THREE ESSAYS IN THEORETICAL, APPLIED, AND NORMATIVE ECONOMICS

Thesis by

Stephen Richard Selinger

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1984

(submitted October 1983)

ACKNOWLEDGEMENTS

I would like to express my gratitude for the help and encouragements I have received from the four members of my committee. My chairman, Burt Klein, made valuable comments on all three essays in their different stages of development. Bruce Cain provided valuable strategic, stylistic, and substantive advice (not necessarily in that order) throughout the writing of this dissertation. Quang Vuong was instrumental in formulating the econometric tests relevant to an efficient market test of the real estate market. David Grether joined my committee after much of the dissertation had already been written and was able to offer a valuable perspective on how the papers read to someone who hadn't been close to them in their inception. It has been both a personal and professional pleasure to work with all four of these individuals.

ABSTRACT

The first essay of this dissertation treats the question of ethical fairness towards future generations. It is argued that Harsanyi's equi-probability characterization of the original position captures the notion of ethical fairness and that the result of applying this model to a future generations context and also satisfying the axioms of the expected utility theorem results in classical utilitarianism being chosen. This is in contrast to the average utilitarianism which is widely thought to be the more plausible utilitarian position in a short run framework. It is also argued that classical utilitarianism does not entail a situation where individuals would exist at a subsistence level as some (Parfit) have assumed.

The second essay is an efficient market test of the real estate market. The question of whether lagged real interest rates contain statistically significant information about future housing prices is examined. It is found that the coefficients of lagged real rates of twelve and eighteen months were negative and statistically significant; thus efficiency is rejected. A Hausmann test was then run to see if it was permissible to use an ordinary least squares approach; such an approach was valid.

The third essay examines the effect of inflation on rates of return in different socieo-economic areas. Measures of expected and unexpected inflation are defined. The rates of return from holding real estate in different areas are then regressed upon the measures of expected and

unexpected inflation. A Chow test was then run to see if it was permissible to pool the coefficients of expected and unexpected inflation. The pooling is permissible and so we can say that in a statistical sense, inflation had the same impact upon the different areas.

TABLE OF CONTENTS

FUTURE GENERATIONS, THE ORIGINAL POSITION, AND CLASSICAL UTILITARIANISM

Many of the public policy options that confront us in various areas have costs and benefits that arrive in the future. The costs of disposing of nuclear waste or hazardous chemicals are future costs which will have to be borne by future generations. Similarly, when we deplete non-renewable resources, the opportunity cost of not having the resources available will have to be borne by future generations. In deciding public policy questions about such issues, to what normative standard are we to appeal?

I will answer this question within a contractarian framework, and since the motivation for this framework has been discussed at length elsewhere, I will touch upon it only briefly here.[1] In the hypothetical contracting situation, individuals are deprived of all information about their particular situation, e.g. about their specific preferences, sex, race, religion, etc. This ignorance is dictated by the two pre-theoretical intuitions which motivate the adoption of a contractualist framework -- (1) impartiality and (2) autonomy.

(1) Impartiality regarding choice of the principles which are to govern society follows from the fact that since people don't know the particulars of their own situation, they can't possibly be partial to them. (2) Autonomy of choice follows in the same way: since the contingent facts of our own situation are unknown they cannot influence us. We will be free from what Kant called heteronomous influences due to our lack of knowledge of the contingent facts.

This notion of autonomous choices can be seen as being compatible with physical determinism; this notion does not require us to repeal the first and second laws of thermodynamics.[2] Rather the notion of autonomous choices can be seen as being compatible with physical determinism if the choices are determined under appropriate circumstances. These circumstances are those which would not cause an individual to be influenced by the particular contingent facts of his own existence. And since an individual is deprived of knowledge of these particular facts, he can't possibly be influenced by them.

Since an individual does not know what his particular position in society will be, the choice of principles in the original position can be thought of as a choice between uncertain prospects. The prospects are the different positions in society and the different temporal societies an individual might occupy. Let us say that there are n individuals in society and m different time periods in which an individual might occupy them. There are thus m x n positions an individual might occupy. Since an individual is ignorant of which society he will live in, and what his position in society will be, he will assign each prospect the probability $1/(m \times n)$. Now it follows from the Marschak postulates that when an individual chooses a social state under these conditions, he will pick that one which maximizes the arithmetical mean. For if an individual's choices satisfy the Marschak postulates, then he behaves as if he were maximizing his expected utility, and a given social policy A yields an individual the expected utility

$$U(A) = (1/(m \; x \; n)) \sum_{j=1}^{n \, x \, m} Uj(A)$$

This is so because an individual would have 1/(m x n) chance of being put in that place of each individual j(j = 1, . . . m x n) and hence of obtaining the utility amount Uj(A), i.e. j's utility in situation A. Thus individuals in the original position will choose the moral principle requiring us to evaluate each social situation A in terms of the long run average utility that the n x m members of society would enjoy in this situation. Social policies would then be ranked according to their ability to produce desired social states (as defined by their average utility level).

The proposal outlined above is the straightforward extension of Harsanyi's 1955[3] powerful argument to a society which exists through time. Harsanyi considered a society of n individuals all of whom exist contemporaneously; above I considered n individuals who exist in m successive time periods. This characterization seems the most natural first attempt to extend a contractarian framework to a future generations context.[4] But it is unsatisfactory as it stands, and for two main reasons.

First, the number of people and the number of societies should not be thought of as exogenously given. Rather, they are part of the problem we are investigating because they will be to a large extent determined by our actions. Thus it is a mistake to specify the problem by assuming, as we have so far done, that there are m x n people (since there are m people in each period and n different time periods). The number of people in each period will be determined to a large extent by the institutional arrangements regarding birth control, abortion, free education, health care, etc., which as they vary cause there to be more or less than m people in each society. Similarly, the number of different societies will be determined by actions we take. If a current generation decides to solve the problems of chemical disposal, nuclear waste disposal, or arms control in risky ways, the number of periods may turn out to be far fewer than n. There is thus a problem of self-referentiality: appeal to the original position is supposed to guide our actions, but our actions affect the original position as it has so far been formulated.

The fact that the number of people in each generation and the total number of generations is not independent of our actions vitiates any straightforward use of the expected utility theorem such as the kind we had in mind in our straightforward extension of Harsanyi's theorem. For one of the axioms of the expected utility theorem is that the states of nature are independent of one's actions. In assessing the action of planting seeds, for example, we are to look at how probable the chances of rain or sun are. The likelihood of the states of nature -- in this case, rain or sun -- are rightfully taken to be independent of our actions. This is not so in the case we are considering, however. For here the probability of being placed in the position of any state of nature is $1/(m \times n)$ and this is surely dependent upon our actions. This dependence of the states of nature upon our actions seems a decisive logical reason why Harsanyi's theorem cannot straightforwardly be applied to a future generations context.[5]

The second main problem with the straightforward extension of Harsanyi's theorem to a future generations context is that this extension violates a consistent application of the impartiality constraint motivating the adoption of a contractualist framework. Thus the second is a moral, rather than a logical objection. In the above specification of the contractualist framework it is implicitly assumed that we already exist and we are asked to pick principles, given that we do not know who we are. The mere fact, however, that we exist and know that we do, is morally relevant information which affects the choice of principles in the original position. Knowing that I exist, I will be concerned only with improving the quality of life of already existing people -- providing, of course, that I act as a rational agent in the original position is supposed to act. My knowledge that I already exist (or will definitely exist in the future), however, is akin to knowing, e.g., that all people in the original position are Caucasian. If the parties knew this, and then acted as rational agents act, they would then select institutional principles which maximize the average utility of Caucasians in society. One would naturally object that the choice of principles is biased by the informational assumption that all individuals in the original position are Caucasian. A similar bias is introduced by the knowledge of all parties in the original position that they will exist in the present or future societies. In a word, choice of principles is biased by including morally prejudicial information. People who might exist under different institutional arrangements are given no say in the formation of the background institutions of society. These individuals are left out in the cold the way non-Caucasians would be left out were all parties in the original position told that they were Caucasians.

One might object that there is a misplaced analogy about being unbiased towards existing people such as non-Caucasians (who will definitely be harmed by including prejudicial information into the original position) and non-existent people who will not be harmed (because they don't exist) if they are not brought into existence; in the first case, actual existing people will be made unhappier and worse off while in the second case, no one will be made worse off because these potential people don't even exist.

The problem with this objection is that it appeals to a notion of existing people that is supposed to be well defined independent of our actions. We can't without circularity say that the non-existing people will be no worse off because whether these potential people are non-existing or not depends upon our actions and we as yet have not determined what is the morally correct action to take. Thus the objection about there being a morally relevant difference between (allowably) including information about existing people versus (not allowably) including information about whether people do exist fails because the "information" about whether people exist is determined by our choice of actions and this has not as yet been determined.

To summarize, then, there are two main problems with the straightforward extension of the single period contractualist framework to a future generations contractualist framework: first, the number of people and periods isn't even well defined in the straightforward extension because the number of people and periods is dependent on our own current actions and yet we aren't told how many people or periods to produce but only told that m individuals and n periods exist; thus one of the axioms required for the expected utility theorem is not satisfied and so the theorem cannot be applied to achieve a straightforward extension of Harsanyi's theorem. Second, the straightforward extension biases the choice of principles by having the members assume that they will exist regardless of the principles chosen.

It seems that a more sophisticated way to model a multi-generational contractarian situation would be to assume that the number of people and generations is allowed to vary depending upon the institutional principles chosen and can vary between 0 and some large but finite number. It seems plausible that there are some obstacles of nature which would prevent the infinite continuation of the species, e.g., the sun will burn out so many years in the future. It does not seem objectionable to take the number of people as exogenously given in the weak sense that there are certain natural and technological forces preventing the infinite perpetuation of the species. Some of these seeming absolute constraints could be overcome by, e.g., sending people to other galaxies to exist even after the sun burned out but perhaps all could not be so overcome. At any rate, to simplify the analysis, it will here be assumed that the choice of principles is subject to broad exogenously given natural forces and technological forces which serve to keep the number of possible people finite. This is largely an assumption made for the sake of convenience and I don't think it would change the analysis if the possible population was allowed to be infinite.

Let the exogenous limit on the maximum number of people the world could possibly support equal Q. Let the number of generations which result from the choice of principles in the original position be R and the number of people in each generation be S (we'll assume for the sake of simplicity that the number of people in each generation is the same). R x S must thus be less than (or possibly equal to) Q. A given social policy A now yields an individual in the original position the expected utility

$$W(A) = (1/Q) \sum_{j=1}^{(S \times R)(A)} Uj(A)$$

since an individual in the original position would have chance 1/Q of being in the place of one of the S x R members that policy A causes to exist. This principle would have individuals in the original position rank social policies according to their ability to result in the highest average utility of all individuals that would exist conditional on the policies chosen. In effect it amounts to the classical utilitarian position of seeking to produce the maximum amount of happiness because since the number of individuals is constant in the original position (Q), the greatest average happiness of people in the original position is produced by producing the greatest amount of total happiness in the real world (outside the original position).[6]

It is permissible to use the expected utility theorem and apply it to a multi-generational context in the way we have above because the states of nature as now defined are independent of our actions. There are now taken to be Q states of nature where Q is the maximum number of people the world could possibly support. Q is thus exogenously given and is independent of our actions. The axioms of the expected utility theorem are thus satisfied and so it is permissible to appeal to it in establishing our result of classical utilitarianism. It does seem rather surprising to me that when the logical objection of the dependence of the states of nature upon our actions is removed, which at the same time removes the moral objection about the partiality of information in the original position, the strikingly different conclusion of the classical utilitarianism rather than average utilitarianism results from the original position.

The conclusion that the classical utilitarian position of producing the greatest amount of happiness possible would be chosen in the original position contrasts with the claim Rawls makes when he says that the original position construction serves to highlight the differences between classical and average utilitarianism:

> From the standpoint of the persons in the original position, it would appear more rational to agree to some sort of floor to hold up average welfare. Since the parties aim to advance their own interests, they have no desire in any event to maximize the sum total of satisfaction. I assume, therefore, that the more plausible utilitarian alternative to the two principles of justice is the average and not the classical principle.[7]

The reason Rawls is led to believe that the original position construction favors an average utilitarian position which is at odds with a classical position is because he does not follow the methodology of the original position rigorously enough. He has all individuals assume that they will exist no matter what policies are followed which naturally biases the decision process in favor of an average utilitarian view. A more rigorous following of the original position, however, results in the conclusion that the average utility of all possible people would be maximized; and as has been pointed out above, this is equivalent to maximizing total happiness since the denominator Q is constant across possible policies.[8]

FURTHER CONSIDERATIONS

The argument presented in the first section needs to be checked for the intuitive appeal of its consequences; otherwise if the consequences are too unintuitive, we might just reject the reasoning or premises which led to our conclusion. One might also reject the original position framework if one thought the conclusion was too counter-intuitive.

