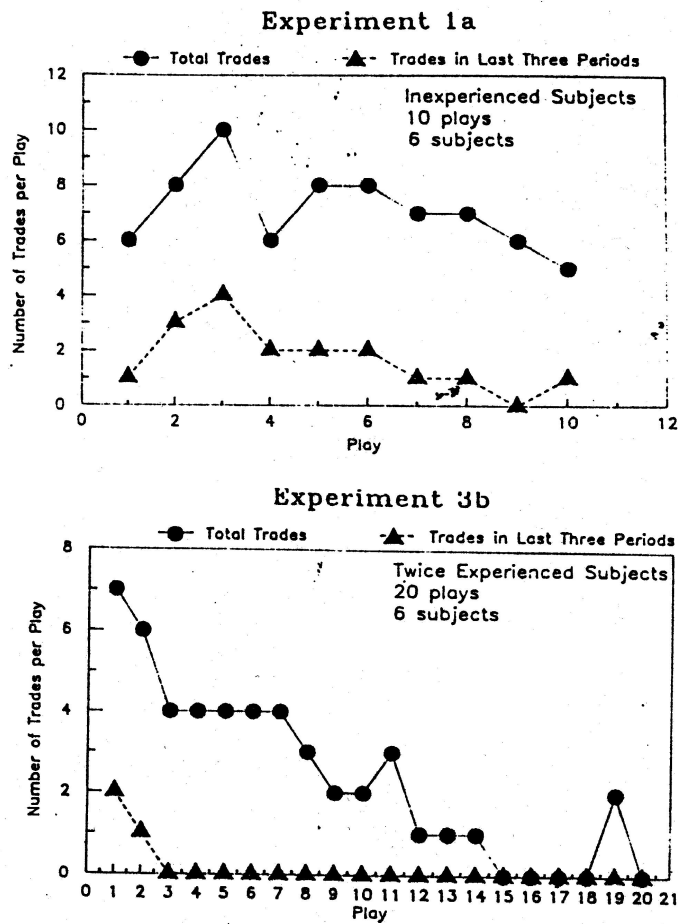


Figure 2 – Trades: experienced and inexperienced subjects.
Source: McCabe (1989)

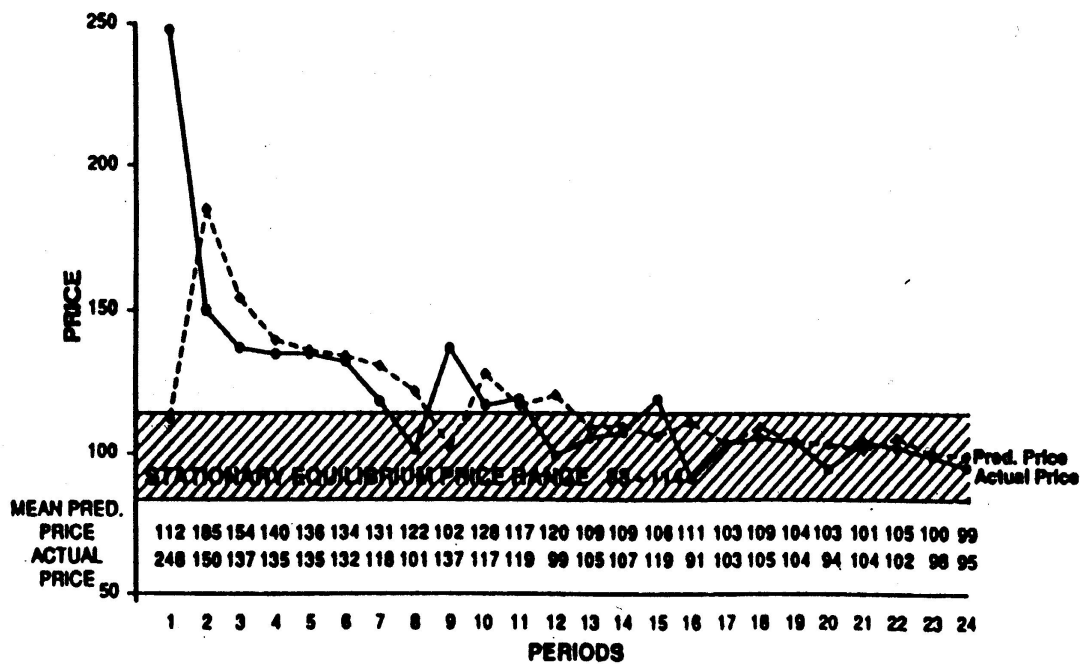


Figure 3 – Transaction prices in economy 4.
Source: Lim *et al.* (1994)