One initially intuitively appealing solution to the problem of intergenerational justice is that of equal opportunity for all generations.[10] This standard says that justice requires that all future generations be provided with a resource base that allows them the same opportunity earlier generations had. This standard also does not imply that individuals might sometimes be required to have children to increase happiness even if they don't want them so it may seem intuitively superior on that score to the total utilitarian position of part I.

Though the equal opportunity standard has the above mentioned strengths, it also has certain drawbacks. First, as an internal criticism, the standard should probably be modified so that it focuses upon the capital stock and technological expertise as well as the natural resource base. If a later generation had fewer natural resources than an earlier one, but a significantly larger capital stock and technological expertise, it would seem that the equal opportunity standard should say that they were justly treated.

Second, and more importantly, the equal opportunity standard suffers from the same distributional implausibilities that have plagued Rawls' difference principle. Imagine that it was extremely costly to maintain the resource base after a certain point in time. If the resource base was depleted according to one plan the people in the year 2100 would have utility 100 and those in the year 2150 would have utility 10. However if we wish to fulfill the equal opportunity standard we would have to follow a plan (if one existed) which would yield the generation of the year 2100 a utility of 11 and the same utility for generation 2150. This seems an implausible restriction. Rather it seems more plausible to follow the utilitarian position of favoring equality insofar as it is reflected in people's utility functions due to diminishing marginal utility but not to give it the absolute weight that the equal opportunity standard does.

Third, to be plausible the equal opportunity standard must be interpreted relative to some population base. That is, consider if no one wanted to have children in later generations. If this happened it would be foolish to keep the resource base intact for them. Similarly, if the population was going to be much larger in the future, the resource base would have to be much greater. The size of the population in the future will be to a large extent dependent on the actions of people today and yet the equal opportunity standard does not give us any guidance on this question. Thus the standard is fundamentally incomplete as a standard of justice for future generations since it is logically dependent upon the size of the future population which in turn depends upon our actions and yet the equal opportunity standard does not give us any guidance about how many people to produce.

One response the equal opportunity standard advocate might make is to say that the standard is indifferent as to the population base as long as each individual throughout history has equal opportunity. We can wonder, however, why the standard should not favor the creation of another person if this person will have a worthwhile life. That is, why shouldn't the standard be moved from indifference about the population base (as long as there is equal opportunity) to strict preference? And if we decide to opt for a theory which says that we should strictly prefer the larger popluation, doesn't it seem permissible to opt for this larger population if it will only cause a very slight decrease in the resources available for future generations? We will turn to these questions as we examine the positive intuitive attractions of a classical utilitarianism position.

Although the equal opportunity standard does not seem as appealing upon reflection as it initially did, we still must check upon the appeal of the proposal put forward in the first part. It must be considered whether the conclusion of the deductive argument is in accord with our considered judgements about these matters; and if it is not initially in accord with our judgements, we must see whether our judgements are changed upon reflection. One intuitive argument which can be made in favor of a total utilitarian view is to consider our intuitions about what would be the right course of action if individuals in a society were perfectly indifferent about whether to bring about an increase in population via the birth of a child. If they were so indifferent, doesn't it seem intuitively plausible to think that the proper course of action would be to have the child if this child would lead a happy life? The argument would then say that even if the people have a mild preference for not having the child, if the gain in utility to the child (or children) is great enough, the correct course of action would be to have the child. The utilitarian would then seek to guide this balancing by appealing to considerations of total utility. Considerations of average utility couldn't guide the balancing because by hypothesis it is admitted that sometimes it is better to have a child even if having the child will result in a lower level of average utility (if existing individuals have only a mild preference against having the child and the child will only be brought up to the former average utility level). Hence unless our intuitions are to make this judgement unguided by any rule, we will have to appeal to a total utilitarian rule. In essence, a variation of the argument that Sidgewick used so persuasively so many years ago.

This appears to be the sort of argument that Derek Parfit uses and calls the "mere addition" argument. Parfit asks us to consider a situation where a group of people could be added to the population at a lower level of utility than the people than the people already in the population. These additional people would not cut into the resources of the previously existing people or lower their utility in any way. They would, however, lower the average utility of society since they are added to the population at a level that is lower than the average. Parfit's mere addition example compellingly illustrates the implausibility of an average utilitarian viewpoint vis-a-vis a total utilitarian in certain situations; if these additional individuals will have lives that are worth living, and they won't harm the people already living, it does seem rather implausible not to add them to the population just because they will lower the average utility.

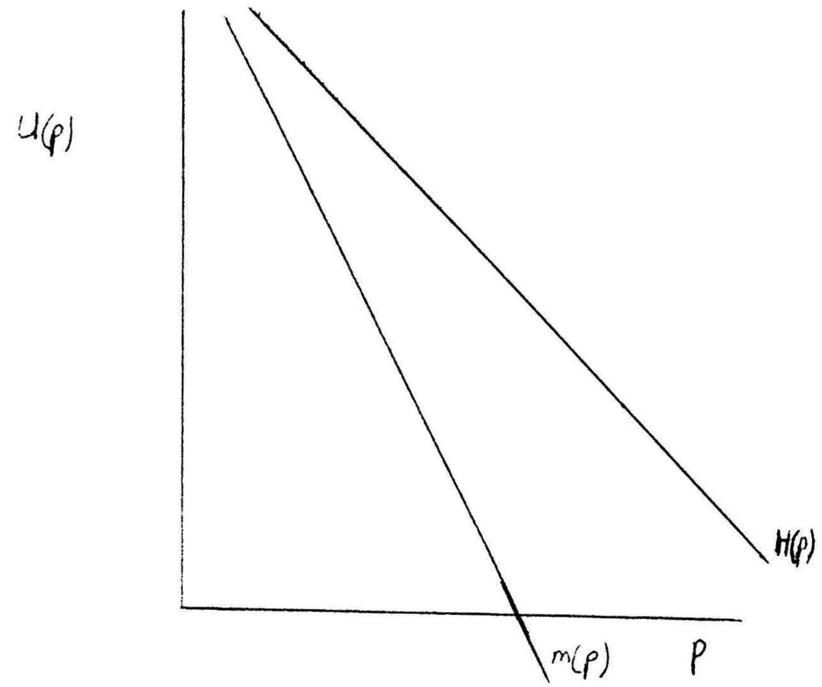
However, the logic of the mere addition argument and some other assumptions which Parfit considers rather plausible, lead him to a conclusion which he considers repugnant. Parfit observes that if one assumes an egalitarian viewpoint, then a society B (in which the utilities of newly added people are combined and averaged with the utilities of the people who already existed) is superior to the former society A (with the higher utilities of the formerly well off separated from the lower utilities of the newly added people). Parfit observes that if one favors equality as a goal this new society should be preferred to the inegalitarian one which separated the old and newly added groups. However if we continue this line of reasoning, we realize that this newly formed egalitarian society should be able to be improved upon by adding new people who are at a lower level of utility and yet whose addition doesn't harm these people;[11] and this society could be improved upon by a more egalitarian one, etc., etc. The conclusion is the repugnant conclusion that any society with a certain amount of people can be improved upon by another society with sufficiently additional amounts of people. Parfit weakens this statement to saying that the new society with sufficiently more people can't be judged worse off than the old society to arrive at his new repugnant conclusion:

> THE NEW REPUGNANT CONCLUSION: For any possible and large population, say of eight billion, all with a very high quality of life, there must be some much larger imaginable population whose existence, other things being equal, would not be worse, even though its members have lives that are not much above the Restricted Level.[12]

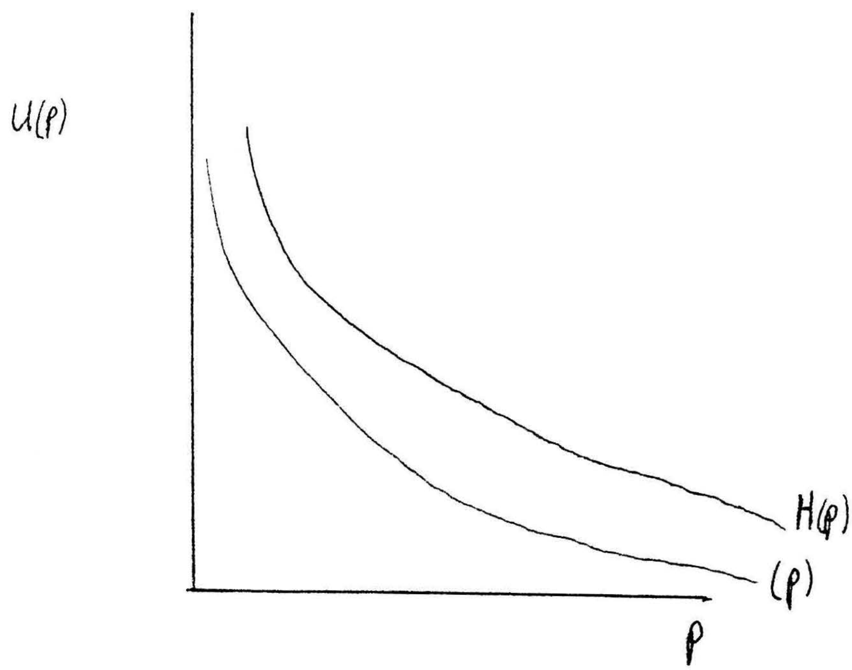
Parfit tells us the restricted level is one at which lives "are worth living but are gravely deprived, crimped and mean -- not much above the level where it would be intrinsically bad that these people are alive." He thus argues that the logic of following a total utilitarian view (as given in the "mere addition argument"), along with an assumption of egalitarianism, leads to the repulsive conclusion that we can't rule out a society with many more people than ours but whose members have lives that are barely worth living. Parfit's argument, if successful, might lead us to rethink the line of reasoning which leads to the total utilitarian view proposed earlier.

A difficulty in Parfit's argument is that it is not clear how the "must" in the statement of the repugnant conclusion is to be interpreted. When he says that there must be some larger population which would not be worse than the previous one -- even though its members all have lives not much above the restricted level -- it is not clear whether he means the "must" as an empirical or logically possible "must." If the argument is that it must empirically be possible to increase the total happiness of any society by adding more people until these people reach the restricted level, then the argument seems false. In order to reach the maximum amount of happiness, the amount of people should be increased until the loss in individual utility more than offsets the gain in utility from the extra individuals now existing. As economists we would say that to maximize happiness the number of individuals should be increased to the point where the marginal happiness curve is 0. Beyond this point, any increase in individuals causes happiness to decrease. Similarly, it is false to say that it is possible to merely add people beyond this point without harming others because as total happiness will be less, the pre-existing people will also be worse off. The situation is analogous to a monopolist seeking to maximize total revenue; the monopolist increases production to the point where the gain from selling an increased unit is exactly offset by the loss in price in the previously existing units (or more precisely, in individuals higher up on the demand curve who are buying at the same time). Seen in this way it is just empirically false to say that beyond this point (where the marginal happiness of individuals is 0) it is possible to increase total happiness by adding more people. The mere addition argument fails because beyond the point where the marginal of happiness of individuals is zero, people can't be added to the population without harming the utility of other people in the population. Beyond this point, any further increases in population harm already existing people and will also decrease total happiness.

Parfit seems to have in mind a situation where the marginal happiness curve never falls below 0. In the graphs on the following page, U(p) is the average level of an individual's utility when the population is p. Total happiness is thus equal to p x U(p) and is equal to the area under the curve Hp. The marginal happiness schedule is Mp and when this crosses below 0, further increases in population decrease total happiness. In the first graph, the marginal happiness schedule falls below 0, while in the second graph it does not fall below 0. Thus if a graph such as the first reflects the happiness of a society as a function of the population of a society, it is not possible to indefinitely add individuals to the population without decreasing total happiness. If a graph such as the second describes the relationship, then it is possible to indefinitely add people without decreasing total happiness. The possibility of societies such as those graphed in (a), however, enables us to reject Parfit's claim that for *any* possible and large population, it must be possible to add more people to form a society which would not be worse (since its former members could be kept at their old level of utility and new members added, or we



(a)



b)

could use the egalitarian variant of this society). For in the case of societies such as those graphed in (a), it is not possible to continually add people without harming those already existing.

Though Parfit's Repugnant Conclusion does not seem to follow from his premises, we can still ask the question of whether a total utilitarian position would result in a large population at a very low level of happiness. In maximizing total happiness at the point where the elasticity equals one, will each individual be at a very low level of utility close to Parfit's restricted level? If the happiness maximizing level is near this restricted level, then we still might have an intuitive objection to total utilitarianism.

It should be stated that although total utilitarianism may make more stringent demands upon existing individuals than average utilitarianism, this is not sufficient grounds upon which to reject it. The original position is supposed to convince existing individuals who are or will be brought into existence not to be partial to their own positions by depriving them of the knowledge that they will definitely be brought into existence. Thus I do not want to claim that the total utilitarian conclusion will not make any new demands upon existing individuals; rather it will make such new demands but these demands should not be rejected merely because existing individuals want to be partial to their own situation.

What if someone agrees with this but objects that total utilitarianism is still too demanding in that it requires that individuals must sometimes have children when they don't want to? Doesn't this objection seem to have some intuitive plausibility? While total utilitarianism will sometimes say that individuals should have children when they don't want to, it seems doubtful that it would often require couples to have children when they have a strong preference not to because new children are generally not very happy if brought into the world with no one wanting them. It also seems better to admit that it would be better for society to be organized so that maximum happiness was produced (even if it meant that sometimes individuals had to have children when they did not want to) and to admit that society was falling short of that standard (which was derived and justified by a consistent application of the original position) and could do better than to say that the best society is one in which individuals are allowed to be partial to those currently existing. Why not admit that society could do better by producing more people but that it is not doing so? This does not mean that the society is evil or wicked as it is but rather only that it could be improved by moving in the direction of a total utilitarian position. This seems a better course than saying that merely because a moral theory makes demands upon us that we don't entirely fulfill, we should scrap the theory. Rather we should admit that we could be doing better and try to do so.

We might still wonder about the point where the elasticity of total happiness is 1. Is this a point where individuals' utility levels will be close to the restricted level? Even if we admit that total utilitarianism can permissably make demands upon us, doesn't it still seem objectionable if individuals end up near the restricted level? Parfit is bothered by this when he says:

> On the New Conclusion, Higher Z [a society with lives near the restricted level] is only claimed to be not worse than A [a society with a fewer number of people but a high average level of utility]. But this still seems, to me at least, pretty repugnant. Lives that are not much above our Restricted Level cannot be well worth living. Even if worth living, they must be devoid of most of what gives life personal value – value to the person whose life it is. If we cannot avoid this New Conclusion, this undermines what most of us believe when we consider overpopulation.[13]

Parfit's lament, however, seems misplaced. It begs the question to complain against a total utilitarian theory that the lives that result from its being implemented are deprived of personal value. For some of the lives resulting from it wouldn't even have been brought into existence if the alternative society (having a small population with a high average utility) had been opted for. Thus although lives in this new society may not have as much value per life as alternative societies, those people who are now existing (and wouldn't have otherwise) won't feel their lives are deprived of personal value for they wouldn't even have been living under a different regime.

Though Parfit's above complaint against classical utilitarianism seems to beg the question, we can ask whether it is empirically plausible to assume that total utilitarianism would have people existing at the restricted level. I think Parfit's concern that classical utilitarianism would result in masses of individuals living at the subsistence level is misguided. Individuals in their own lives don't place such a high premium on continued existence that they sacrifice indefinite amounts of quality for increased quantity. Many individuals smoke cigarettes and drink alcohol -- activities which surely decrease the quantity of life. Other individuals eat too much and don't exercise enough or drive too fast in their cars. If individuals in their own lives opt to maximize happiness by sacrificing some of the quantity for more quality, it seems unreasonable to think that a classical utilitarian position would require individuals to make these and even more severe sacrifices in the quality of life (so that they end up at the restricted level) so that they can increase the quantity of life of someone else. (It is of course possible that these decreases in quality in a person's life would bring about a greater increase in quantity than his own extension of life but the limits to this would still seem to place the quality of life above the restricted level.) People aren't willing to make sacrifices that would put them near the subsistence level even if this would greatly increase the quantity of their own life and so it seems mistaken to believe that they would be required to make these same sacrifices to increase the quantity of someone else's life. Thus the concern that total utilitarianism would have people subsisting at a very low level seems mistaken to me.[14]

SUMMARY AND CONCLUSION

This paper argues that a total utilitarianism position would be chosen if the constraints of the original position are consistently adhered to. A straightforward generalization of Harsanyi's characterization of the original position suffers both from the logical flaw of having the states of nature be dependent upon the actions of individuals as well as the morally objectionable feature that individuals presumed to exist will be able to be partial to their own situation in an original position framework which is designed to insure an impartial choice of principles. A more sophisticated extension of Harsanyi's characterization avoids both these problems and results in total utilitarianism being chosen in the original position.

The "justice as equal opportunity" position is argued against by pointing out that its standard of an equal resource base must make reference to the size of the population to have any intuitive appeal and the equal opportunity standard is incomplete in that it just treats as indifferent all policies which preserve equal opportunity; it doesn't opt for a greater population if all the members will have good lives and the standard will still be satisfied (and doesn't give any guidance as to permissible tradeoffs if the larger population won't completely satisfy the standard). The equal opportunity standard is fundamentally incomplete in this respect. Moreover the equal treatment standard inherits the implausible features that have plagued Rawls' difference principle.

A total utilitarian position does not result in Parfit's "Repugnant Conclusion"; his repugnant conclusion overlooks the fact that we may well be in a society where the marginal happiness curve at some point equals 0 and that beyond this point further increases in population cause total happiness to decrease. There is also no good reason to believe that a total utilitarian position would have people existing at the subsistence level. Although a total utilitarianism position will make more demands upon existing individuals than average utilitarianism, the theory should not be rejected for this reason. The total utilitarian view has the intuitively acceptable consequence of not having people existing at the subsistence level as well as following from a consistent and rigorous formulation of the original position.

yields an individual in the original position utility

$$W(A) = (1/Q)\sum_{j=1}^{(S\&R)(A)} U_j(A) + \sum_{j=1}^{Q} Q - (S\&R)(A)U_j(A).$$

Because I have assumed the utility of not being born is zero, the second term in the above summation is zero. If it was not assumed to be zero, the second term is not zero. The conclusion would not follow, however, that the average happiness of *existing* people should be maximized. Since Q is constant, both the total and average happiness of individuals in the original position will be maximized by the choice of A which maximizes the above summation. However, the conclusion (believed by Rawls and indicated by a straightforward generalization of Harsanyi's theorem) that the average utility of existing individuals would be the most plausible utilitarian position chosen in the original position does not follow. Thus our assumption that the utility of unborn individuals is zero is not necessary for our conclusion that the principles chosen in the original position would be ones which maximize the total and average utility of individuals in the original position and *not* the average and total utility of existing people. What is crucial for this conclusion is that the people who do (and will) exist in the world *not* be automatically indentified with the people in the original position; the two arguments against the identification have been given in the text.

REFERENCES

I would like to gratefully acknowledge the financial support provided by Toby Page's N.S.F. grant. I have also benefitted from many conversations with Toby on this subject. This paper has also benefitted from the comments of Bruce Cain and Burt Klein.

1. See John Rawls' *A Theory of Justice*, pp. 17-22.

2. Burt Klein put the compatiblist position in this rather pithy way.

3. See John Harsanyi's "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy* 1955.

4. Rawls' attempt to handle the future generations problem by assuming that all individuals in the original position care about future generations is rather ad hoc and has been rightly criticized by David Richards in "Contractarian Theory, Intergenerational Justice, and Energy Policy," University of Maryland Working Paper EP-1.

5. See the discussion and proof of the expected utility theorem in Roy Radner's and Jacob Marschak's *An Economic Theory of Teams* as to why the states of nature must be independent of actions.

6. I have for simplicity made an assumption that the utility of not being born is zero. Without this assumption, a given social policy A

7. c.f. *A Theory of Justice*, p. 162.

8. Note that this solution to the future generations problem allows discounting of their future benefits only to the extent that there is exogenous uncertainty about their receiving these benefits.

The proposal in this section assumes that individuals can make meaningful interpersonal comparisons of utility. For a discussion of how these are possible, see Harsanyi, op. cit., and Stephen Selinger's *A Defense of a Preference Based, Long Run Utilitarianism*, unpublished Ph.D. dissertation, Princeton University, 1981.

9. This proposal has been put forward by Talbot Page, "Intergenerational Justice as Opportunity," and Brian Barry, "Justice Between Generations."

10. c.f. "Future Generations: Further Problems," *Philosophy and Public Affairs* 1982.

11. See ibid., pp. 162-163. Parfit's reply to the charge that this added group ends up harming the previously existing groups seems rather unconvincing to me. It does not enable him to consider where we have to actively consider a population policy rather than passively judge how we would have liked a history of the world to have developed.

12. Ibid., p. 168.

13. Ibid., p. 168.

14. The objection is also sometimes made that utilitarianism would have present generations existing at a very low level of satisfaction because by these generations saving more, later generations will be able to be much better off due to this saving being productively invested over a number of years.

Two points should be kept in mind, however. First, the tendency to have earlier generations starve so that later generations will be better off is mitigated to a large extent by considerations of diminishing marginal utility; because these earlier generations will have less goods, they will derive more utility from these goods than later generations. Second, to the extent that considerations of the additional productivity from earlier generations saving outweigh considerations of diminishing marginal utility, the utilitarian has the same convincing reply he has in the intragenerational case. The utilitarian will recommend that if a very sick rich person will benefit more from a drug than an only mildly ill poor person, the drug should be given to the rich person as it will do more good there. Similarly, if later generations, who will already be better, will benefit more from the goods saved by earlier generations than the earlier generations would by spending them, the goods should be so saved.

REFERENCES

Barry, Brian. "Justice Between Generations."

Harsanyi, John. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility." *Journal of Political Economy* 1955.

Page, Talbot. "Intergenerational Justice as Opportunity."

Parfit. "Future Generations: Further Problems." *Philosophy and Public Affairs* 1982.

Radner, Roy and Marschak, Jacob. *An Economic Theory of Teams* (Yale University Press, 1974).

Rawls, John. *A Theory of Justice* (Harvard University Pess, 1973).

Richards, David. "Contractarian Theory, Intergenerational Justice, and Energy Policy." University of Maryland Working Paper EP-1.

Selinger, Stephen. *A Defense of a Preference Based, Long Run Utilitarianism*, unpublished Ph.D. dissertation, Princeton University, 1981.

ASSET RETURNS OF REAL ESTATE IN DIFFERENT SOCIO-ECONOMIC AREAS AND INFLATION

I. Introduction

There have been a number of studies, both theoretical and empirical, on the effects of inflation on assets. Several of these studies have looked at real estate in particular and most of them have treated real estate as a homogenous commodity. However, it is reasonable to ask whether real estate should be so treated. The commodity of real estate is a differentiated one that can take on many different characteristics, the most prominent of which is the property's location. Thus, a research strategy that studies real estate in different socio-economic areas and does not combine them together will better reflect the diffentiated nature of the product. There are three reasons for separating real estate into different socio-economic locations when studying the effect of inflation on real estate values. One reason is that the location of a piece of real estate is undoubtedly correlated with the marginal tax bracket of homeowners in an area. A house in a relatively poor socio-economic area will be surrounded by other houses who have owners who are relatively poor. Inflation increases the marginal tax brackets of individuals and since home mortgage payments are deductible, real estate values in different areas may be affected differently by inflation. The relative after-tax cost of housing may change in different areas as a result of inflation and this may affect the price of housing in the different areas. A second reason for studying real estate in different socio-economic areas is that one of the characteristics affecting the value of a home is the bundle of local public services, e.g., local schools, police protection, road repairs, etc. that are provided with the house. The relative amounts of public services provided to each area could possibly vary with the inflation rate. This would result in inflation impacting differently upon the different areas since one of the determinants of housing values is the amount of local services provided with the house. In effect, by combining areas with different local services, we could possibly be introducing a misspecification into our model. Local services are a determinant of housing prices and by neglecting these services we would be omitting a relevant variable. A third reason for studying real estate in different areas is that there is no reason a priori to combine real estate in different locations without testing to see if this pooling is permissible. We should no more combine the assets of real estate in different areas than we should combine the assets of gold and common stocks. It may in fact be permissible to treat the different socio-economic areas as one area in certain cases but the permissability of this pooling should be tested.