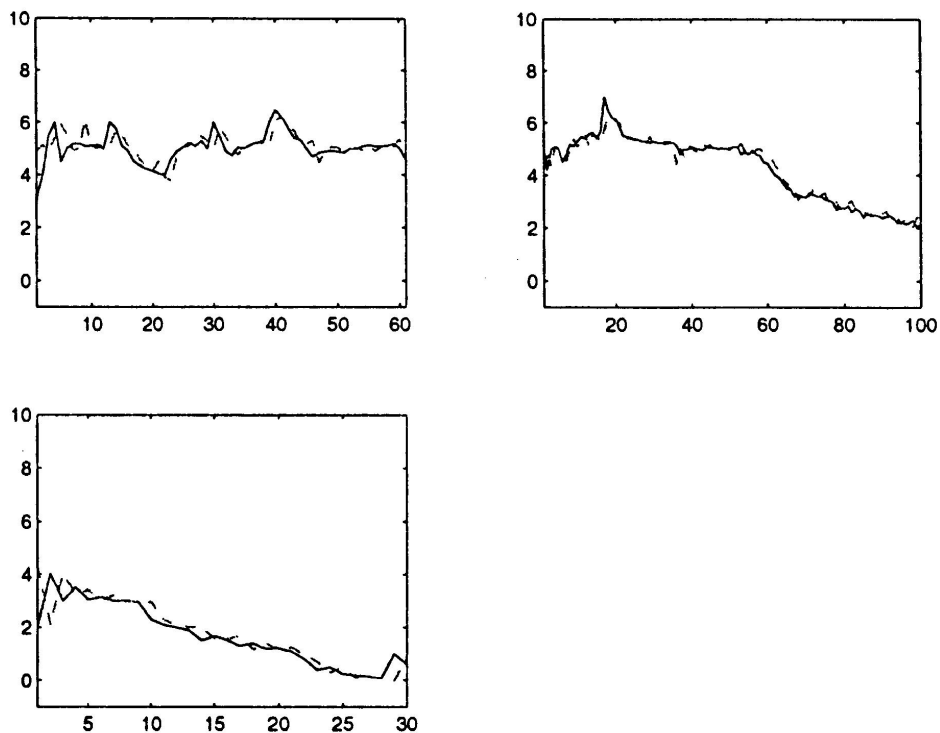


Figure 4 – Inflation forecast (dotted) vs. actual inflation (solid) in economy 1.
Source: Arifovic and Sargent (2001)

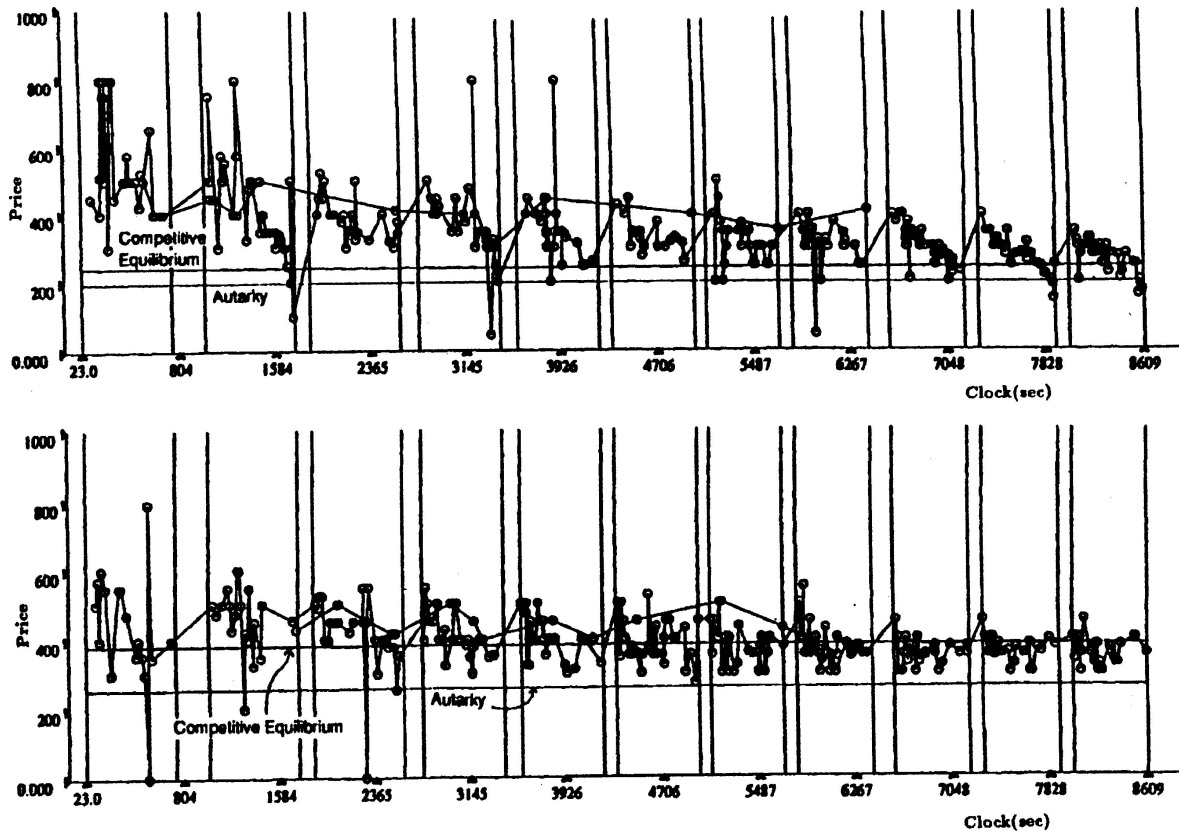


Figure 5 – Output price time series, experiment 041391A: prices of Y in both countries (upper) and prices of Z in both countries (lower)
 Source: Noussair et al. (1995)