The goals of this paper are to test for the effects of inflation on rates of return on residential owner occupied in different socio-economic areas. Measures of expected and unexpected inflation will be defined and discussed. Rates of return from owning real estate in different areas will then be computed. These rates of return will then be regressed upon the measures of expected and unexpected and inflation. A test will then be performed to determine whether it is possible to pool the rates of return in the different areas. We will also be able to examine whether expected or unexpected inflation had a larger impact upon returns in the different areas.

II. Literature Review

There has been a significant amount of study concerning the return of common stocks and treasury bills when there are changes in the rates of expected and unexpected inflation. Lintner (1975) was one of the first to observe the negative relationship between expected inflation rates and the return to common stocks. Later this relationship was confirmed by Jaffe and Mandelker (1976), Body (1976), Nelson (1976), Fama and Schwert (1977), and Cohn and Lessard (1981). In order to study the effects of inflation upon real estate returns, it is necessary to develop a measure of housing prices so that rates of return can be calculated. We thus need to review some of the literature regarding the determinants of housing prices.

There have been several studies concerning the determinants of housing values. These studies have been primarily cross sectional within the same time period. Thus Bailey (1966) examined the influence of a neighborhood's racial composition and population density on housing prices. Grether and Mieszkowski (1974) analyze both the particular characteristics of a house that determine its value, e.g., the square feet of the house, age of the house, number of bathrooms, quality of carpet, etc. as well as the local characteristics that determine its quality, e.g., the pupil/teacher ratio in the schools, traffic flow, racial composition, etc. More recently, Noland (1979) has studied the different determinants of owner occupied and rental housing value. Grether and Mieszkowski used both linear and semi-log models and found that the estimated coefficients of both models were rather similar. Grether and Mieszkowski dealt somewhat with the time series nature of their data in that they computed rates of return on land values; the primary focus of both of these papers, however, was on the determinants of real estate value from a stationary viewpoint (although to be able to assemble the data base, it was necessary to pool houses over a certain length of time.) The articles did not primarily deal with any of the determinants of residential housing value over time, which is what we shall do when we study the impact of inflation on different socio-economic areas over the last fifteen years.

There have been several studies -- both theoretical and empirical -- of the effect of inflation upon the demand for housing. On the theoretical side, Titman (1983) modeled the behavior of high and low income individuals in response to an increase in anticipated inflation. He found that with a fixed supply of housing, higher income individuals will increase their consumption of housing while lower income individuals decrease their consumption of housing.

Titman's results, however, are not really applicable to the situation we are studying. He assumes that there is just one, homogenous commodity called housing and that its price increases are the same in all areas. Thus, it does not tell us whether to expect that the price will go up more or less in "good" areas with high income individuals or "bad" areas with low income individuals.

On the empirical side, Kearl (1979) found that an increase in the rate of anticipated inflation reduces housing demand and housing prices. Using data from the Federal Reserve Board, Kearl regressed the price of housing upon a constant, the initial quarterly mortgage payment, the cost of capital, the stock of housing, and the income level of households. He found that through its effect on initial mortgage payments, an increase in inflation reduced housing demand and housing prices. Kearl did not investigate the change in housing prices of different socio-economic areas due to an increase in inflation and so his results are not directly relevant to our study.

On the other hand, Rosen (1979) argued that the value of residential real estate should rise in an inflationary environment because the income tax system is not indexed. Because nominal rather than real interest payments are tax deductible, he argued an increase in inflation decreases the after tax cost of capital for homeowners, which in turn increases the demand for housing and increases its real price. Rosen and Rosen (1980) studied a time series of housing prices from *Historical Statistics of the United States: Colonial Times to 1970*. They regressed the proportion of households owning houses upon permanent income, credit availability, and certain demographic characteristics of families. They found that not using the *after* tax cost of capital leads to underestimates of the percentage of homeowners.

Follain (1979) was one of the few to investigate the effect of inflation upon the housing demand of different income groups. He did this only using a cross sectional analysis consisting of part of the 1975 Annual Housing Survey. Follain used a maximum likelihood approach to estimate the demand function of housing. He found that an increase in the rate of anticipated inflation increases the demand for housing among individuals in the 40% or higher tax bracket and reduces the demand for individuals in lower tax brackets. However, he also did not study the behavior of housing prices in high and low income areas over time in an inflationary period. Rather he just studied the behavior on the demand side of the market in a single period of one cross section. There are thus two reasons why Follain's results are not directly relevant to ours. The first is that he considered only one time period and did not analyze a time series. The second is that he only studied the demand side of the market in different socio-economic areas and did not study the behavior of housing prices in different areas. It could be the case that the demand for housing was decreased relative to other areas but that prices in this area rose the same or more than other areas depending upon the impact of inflation upon the supply curves in the different areas. Thus the second main difference between our study and Follain's study is that he studied only the demand side of the market while we are studying prices--which are the result of both the supply and demand sides.

Particularly in view of the fact that the value of the assets in real estate exceeds the value of the assets contained in common stocks, relatively little study has been made of the effect of expected and unexpected inflation on real estate prices. In their 1979 study, Fama and Schwert found that private residential real estate was a complete hedge against expected inflation and a partial hedge against unanticipated inflation. The Fama-Schwert test for the return of an asset as a hedge against expected inflation is to regress the percentage change in the price of the asset against the expected and unexpected rates of inflation. This study will follow the test of the Fama-Schwert model so it is worthwhile discussing their model more fully.

Fama-Schwert define the return on the i-th asset in period t to be the percentage change in the price of the i-th asset in period t.

$$R_{i,t} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} \qquad (1)$$

In a previous study, Fama (1975) argued that the rates on United State Treasury bills were a good measure of the rate of expected inflation. Fama found that if we assume that real rates of return on treasury bills are constant through time, then the changes in the nominal rates of return reflect changes in the expected rate of inflation. Fama argued that if the expected real return on a treasury bill is constant through time, and if the bill market is efficient, the nominal return is equal to the constant expected real return plus the expected inflation rate. Let B_t be the nominal return on a treasury bill which matures at time t. Because this bill is sold at time t-1, the nominal return is known at time t-1. Let $E(\tilde{I}_t \mid K_{t-1})$ be the expected rate of inflation at time t conditioned on our knowledge at time t-1. Let E(i) be the constant expected real return. If the nominal return is equal to the constant expected real return plus the expected inflation rate, we have

$$B_t = E(\tilde{i}) + E(\tilde{I}_t \mid K_{t-1}) \qquad (2)$$

Subtracting and rearranging we have

$$E(\tilde{I}_t \mid K_{t-1}) = B_t - E(\tilde{i}). \qquad (3)$$

If $E(\tilde{i})$ is constant then we can substitute a for it and (3) can then be tested from estimates of (4).

$$\tilde{I}_t = a + bB_t + \tilde{e}_t \qquad (4)$$

where the proposition of (4) is that b=1.0 and $E(\tilde{e}_t / Ksubt - 1) = 0$; that is, if we suppose that the expected real rate of return is constant, all variation in the nominal rate of returnB_t at t-1 reflects variation in the expected value of the inflation rate to be observed at t (or $E(\tilde{I}_t \mid K_{t-1}) E(\tilde{I}_t \mid K_{t-1}.)$ The unexpected value of the inflation rate is then the disturbance term $\tilde{e}_t$.

In their 1975 study, Fama-Schwert found that the regression of the annual inflation rate on the interest rate of a treasury bill which matures at the end of the one year period indicates that the nominal rate on a treasury bill is a good proxy for the annual expected inflation rate. The estimate of b is 1.06 with a standard error of 0.10, R^2 is 0.82, and the residuals do not seem autocorrelated.

Fama's 1975 study is relevant for present purposes because we need some measure of the rate of expected inflation if we are to determine the ability of an asset to serve as a hedge against expected inflation. His study justifies taking the nominal rates on treasury bills as such a measure.

One key assumption of Fama-Schwert's 1975 study is that the expected real returns of treasury bills remain constant. That is, the expected return of real interest rates is expected to remain constant and any variation in nominal rates is to be explained by changes in the expected rate of future inflation. (Note that this assumption of a constant expected rate of interest is consistent with our methodology in chapter two where it was assumed that the ex post real rate of interest could vary. The real rate of interest was there defined to be the difference between the current interest rate and the rate of inflation realized over the past year. This rate can vary depending upon how much (or whether) the nominal rate of interest exceeds the realized rate. This is consistent with holding that the future rate of expected real returns is constant however.)

We might wonder, however, if this assumption of constant expected real rates is valid. For instance, in September 1983, rates on three month treasury bills have been approximately 9% (give or take .5%). The inflation rate in the preceding year has been about 2.4 % on an annualized basis. If we consider real rates of interest to generally be about 2-3% above the inflation rate, then we would have to say that investors expect that the inflation rate is to jump to about 6 to 7 over the next three months. If the real rate is constant (at 2-3%) only an expectation that inflation will jump back to 6-7% over the three month period from Sept. to Dec. would explain the nominal rate of 9%. Given that the inflation rate has been 2.4% on an annualized basis, it may seem unlikely that investors really expect inflation to increase to 6-7% over the next three months. One might rather hypothesize that real expected rates of return have been higher at some times than at others.

The problem with the hypothesis that real rates of return have varied is that it is difficult to explain what would account for the changes in these real rates. The ability of the United States Government to redeem the notes at maturity does not seem to change and so it is dificult to see how a higher risk premium at certain times might account for the different real rates. In the absence of an explanation why real rates of return should be higher, it seems plausible to assume that investors may not have been convinced that inflation will stay down (if rates seem abnormally high.) This assumption of a constant expected real rate of return is crucial for the measure of expected inflation we employ (following Fama-Schwert) and it is not entirely unproblematic; however as yet there does not seem to be a good theory which relates the real rates on treasury bills to both the supply and demand for credit in the market. Thus lacking any good explanation of the factors causing variation in real rates, we will note the possible shortcomings of Fama's method but continue to use it measure expected inflation. The falsity of his assumption of constant expected real rates would cause a misspecification in our model however.

We will thus take the expected rate of inflation during period t to be equal to the nominal interest rate on a treasury bill which sells at the beginning of the t-th period and matures at the end of the t-th period. The unanticipated rate of inflation is then defined to be the difference between the inflation realized ex post and the expected rate of inflation (as measured by the nominal interest rate). If we let B_t be the expected rate of inflation, and D_t be the inflation rate realized ex post, the unanticipated rate is $D_t - B_t$.

In order to estimate the effectiveness of an asset as a hedge against inflation we must see how the return of this asset varies with the expected and unexpected rates of inflation. Thus we will run a regression of the asset return upon the expected and unexpected rates of inflation.

$$R_{jt} = a_j + b_j B_t + c_j (D_t - B_t) + u_{jt} \tag{5}$$

In the above equation R_j stands for the return on real estate in the j-th area and j equals 1, 2, 3, because there are three socio-economic areas; t is the time period and equals 1, 2, ... 27 because we are looking at semi-annual returns over an approximately fourteen year period. The coefficients b_j and c_j measure ability of real estate in the j-th area to serve as a hedge against expected and unexpected inflation.

The Fama-Schwert paper is an interesting and useful one and we will follow its general framework in this study. However with respect to their treatment of the returns from residential real estate, there are two main drawbacks. First, the sample of houses in the Fama-Schwert study consists only of houses that were sold through loans made by the Federal Housing Administration (FHA). In order for a house to qualify for an FHA mortgage, it must be a relatively low priced house. For instance, as of January 1983, a house had to be priced under $108,000.00 in order to qualify for an FHA loan. This causes the houses in the sample of Fama-Schwert to be all of the lower price variety. We might wonder if the return on middle and upper income residential real estate was a similar hedge against inflation.

The second major difficulty with the Fama-Schwert study (which they note) is that the prices they use to construct the rate of return are the prices of the average square foot of housing; the average transaction price of the houses that sold with an FHA mortgage is divided by the amount of square feet in the average house to obtain a price per square foot of housing. What is desired, however, are prices that allow us to calculate a rate of return *net* of both interim costs and benefits associated with holding real estate. The relevant costs are mortgage expenses, property taxes, and imputed maintenance and management fees. The benefits are the imputed rental value of the house and the tax benefits of the deductibility of interest on mortgage payments.[2] These benefits are not given in the data used to construct the FHA Home

Purchase Price Index. Because the Fama-Schwert article uses gross transaction prices (as reflected in the square foot prices) to construct the rate of return, it will not yield an accurate measure of the real economic rate of return.

This study attempts to rectify these two problems in the Fama-Schwert study. The first problem is addressed by looking at the impact of inflation upon real estate returns in both middle income and upper income areas as well as lower income areas. The second problem is rectified because the prices used to compute the rates of return reflect all of the relevant economic costs and benefits that go into buying and owning a house and not only the costs and benefits captured by the transaction prices of the buying and selling of the house. The transaction prices are adjusted in the next section to reflect all of the interim costs and benefits associated with owning a house. However, for the purposes of comparison, we shall also calculate a rate of return based upon the method of prices which Fama-Schwert used. Their measure of the price of housing was the average price per square foot of a house. We will also calculate rates of return based upon these prices and compare them to rates of return based upon all of the relevant interim costs and benefits.

II. The Data

The data in this study come from the Pasadena Board of Realtors Multiple Listing Books. The data were compiled semi-annually – from the first half of 1968 to the second half of 1982 – in volumes that list which houses sold. There are six pieces of information collected on each transaction. These are the price for which the house sold, the number of square feet in the house, the number of square feet on the lot on which the house rested, a rental price at which the house would have rented, the date the house sold, the interest rate on the mortgage of the house when it was bought, and a dummy variable for the area in which the house was located.

The rental price was constructed for each house by regressing the rental price of houses that rented in each period in a given area upon a constant and the square feet of the houses. Rental prices for the houses that sold were then estimated by multiplying the number of square feet in their house times the rental coefficient of house square feet in that period and adding the constant of that quarter.

There are three different areas that are being studied. The first is the area consisting of the northwest area of Pasadena. This area is a lower income area with a high percentage of minorities (primarily black and Latin). The second area consists of part of southeast Pasadena. This is a largely middle income area with a predominantly Caucasian population. The third area is the community of San Marino. This area is upper income and almost exclusively Caucasian. Table I gives the income of homeowners and ethnic compositions of the areas on the basis of data given in the 1960, 1970, and 1980 censuses.

TABLE I

% of Minorities	San Marino	Northwest Pasadena	Southeast Pasadena
1960	00.6	47.1	00.5
1970	00.7	69.1	2.0
1980	8.2	76.2	13.7

Income	San Marino	Northwest Pasadena	Southeast Pasadena
1960	17,055	5,210	7,890
1970	24,096	9,431	16,710
1980	51,011	15,741	29,034

There were some items that were not recorded for every transaction. For instance, sometimes the square feet of the houses that sold were not listed and an educated guess had to be made about the size of the house. This procedure seemed preferable to reducing the size of the sample. There is no reason to think that there was any systematic bias in the guesses. Thus the estimates of the coefficients that result from combining transactions with partial data with those with complete data should be unbiased.[3]

One possible bias in the model exists because the houses which were selected in the Pasadena areas were always selected from the lowest prices which sold while those which were selected from the San Marino area were selected from the highest price houses which sold. The reason for doing this is to make the neighborhood effect in each of the areas more uniform. The higher priced houses in the southeast Pasadena area are in a significantly nicer area than the lower priced houses. The higher priced houses are the estate type houses found in the sample of houses selected from San Marino. Moreover the homeowners living in these houses could reasonably be assumed to have significantly higher incomes than those individuals in the lower priced houses. Thus in order to keep the neighborhood effect and the marginal tax bracket as uniform as possible, houses were selected from the lower priced houses which sold in both Pasadena areas and from the higher priced area in San Marino.

The data used to estimate the income and tax bracket of the homeowners in the areas came from the Census Bureau data of 1960, 1970, and 1980. The overall growth in income that occurred in each decade was assumed to have taken place according to the rate of inflation. Thus, for example, if the growth in income over the decade in an area was 50%, and inflation was 10% in the first year, it was assumed that wages rose 5% in the first year. The income tax bracket of individuals in each area was estimated to be the same. It would be desirable to have information about the tax rate of each individual, but such information could not be located. The Annual Housing Survey contains information about the tax brackets of individual homeowners, but does not contain such information over a fifteen year time period in any particular cross sections. The Multiple Listing Books contain information over a fifteen year period, but do not have the individual's tax rates. The City of Pasadena also did not have any data of income growth beside that of the census.

The transaction price for which a house sold was modified according to the following formula that adjusts the house according to the positive or negative cash flow the house produces.

$$P_{it}^{A} = P_{it} + NM[(1 - T_{it})(.85P_{it}/1000 \, x \, ppt_{it} + (.02/12)P_{it}) - .94rental_{it}] \quad (6)$$

Particularly to those unfamiliar with the business of buying, renting, and then selling houses, the above expression (6) may appear rather mysterious and ad hoc. In fact, however, it is the natural way to adjust the nominal selling price of a house to reflect the monetary effect of carrying the house for a year. We will be studying carrying costs over a one year period because we are interested in the rate of return over a one year period and so must calculate the carrying costs over this length of time. If we take the actual price someone paid for a house to be his real economic cost of holding the house for a year, we would be making a mistake. For this cost of buying the home can be greater or smaller than the relevant economic cost depending upon whether or not the house yielded a positive or negative cash flow over the time period in question. The situation is analogous to determining the relevant price of a security on the stock market. If one stock yielded higher dividends than another stock even though both sold for the same prices over a one year period, they would not have the same relevant economic prices. The one with the higher dividend stream would have had a lower economic price at the beginning of the period because its net cost was lower than the other one (or alternatively one could say it had a higher economic price at the end of the period). Similarly, if two houses sold for the same prices over a one year period but one house had a higher positive cash flow than the other one, we would want to say that the house with the higher cash flow had a lower economic cost than the other house. This is the motivation for adjusting the data according to (6). The formula for adjusting the data according to (6) is similar to the formula used in chapter two except that the tax effects are now included in the formula.

Let us go through expression (6) term by term. P_{it} is the price the i-th house sold for in period t; NM stands for the number of months the house was owned and will be equal to 12 because we will study the returns from holding patterns of one year; T_{it} is the marginal tax bracket of the owner of the i-th house in the t-th period and since the interest on house payments is tax deductible, $1 - T_{it}$ is the after tax percentage of mortgage payments the homeowner pays; ppt_{it} is the price per thousand dollars that an owner would have to pay on a loan amortized over thirty years -- it is multiplied by .85 of the price because it assumed that on average individuals make a down payment of 15%; $(.02/12)P_{it}$ is the monthly portion of the yearly property tax if the property tax is figured at a rate of 2% of the sales price;[4] $rental_{it}$ is the rental value of the i-th house in period t.

The rental value is multiplied by only .94 to allow for the implicit maintenance and managerial services which a homeowner must face. The rental value is not multiplied times the marginal tax bracket of the average homeowner in the area, because the homeowner does not have to pay any taxes on the implicit income he receives from living in the house. It is sometimes said that the deductibility of home interest payments provides the taxpayer with a subsidy. This is really not the case, however. For a landlord who rents property to tenants also is able to deduct interest payments. Rather the real tax benefit to homeowners resides in the fact that the homeowner does not have to declare as taxable income the benefit he receives from staying in the house. This is why the rental value of the house is not multiplied by the homeowners marginal tax bracket.

Perhaps a word of explanation is in order about the inclusion of the rental value of the house in the adjustment of the house. Even though most of these houses were owner occupied, the owner was receiving a benefit in kind from staying in the house which must be figured into the adjusted price; thus it is necessary to enter in the rental value of the house in constructing prices even for owner occupied housing.

The adjusted prices obtained from (6) are then regressed upon the characteristics of the houses to obtain values for the regression coefficients of house square feet and lot square feet.

After the transaction prices have been modified to incorporate the positive or negative cash flow from holding them for a twelve month period, the houses in each area and each quarter were pooled and the regression in (7) was run.

$$P^A_{ijt} = a_{jt} + a_{1jt}HF_{ijt} + a_{2jt}LF_{ijt} + u_{ijt} \qquad (7)$$

and where HF_{it} and LF_{it} stand for (respectively) the square feet of the i-th house and lot in the j-th section in the t-th period.

We will also experiment with a functional form in which we regress the price of a house upon a constant, the log of square feet of the house and the log of the square feet of the lot as in (8). The reason for experimenting with these different functional forms is because the price of a house might be a linear function of the square feet of the house and lot on which it rests; or on the other hand, the price of a house might not increase as fast with each additional square foot of house size and lot size that is added. For instance, the price of a house might increase more if its size is increased from 1000 square feet to 1500 square feet than if its size is increased from 5000 square feet to 5500 square feet. If the derivative of price with respect to size is not constant but rather is positive but decreasing, we would want to use a functional form such as (8).[5]

$$P^A_{ijt} = a_{jt} + a_{1jt} logHF_{ijt} + a_{2jt} logLF_{ijt} + u_{ijt} \qquad (8)$$

The regression coefficients a_{1jt} and a_{2jt} that are obtained from (7) (or (8)) are then multiplied by the mean value of the square feet of the house in each area over the fifteen year period and summed together with the constant to obtain the price of buying a house and holding it for a specified period of time. They are multiplied by the same mean because it would not be sensible to just average the prices of the houses in each period because in some periods there might be large lots or large houses which, if the transaction prices were just averaged, would appear to make the price of housing increase more than it actually did. The regression coefficients are multiplied by a constant number in every period to avoid this problem. The buying prices are thus computed as in (9).

$$P^B_{jt} = a_{jt} + a_{1jt}\overline{HF}_j + a_{2jt}\overline{LF}_j \qquad (9)$$

The buying prices as constructed represent not just the cost of buying an average size house situated on an average size lot in a given area at a given time but also include the costs of carrying the house for a one year period. These are the relevant prices we want when constructing rates of return.

The selling prices for each area are computed in a similar way except that the cash flow of the house is not figured into the adjusted price. When one buys a house and holds it for a period of time, the positive or negative cash flow must be figured into the price but when one sells a house, all one gets are the proceeds from the transaction price. Thus in determining selling prices, the unadjusted transaction prices were regressed upon the house characteristics to determine regression coefficients of house square feet and lot square feet (and also the coefficients of the logs of house feet and lot feet as in (11)).

$$P_{jt} = \tilde{a}_{jt} + \tilde{a}_{1jt} HF_{ijt} + \tilde{a}_{2jt} LF_{ijt} + u_{ijt} \qquad (10)$$

$$P_{jt} = a_{jt} + a_{1jt} log HF_{ijt} + a_{2jt} \log LF_{ijt} + u_{ijt} \qquad (11)$$

These regression coefficients are then multiplied times their respective mean values as in (9) to arrive at the selling price of housing in each period as in (12).

$$P^S_{jt} = a_{jt} + \tilde{a}_{1jt} \overline{HF_j} + \tilde{a}_{2jt} \overline{LF_j} \qquad (12)$$

III. The Hypothesis to Test

The hypothesis we wish to test is that the different socio-economic areas have served equally well as hedges against inflation. To test this hypothesis, we must calculate the rate of return from owning real estate in the different areas and regress this rate of return upon the inflation rate.

The rate of return to real estate in the different areas is defined as the percentage change in the price of real estate in going from one period to another. We will thus define R_{jt} as the return on real estate in the i-th area of the t-th period and it is equal to the percentage change in the price of real estate in the i-th area in going from the t-1 to t period.

$$R_{j,t} = \frac{P^S_{j,t} - P^B_{j,t-1}}{P^B_{j,t-1}} \tag{13}$$

The buying price is subtracted from the selling price because the buying price has been defined to include not only the costs involved in the transaction price but also all of the interim costs and benefits associated with holding the house.

The returns from real estate will then be regressed upon both the expected and unexpected rate of inflation. As we said in our introduction, we will follow Fama in taking the expected rate of inflation during period t to be equal to the interest rate on a treasury bill which sells at the beginning of the t-th period and matures at the end of the t-th period. The unanticipated rate of inflation is then defined to be interest rate realized ex post and expected (ex ante) interest rate. If we let B_t be the expected rate of inflation, and D_t be the inflation realized ex post, the unanticipated inflation rate is $D_t - B_t$.

There have been some exogenous shocks that may have changed the value of residential properties in different areas of Pasadena. In particular, the two exogenous shocks we will examine (and control for if significant) are the construction of a new freeway built in a portion of the northwest area of Pasadena and a court ordered school busing that changed the previous pattern of local neighborhood schools in 1970. We will test for the statistical significance of price changes due to these exogenous shocks by using a dummy variable to test if the rates of return are abnormally higher or lower in the years the freeway was built or busing was implemented. When the rates of return are regressed upon the expected and unexpected rates of inflation as in (11), there will also be a dummy variable equal to one if the year is 1970 or later (the first year of court ordered busing) and the cross sections are either of the Pasadena areas and zero otherwise; and one if the year is 1972 (the year the freeway was begun) and the area is northwest Pasadena.

The reason for studying the effect of the highway and busing upon the prices of the different areas is to determine if factors other than inflation may have impacted upon the areas differently and be responsible for some areas appreciating more than others. If these other omitted factors were correlated with the inflation rate, a finding that one area was a better hedge against inflation might be in reality due to the omitted variable (that was correlated with the inflation rate) and impacted differently upon the different areas.

The reason I chose the years busing was actually implemented and the freeway was actually begun and finished for the dummy variables is because it would have been very difficult and somewhat arbitrary to pick another year, for instance, the year in which information about the freeway and busing first became available to the market. The freeway had been studied for a long time and the possibility of busing had existed for many years before it was actually ordered.

It is possible to object that the market may have absorbed the information about busing and the freeway before they actually took place and so using the year in which they were actually implemented to measure their impact may be a mistake in that the information may have already been incorporated by the market. However, since it is practically impossible to say when the market first started incorporating rumors about the possibility of busing and the freeway, it seemed best to use the year in which they were actually implemented (or started to be implemented in the case of the freeway being built). Even if this information had already been taken into account by the market, there was always the chance that the freeway construction or busing order might have been stopped at the last minute. Thus, by testing for the significance of the dummy variable in the year of implementation, we are able to assume that the market received the new information that there was no chance that the freeway construction or busing order would not be implemented.

To summarize, in addition to regressing the rates of return upon the expected and unexpected inflation rates, we will also regress them upon the two dummy variables previously discussed. These are (respectively) 1 if the return occurred in 1970 in either of the Pasadena areas (the year of school busing) and 0 otherwise; and 1 if the return is in 1972 and the area is northwest Pasadena (the year the freeway was built) and 0 otherwise.

Both of these dummy variables turned out to be insignificant. Thus we will not use them as explanatory variables in the prices we use to construct rates of return.

The regression to estimate the effect of anticipated and unanticipated inflation (and the other variables we are controlling for) on the rate of return from real estate is thus

(14) $R_{jt} = a_j + b_j B_t + c_j (D_t - B_t) + u_{jt}$

$j = 1, 2, \cdots 27$

In the above equation j stands for the j-th area and equals 1, 2, 3, because there are three socio-economic areas; t is the time period and equals 1, 2, ... 25 because we are looking at semi-annual returns over an approximately thirteen year period.

The ability of the different investments to serve as a hedge against inflation is then measured by the coefficient b_j; the ability of an asset to serve as a hedge against unanticipated inflation is measured by the coefficient c_j.

In order to test the hypothesis that the different socio-economic areas are affected to the same extent by inflation, we need to also run the regression that pools the different areas. The following is this regression.

$$R_t = a + bB_t + c(D_t - B_t) + u_t \qquad (15)$$

A test of whether inflation affects the real estate values in different socio-economic areas equally is then to compare the value of the unrestricted residual sum of squares obtained from (14) with the value of the restricted sum of squares obtained from (15). A high F value of this ratio will allow us to reject the hypothesis that inflation affects different socio-economic areas to the same extent. The F ratio is

$$F = \frac{S_2 - S_1/r}{S_1/n - k} \tag{16}$$

where S_2 stands for the restricted residual sum of squares obtained from (15), S_1 stands for the unrestricted residual sum of squares obtained from (14), r stands for the number of restrictions and equals 6, n is the number of observations and equals 75, and k is the number of explanatory variables in the unrestricted model and equals 9 (a constant, expected, and unexpected inflation for each of the three areas.) The hypothesis we are testing is thus:

$$\begin{aligned} H^o = a_1 &= a_2 = a_3 \\ b_1 &= b_2 = b_3 \\ c_1 &= c_2 = c_3. \end{aligned}$$

That is, we are testing whether the constant terms and coefficients of expected and unexpected inflation were the same in the three areas.

IV. RESULTS AND CONCLUSION

The principal results of this study are contained in Table 1 through 12. Tables 1 - 3 show the dependent variable of the return from housing when one year holding costs are incorporated into the price of a house; a price per square foot is then used to calculate the rate of return from housing. The price per square foot is the measure of housing prices that Fama-Schwert used except that they did not take into account the costs and benefits of carrying the house. Tables 1 - 3 show that the coefficients of expected and unexpected inflation for San Marino, northwest Pasadena, and southeast Pasadena are (respectively) 2.20, 1.54; 2.68, 3.58; .88,and 3.40. The coefficient on unexpected inflation was significant at the 95% level in both Pasadena areas but not significant in San Marino. The coefficient of expected inflation was not significant in any of the areas.

The pooled regression shown in Table 4 results in coefficients for expected and unexpected inflation of 1.92 and 2.84. The coefficient of expected inflation is not significant at the 95% level (though it is at the 90% level.) The coefficient of unexpected inflation is significant at the 95% level. The critical value at the 95% level for rejecting the hypothesis that the coefficients of expected and unexpected and inflation are equal in the different areas is 2.50 and the value of the Chow test in equation (13) is .82. Thus we can accept the hypothesis that the constant terms and the coefficients of expected and unexpected inflation are the same in the different areas if we use rates of return based upon a price per square foot of housing.

Tables 5 - 8 show the results when rates of return are based upon prices which are constructed upon semi-annual coefficients of lot feet and house feet as in Equation (6). The coefficients for San Marino, northwest Pasadena, and southeast Pasadena are respectively -.15, 1.65; -.19, 2.84; 4.51, and 1.66. The coefficient of unexpected inflation was not significant in either Pasadena area but was significant in San Marino. The coefficient of expected inflation was not significant in the San Marino area or northwest Pasadena area but was significant at the 95% level in southeast Pasadena.

The results of running the pooled regression are given in Table 8. The coefficients of expected and unexpected inflation were 1.45 and 2.05. Only the coefficient of unexpected inflation was significant (although the coefficient of expected inflation was significant at the 90% level.) The value of the Chow test given in (13) is equal to 1.55. The critical value for rejecting the hypothesis that the coefficients of expected and unexpected inflation are equal in the different areas is 2.50 at the 95% level so we can accept the hypothesis that they are equal.

Tables 9 - 12 show the results of using prices which are based upon coefficients derived from regressing prices upon the log of house feet and log feet and then multiplying times the log of house feet and lot feet (as in Equation (11)). The coefficients of San Marino, northwest Pasadena, and southeast Pasadena for expected and unexpected inflation are (respectively) None of the coefficients of expected inflation were significant at the 95% level but the coefficients of unexpected inflation were significant for northwest Pasadena and San Marino.

The results of the pooled regression are given in Table 12. The coefficients of expected and unexpected inflation were 2.08 and 2.60. The coefficient of expected inflation was not significant but the coefficient of unexpected inflation was significant. The value of the Chow test given in (13) was 1.46. The critical value for rejection at the 95% level was 2.50 so we can accept the hypothesis that the coefficients of expected and unexpected inflation were the same in the different areas.

The conclusion that the coefficients of expected and unexpected inflation were the same in different areas thus seems to be robust with respect to the measure of real estate prices with which we use to construct rates of return. Given that some of coefficients of expected and unexpected inflation were not significant, we might wonder, however, if the coefficients of expected and unexpected inflation are both zero and that they don't influence the rate of return on real estate at all. I have tested this hypothesis for each of the pooled models contained in tables 4, 8, and 12. The critical value for accepting the hypothesis that expected and unexpected inflation influence the rate of return is 2.37. The values in tables 4 and 8 are (respectively) 4.42 and 2.61. However the pooled model of table 12 fails to meet the critical value and only achieves a value of 2.34. The differences between the constant terms and coefficients in the different areas were statistically insignificant regardless of whether we used prices based on a price per square foot of housing or prices based upon coefficients multiplied by the mean house feet and lot feet.

Moreover the hypothesis that none of the explanatory variables influence the dependent variable was tested and rejected. However the regressions of rates of return prices based upon coefficients obtained from the logs of house feet and lot feet turned out not to be influenced by the rates of expected and unexpected inflation.

One significant way in which my results differ from those of the Fama-Schwert study is that the coefficient of unexpected inflation is generally larger than that of the coefficient of expected inflation and the t-statistics are also generally larger in the different models. The coefficients of expected and unexpected inflation in the Fama-Schwert model were 1.19 and 0.31 and the t-statistics were 7.4 and 2.8. The history of real estate in Southern California seems to bear out the fact that it is unexpected inflation which has had a more significant impact upon increases in housing prices (and thus upon rates of return) than expected inflation. Unexpected inflation was rather high in the years of 1976-1979 when the realized rate of inflation was often as high or higher than expected inflation as measured by the rates on treasury bills. These were also the years of most rapid appreciation in housing prices. Expected inflation has been high in the years 1980-1982 when rates on treasury bills soared. Yet even though there was much expected inflation by this measure, there was little appreciation in houses. Unexpected inflation was negative during much of this time as interest rates were much higher than the realized rate of inflation. Thus the low (negative)ship level of unexpected inflation was closely correlated with the low returns of 1980-1982.

FOOTNOTES

(1) One possibility is that an increase in the size of the federal deficit may cause higher real returns on treasury bills. The evidence on this is hardly conclusive, though.

(2) One might think that there is some intrinsic value to homeownership above and beyond that yielded by the rental value of the home. Such value would be difficult if not impossible to measure so it is ignored here.

(3) I borrow this argument about missing data from Grether and Mieszkowski (p. 127). See their article for a more extended discussion.

(4) This was changed to 1% of the sales price after Proposition 13.

(5) This functional form also has the advantage of not using any more degrees of freedom than the linear model.

(6) It would be desirable to know the impact of students' scores on standardized tests on real estate values. Pasadena, however, did not keep track of scores in a systematic way that would allow us to compare students' scores in a continuous series over the time period we are studying. There are two reasons for this. The first is that the type of test was changed in 1975. The second is that before 1975, the dates on which tests were

administered were not kept constant across the different years. Students will invariably do worse on tests that are administered in the very beginning of the year and better later in the year. Because the tests were not administered at standardized times over this time period, comparisons of studnts' scores over this period are not particularly meaningful. Mr. Bibiani, who is in charge of testing for the city, said that scores generally fell for all ethnic groups from 1970 - 1975; these were the first five years of school busing. Since 1975, the test scores have been rising for all ethnic groups. The table below shows this.

Median Percentile Scores
1975 through 1981
CTBS and CAT

	1975	1976	1977	1978	1979	1980	1981
District	45	49	49	51	53	54	56
Spanish-surnamed	35	40	40	41	45	49	51
Black	27	31	32	36	42	44	46
White	65	69	70	72	69	72	76
Others	68	71	69	73	74	76	78

These figures are based on weighted median standard scores across all grade levels and subjects. They present an accurate depiction of trends within the district but do not conform to rigorous statistical practice.

The test results in San Marino are kept confidential and not released.

(7) The test of the hypothesis that $B_1 = B_2 = 0$ is given by

$$F = \frac{R^2 \; x \; n-k-1}{1-R^2 \; x \; k}$$

which has an F distribution degrees of freedom k,n-k-1. See Madalla, p. 121.

(8) Lenders may have had different expectations about future inflation than borrowers; perhaps this is why for several years unexpected inflation was high and housing prices increased. If this is so, then the expectations of borrowers would be somewhat more complicated than our model which depicts the expectations of both borrowers and lenders as reflected in the same treasury bill rates. Another possibility is that lenders were taking a longer term view of inflation than borrowers were. In this case, the relevant expectations of lenders might be different than the relevant expectations of borrowers.

BIBLIOGRAPHY

Bailey, M. J. "Effect of Race and of Other Demographic Factors on the Values of Single-Family Homes." *Land Economics* May 1966.

Body, Zvi. "Common Stocks as a Hedge against Inflation." *Journal of Finance* May 1976.

Cohn, Richard and Lessard, Donald. "The Effect of Inflation on Stock Prices: International Evidence." *Journal of Finance* May 1981.

Fama, Eugene. "Short Term Interest Rates as Predictors of Inflation." *American Economic Review* June 1975.

Fama, Eugene. "Asset Returns and Inflation." *Journal of Financial Economics* November 1977.

Fama, Eugene and Schwert, William. "Inflation, Interest and Relative Prices." Working Paper 7720, University of Rochester, May 1977.

Follain, James. "Does Inflation Affect Real Behavior: The Case of Housing." *Journal of Financial and Quantitative Analysis* November 1979.

Grether, David and Mieszkowski, Peter. "Determinants of Real Estate Values." *Journal of Urban Economics* April 1974.

Jaffe, Jeffrey and Mandelker, Gershon. "The Fisher Effect for Risky Assets: An Empirical Investigation." *Journal of Finance* 31 May 1976.

Kearl, James. "Inflation, Mortgages and Housing." *Journal of Political Economy* June 1979.

Lintner, John. "Inflation and Security Returns." *Journal of Finance* 30 May 1975.

Madalla, G.S. *Econometrics*, McGraw-Hil, Inc. 1977.

Nelson, Charles. "Inflation and the Rates of Return on Common Stocks." *Journal of Finance*.May 1980

Noland, Charles. "Assessing Hedonic Indexes for Housing." *Journal of Financial Economics* November 1979.

Rosen, Harvey. "Housing Decisions and the U.S. Income Tax." *Journal of Public Economics* June 1979.

Rosen, Harvey and Rosen, Kenneth. "Federal Taxes and Homeownership: Evidence From Time Series." *Journal of Political Economy* February 1980.

Titman, Sheridan. "The Effect of an Increase in Anticipated Inflation on the Demand for Housing." *Journal of Finance* May 1983.

APPENDIX TABLE 1

ORDINARY LEAST SQUARES

DEPENDENT VARIABLE: OYPER1

SUM OF SQUARED RESIDUALS = 0.807744

STANDARD ERROR OF THE REGRESSION = 0.183456

MEAN OF DEPENDENT VARIABLE = 0.170986

STANDARD DEVIATION = 0.184044

R-SQUARED = 0.828109E-01

ADJUSTED R-SQUARED = 0.637847E-02

F-STATISTIC(2., 24.) = 1.08345

LOG OF LIKELIHOOD FUNCTION = 9.06484

NUMBER OF OBSERVATIONS = 27.

SUM OF RESIDUALS = 0.298023E-07

DURBIN-WATSON STATISTIC (ADJ. FOR 0.GAPS) = 1.3840

RIGHT-HAND VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC
C	0.160660E-01	0.129984	0.123599
O_Y_EXPE	2.20482	1.75013	1.25980
O_Y_UNEX	1.54111	1.20710	1.27671

APPENDIX TABLE 2

ORDINARY LEAST SQUARES

DEPENDENT VARIABLE: OYPER2

SUM OF SQUARED RESIDUALS = 0.780563

STANDARD ERROR OF THE REGRESSION = 0.180343

MEAN OF DEPENDENT VARIABLE = 0.148325

STANDARD DEVIATION = 0.203556

R-SQUARED = 0.275455

ADJUSTED R-SQUARED = 0.215076

F-STATISTIC(5., 21.) = 4.56211

LOG OF LIKELIHOOD FUNCTION = 9.52694

NUMBER OF OBSERVATIONS = 27.

SUM OF RESIDUALS = -0.2233517-07

DURBIN-WATSON STATISTIC (ADJ. FOR 0.GAPS) = 1.1276

RIGHT-HAND VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC
C	-.343304E-01	0.1277779	-0.268671
O_Y_EXPE	2.67879	1.72044	1.55704
O_Y_UNEX	3.58207	1.18661	3.01873

APPENDIX TABLE 3

ORDINARY LEAST SQUARES

DEPENDENT VARIABLE: OYPER7

SUM OF SQUARED RESIDUALS = 1.33021

STANDARD ERROR OF THE REGRESSION = 0.235426

MEAN OF DEPENDENT VARIABLE = 0.113945

STANDARD DEVIATION = 0.250497

R-SQUARED = 0.184650

ADJUSTED R-SQUARED = 0.116704

F-STATISTIC(2., 24.) = 2.71760

LOG OF LIKELIHOOD FUNCTION = 2.33036

NUMBER OF OBSERVATIONS = 27.

SUM OF RESIDUALS = 0.521541E-07

DURBIN-WATSON STATISTIC (ADJ. FOR 0.GAPS) = 1.7998

RIGHT-HAND VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC
C	0.609694E-01	0.166807	0.365508
O_Y_EXPE	0.883334	2.24593	0.393305
O_Y_UNEX	3.40875	1.54905	2.20054

APPENDIX TABLE 4

ORDINARY LEAST SQUARES

DEPENDENT VARIABLE: OYPER

SUM OF SQUARED RESIDUALS = 3.07327

STANDARD ERROR OF THE REGRESSION = 0.198497

MEAN OF DEPENDENT VARIABLE = 0.144419

STANDARD DEVIATION = 0.213131

R-SQUARED = 0.154299

ADJUSTED R-SQUARED = 0.132614

F-STATISTIC(2., 78.) = 7.11559

LOG OF LIKELIHOOD FUNCTION = 17.5701

NUMBER OF OBSERVATIONS = 81.

SUM OF RESIDUALS = 0.268221E-06

DURBIN-WATSON STATISTIC (ADJ. FOR 0.GAPS) = 1.4561

RIGHT-HAND VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC
C	0.142351E-01	0.811993E-01	0.175311
OYEXP	1.92231	1.09328	1.75829
OYUNEXP	2.84397	0.754056	3.77157

APPENDIX TABLE 5

ORDINARY LEAST SQUARES

DEPENDENT VARIABLE: OYREG1

SUM OF SQUARED RESIDUALS = 0.377910

STANDARD ERROR OF THE REGRESSION = 0.125484

MEAN OF DEPENDENT VARIABLE = 0.156723

STANDARD DEVIATION = 0.133804

R-SQUARED = 0.188143

ADJUSTED R-SQUARED = 0.120488

F-STATISTIC(2., 24.) = 2.78093

LOG OF LIKELIHOOD FUNCTION = 19.3193

NUMBER OF OBSERVATIONS = 27.

SUM OF RESIDUALS = -0.154600E-06

DURBIN-WATSON STATISTIC (ADJ. FOR 0.GAPS) = 0.8509

RIGHT-HAND VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC
C	0.172993	0.889095E-01	1.94572
O_Y_EXPE	-0.150188	1.19710	-0.125461
O_Y_UNEX	1.65047	0.825657	1.99898

APPENDIX TABLE 6

ORDINARY LEAST SQUARES

DEPENDENT VARIABLE: OYREG2

SUM OF SQUARED RESIDUALS = 1.23011

STANDARD ERROR OF THE REGRESSION = 0.226394

MEAN OF DEPENDENT VARIABLE = 0.157783

STANDARD DEVIATION = 0.238189

R-SQUARED = 0.166074

ADJUSTED R-SQUARED = 0.965798E-01

F-STATISTIC(2., 24.) = 2.38976

LOG OF LIKELIHOOD FUNCTION = 3.38659

NUMBER OF OBSERVATIONS = 27.

SUM OF RESIDUALS = -0.134110E-06

DURBIN-WATSON STATISTIC (ADJ. FOR 0.GAPS) = 1.7312

RIGHT-HAND VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC
C	0.168400	0.160408	1.04982
O_Y_EXPE	-0.187484E-01	2.15976	-0.868076E-02
O_Y_UNEX	2.84210	1.48962	1.90793

APPENDIX TABLE 7

ORDINARY LEAST SQUARES

DEPENDENT VARIABLE: OYREG7

SUM OF SQUARED RESIDUALS = 0.908804

STANDARD ERROR OF THE REGRESSION = 0.194594

MEAN OF DEPENDENT VARIABLE = 0.172243

STANDARD DEVIATION = 0.208750

R-SQUARED = 0.197873

ADJUSTED R-SQUARED = 0.131029

F-STATISTIC(2., 24.) = 2.96023

LOG OF LIKELIHOOD FUNCTION = 7.47340

NUMBER OF OBSERVATIONS = 27.

SUM OF RESIDUALS = -0.178814E-06

DURBIN-WATSON STATISTIC (ADJ. FOR 0.GAPS) = 1.5847

RIGHT-HAND VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC
C	-0.149478	0.137876	-1.08415
O_Y_EXPE	4.50968	1.85639	2.42927
O_Y_UNEX	1.66256	1.28038	1.29849

APPENDIX TABLE 8

ORDINARY LEAST SQUARES

DEPENDENT VARIABLE: OYREG

SUM OF SQUARED RESIDUALS = 2.78592

STANDARD ERROR OF THE REGRESSION = 0.188989

MEAN OF DEPENDENT VARIABLE = 0.162249

STANDARD DEVIATION = 0.196138

R-SQUARED = 0.947807E-01

ADJUSTED R-SQUARED = 0.715700E-01

F-STATISTIC(2., 78.) = 4.08348

LOG OF LIKELIHOOD FUNCTION = 21.5457

NUMBER OF OBSERVATIONS = 81.

SUM OF RESIDUALS = -0.640750E-06

DURBIN-WATSON STATISTIC (ADJ. FOR 0.GAPS) = 1.3644

RIGHT-HAND VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC
C	0.639714E-01	0.773101E-01	0.827466
OYEXP	1.44692	1.04092	1.39004
OYUNEXP	2.05171	0.717939	2.85778

APPENDIX TABLE 9

ORDINARY LEAST SQUARES

DEPENDENT VARIABLE: LOYREG1

SUM OF SQUARED RESIDUALS = 0.346760

STANDARD ERROR OF THE REGRESSION = 0.120201

MEAN OF DEPENDENT VARIABLE = 0.165464

STANDARD DEVIATION = 0.135104

R-SQUARED = 0.269339

ADJUSTED R-SQUARED = 0.208450

F-STATISTIC(2., 24.) = 4.42348

LOG OF LIKELIHOOD FUNCTION = 20.4806

NUMBER OF OBSERVATIONS = 27.

SUM OF RESIDUALS = -0.819564E-07

DURBIN-WATSON STATISTIC (ADJ. FOR 0.GAPS) = 1.2414

RIGHT-HAND VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC
C	0.114575	0.851664E-01	1.34531
O_Y_EXPE	0.803358	1.14670	0.700584
O_Y_UNEX	2.26825	0.790897	2.86795

APPENDIX TABLE 10

ORDINARY LEAST SQUARES

DEPENDENT VARIABLE: LOYREG2

SUM OF SQUARED RESIDUALS = 1.33136

STANDARD ERROR OF THE REGRESSION = 0.235528

MEAN OF DEPENDENT VARIABLE = 0.173011

STANDARD DEVIATION = 0.269381

R-SQUARED = 0.294355

ADJUSTED R-SQUARED = 0.235551

F-STATISTIC(2., 24.) = 5.00572

LOG OF LIKELIHOOD FUNCTION = 2.31877

NUMBER OF OBSERVATIONS = 27.

SUM OF RESIDUALS = -0.104308E-06

DURBIN-WATSON STATISTIC (ADJ. FOR 0.GAPS) = 1.9088

RIGHT-HAND VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC
C	0.257381E-01	0.166879	0.154232
O_Y_EXPE	2.24662	2.24689	0.999881
O_Y_UNEX	4.82015	1.54972	3.11034

APPENDIX TABLE 11

ORDINARY LEAST SQUARES

DEPENDENT VARIABLE: LOYREG7

SUM OF SQUARED RESIDUALS = 3.12972

STANDARD ERROR OF THE REGRESSION = 0.361116

MEAN OF DEPENDENT VARIABLE = 0.252183

STANDARD DEVIATION = 0.353337

R-SQUARED = 0.358290E-01

ADJUSTED R-SQUARED = -0.445186E-01

F-STATISTIC(2., 24.) = 0.445925

LOG OF LIKELIHOOD FUNCTION = -9.22027

NUMBER OF OBSERVATIONS = 27.

SUM OF RESIDUALS = -0.119209E-06

DURBIN-WATSON STATISTIC (ADJ. FOR 0.GAPS) = 1.9558

RIGHT-HAND VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC
C	0.224075E-01	0.255862	0.875764E-01
O_Y_EXPE	3.19961	3.44498	0.928774
O_Y_UNEX	0.714495	2.37606	0.300706

APPENDIX TABLE 12

ORDINARY LEAST SQUARES

DEPENDENT VARIABLE: LOYREG

SUM OF SQUARED RESIDUALS = 5.26035

STANDARD ERROR OF THE REGRESSION = 0.259693

MEAN OF DEPENDENT VARIABLE = 0.196886

STANDARD DEVIATION = 0.267674

R-SQUARED = 0.822732E-01

ADJUSTED R-SQUARED = 0.587416E-01

F-STATISTIC(2., 78.) = 3.49631

LOG OF LIKELIHOOD FUNCTION = -4.19682

NUMBER OF OBSERVATIONS = 81.

SUM OF RESIDUALS = -0.864267E-06

DURBIN-WATSON STATISTIC (ADJ. FOR 0.GAPS) = 1.7444

RIGHT-HAND VARIABLE	ESTIMATED COEFFICIENT	STANDARD ERROR	T-STATISTIC
C	0.542401E-01	0.106233	0.510577
OYEXP	2.08320	1.43034	1.45643
OYUNEXP	2.60097	0.986530	2.63648

EFFICIENT MARKETS, INTEREST RATES, AND REAL ESTATE

I. Introduction and Literature Review

In the past few years (1980-1982), interest rates have fluctuated widely. When the rates have gone up, they have increased the monthly housing payments that would be required if the principal on a loan remained constant. Because consumers' incomes are relatively fixed, the price of housing became soft after the rates went up. (In real terms, the price of housing went down for periods of time when the rates were very high). A natural question that poses itself is whether one could have done better than the market average by conditioning one's buying and selling of real estate upon the movement of home mortgage rates. If one had bought when the rates were high and prices were soft, and sold after rates had gone down and prices had gone up, would one have earned more than from following a buy and hold strategy -- a strategy which said to buy real estate at the beginning of the time period in question, hold it to the end of the time period in question (irrespective of the movement of interest rates), and then sell it? This paper is designed to answer this question.

If a market is working efficiently (in a sense that will shortly be made more precise), the market must be taking into account all available information. If the market were not taking into account all available information, it would be possible to condition one's buying and selling upon this neglected information set and earn positive profits. In the case we are considering, the commodity in question is housing and the information set used to condition one's buying and selling upon is home mortgage rates.

It is useful to formalize this concept of an efficient market. Let us define the class of expected return theories as follows:

$$E(\tilde{p}_{j,t+1} \mid Z_t) = [1 + E(\tilde{r}_{j,t+1} \mid Z_t)]p_{jt} \tag{1}$$

where E is the expected value operator; p_{jt} is the price of security j at time t; $p_{j,t+1}$ is its price at t+1 (with any cash income reinvested in the security); $r_{j,t+1}$ is the one period percentage return $(p_{j,t+1} - p_{jt})/p_{jt}$; Z_t is a general symbol for whatever information is supposed to be reflected in the price at t; and the tildes indicate that $p_{j,t+1}$ and $r_{j,t+1}$ are random variables at t. Equation (1) tells us that the expected price of security j at time t+1 given the information set Z is equal to its price at time t plus its expected one period percentage return given Z times the price at t.

The efficient market hypothesis tells us that it would not be possible to earn a net return that is in excess of the market equilibrium rate by conditioning one's buying and selling upon a particular information set. Let

$$x_{j,t+1} = E(p_{j,t+1} \mid Z_t) - E(p_{j,t+1}). \tag{2}$$

$x_{j,t+1}$ is the return from conditioning one's purchase of j at time t+1 on the basis of the information in Z at time t; it is the return from speculating on the basis of the information set Z. A necessary condition for a market to be efficient is that

$$x_{j,t+1} = 0. \qquad (3)$$

It should be noted that even if we do positively ascertain that certain kinds of information have been incorporated into the real estate market, this does not enable us to positively say that the market is definitely efficient. For it may be that other bits of information are not being incorporated into the market. In effect we have another example of a point familiar from logic, viz., one counter-example is sufficient to disprove a universally quantified statement but no list of examples is ever sufficient to prove a universally quantified statement. In this case, the universally quantified statement is the assertion that the market takes into account *all* relevant economic information. Thus even if we find that the market is taking into account the economic information that we test for, it might be the case that other information is not being taken into account; hence we would have some positive evidence that the market is efficient if we obtain positive tests but not conclusive evidence. That is, the expectation of what the price of commodity j will be at time t+1 must equal the conditional expectation of the price at time t+1 (when the expectation is conditioned on the information set Z). For only in this case will equation (3) equal 0. (3) tells us that the sequence of returns $[X_{jt}]$ is a "fair game" with respect to the information sequence $[Z_t]$.[1]

It should be noted that even if we do positively ascertain that certain kinds of information have been incorporated into the real estate market, this does not enable us to positively say that the market is definitely efficient. For it may be that other bits of information are not being incorporated into the market. In effect we have another example of a point familiar from logic, viz., one counter-example is sufficient to disprove a universally quantified statement but no list of examples is ever sufficient to prove a universally quantified statement. In this case, the universally quantified statement is the assertion that the market takes into account *all* relevant economic information. Thus even if we find that the market is taking into account the economic information that we test for, it might be the case that other information is not being taken into account; hence we would have some positive evidence that the market is efficient if we obtain positive tests but not conclusive evidence. Our condition (3) should thus be thought of as a necessary but not sufficient condition.

The above formulation of the "fair game" model is due to Samuelson[2] and Mandelbrot.[3] Although Bachelier[4] had anticipated some of the results of the Samuelson-Mandelbrot (S-M) model, it was not until the S-M papers of 1965-1966 that previous empirical studies received a rigorous theoretical underpinning.

The characterization of market efficiency in terms of the conditional and unconditional expectation of price captures the notion that all available information is to be incorporated into the price of a commodity in an efficient market. If the information which was conditioned upon yielded a better prediction of what tomorrow's price would be, then the traders in that market would not have taken that information into account.

Kendall[5] was one of the first to conduct an empirical efficient market test. He examined the effect of weekly changes in nineteen indices of British industrial share prices and in spot prices for cotton and wheat. After much analysis of possible serial correlations, Kendall concluded that these market were indeed efficient in that conditioning one's purchases upon past prices would not have enabled one to more accurately predict future prices and so earn a return in excess of the market.

Sidney Alexander[6] has performed efficient market tests for the stock market. Alexander studied certain trading rules which he called y% trading rules. An example of a y% trading rule would be a 3% trading rule. Such a rule would tell us that if a stock went up at least 3%, buy and hold the security until its price moves down at least 3% from a subsequent high, at which time simultaneously sell and go short. The short position is maintained until the price rises at least 3% above a subsequent low, at which time one covers the short position and buys. Thus the information set conditioned upon in this case is the past price behavior of the stock (including its highs and lows). After extensively studying various y% filters, and then correcting for some initially incorrect assumptions, Alexander concludes in his final paper on the subject:

> In fact, at this point I should advise any reader who is interested only in practical results, and who is not a floor trader and so must pay commissions, to turn to other sources on how to beat buy and hold. The rest of this article is devoted principally to a theoretical consideration of whether the observed results are consistent with a random walk hypothesis.[7]

Alexander found that some profitability existed for the results of very small filters, i.e., filters in the range of 1% and under. The profitability of these small filters is inconsistent with the efficient market hypothesis but the profitability vanishes if one takes into account even the minimum trading costs that a floor trader must pay. Even a trader who owns his own seat on the New York Stock Exchange must pay a clearinghouse transaction on each trade that amounts to about .1% per turnaround transaction, i.e., sales plus purchase. Fama-Blume[8] later showed that because such small filters produce such frequent trades, these minimum trading costs are sufficient to eliminate their advantage over a buy and hold strategy.

The efficient market test in this paper differs from those done in the previous studies in two significant ways. The first is that this study deals with a non-uniform, differentiated commodity, i.e., housing. No two homes are exactly alike if only for the simple reason that they are not in exactly the same location. And as every real estate broker knows, there are three keys to the value of real estate – location, location, location. Houses will also differ in quality, age, quantity, etc. The previous efficient market studies have all dealt with standardized commodities.

The previous definition of efficiency in (3) deals with the case of a homogenous commodity. In dealing with a non-homogenous commodity, it is necessary to control for the characteristics of the commodity. In this case, the condition for market efficiency becomes

$$E(p_{j,t+1} \mid Z_{jt}, R_t) = E(p_{j,t+1} \mid Z_{jt}) \tag{4}$$

where Z_t represents characteristics of the house, and R_t represents any other information one might wish to condition upon. The condition I actually use to test efficiency in my model of differentiated commodities (in addition to (4)) is the implied condition that

$$E(p_{j,t+1} \mid Z_{jt}, M_{t-k}) = E(p_{j,t+1} \mid Z_{jt}) \tag{5}$$

where k is some lag, and M_{t-k} stands for lagged mortgage rates.

It should be noted that if the model is misspecified, my tests do not bear only on the model of market efficiency. The conclusion that (5) tests market efficiency is valid only under the assumption that the model of price determination is specified correctly -- or rather correctly specified to the extent that any omitted explanatory variables are uncorrelated with the variable being conditioned upon. In my case, the explanatory variables are those of the square feet of the house, square feet of the lot, and the quality of location of the house. There are undoubtedly other omitted explanatory variables which would marginally improve one's ability to control for the characteristics of the house, e.g., the number of bathrooms in the house, the quality of kitchen, the presence of central heating and air conditioning. As long as these omitted variables are uncorrelated with the variable being conditioned upon, viz., home mortgage rates, the efficiency test is robust with respect to misspecification. As there is no good reason to think that these omitted variables are correlated with current or lagged interest rates, the problem of misspecification does not seem particularly troublesome.

The second way in which this study differs from other market tests is that the agents in the real estate market are generally not the professionals that have been studied in the previous tests. Although there are undoubtedly a large number of very professional real estate brokers, a significant amount of the residential brokers work only part time as compared to the full time work of the professional security trader. It is also the case that many buyers do not use brokers but rather buy and sell their homes themselves. While it is no doubt the case that these buyers invest significant amounts of time educating themselves since this is probably the largest purchase of their lives, it is nevertheless the case that the expertise developed in two to six months of house hunting is not likely to match that generated by many years of full time work. It is probably not possible to exactly measure the amount of difference in the professionalization of the residential real estate market versus the financial securities market but there does seem nevertheless to be a difference.

These two differences combine to lead one to suspect that if the efficient market hypothesis is to fail, it may well fail in the kind of market we are studying here. In his review article cited above, Fama discussed the generally successful confirmations of the efficient market hypothesis.

We might, however, expect the above two mentioned features to result in the efficient market hypothesis being falsified in the real estate market. Since the product is a differentiated one, it might be the case that agents do not possess very good information about the products and the prices of these products. It is certainly easier to possess complete and accurate information about a product when the product is the standardized type financial securities are. The fact that the economic agents in the real estate market are not as professional as the agents in the financial securities markets also might lead to the efficient market hypothesis being falsified. For it might be that these nonprofessional agents are not taking into account all the relevant information in deciding what the future prices will most likely be.

II. Practical Considerations and Efficient Markets

The information which I propose to examine is information about home mortgage rates. It is frequently said among real estate brokers that the time to buy is when real estate mortgages are high. When the rates have gone up, as they have so dramatically at times in the last few years, it is thought that sellers are begging for buyers and there are plenty of good buys out there. As evidence of this widespread viewpoint, let me quote the following passage from a respected journal in the real estate profession written in 1982.

> Depressed market conditions allow buyers to purchase property at 1980 prices. When interest rates begin to recede, the pent up

appreciation will increase dramatically.

This concept of a buyer's market due to high interest rates is at odds with the economists' notion of an efficient market equilibrium in the real estate market. The economist who feels that there is an efficient market equilibrium in effect feels that the ordinary real estate broker can do both better and worse than he thinks he can. The broker can and will do worse than he thinks he can by buying in a "buyers" market because this information has already been taken into account by other brokers and so the profit opportunities the broker sees are really illusory. On the other hand, because brokers are efficiently taking advantage of all profit opportunities offered by speculation off movements in home mortgage rates, the economist at the same time thinks that brokers on average are doing a better job than the average broker thinks. The economist who believes there is an efficient market equilibrium thinks that brokers are doing such a good job on average that it will not be possible for the individual broker to profitably speculate from costlessly available information such as the influence of home mortgage rates.

There is another possibility which we haven't considered yet. This is the possibility that our statistical study will show that it would have been possible to have earned positive profits by speculating on the movements of home mortgage rates but that this was so only because the market was not in equilibrium. There is nothing that prevents the market from adjusting to a new equilibrium in which there are no excess profits to be earned by speculating on the movement of interest rates, e.g., certain classes are not barred from owning land so there is free entry into the field.10

III. The Data

The data used in this study come from two different sources. One source is the Southern California Real Estate Research Council. The quality of the home in this sample has been held constant throughout the last fifteen years for which there are recorded data. The prices have been compiled semi-annually for the years 1968 to 1982; these are the years being analyzed. I do not know how the location of the houses was kept constant because it would seem difficult to cover a broad area such as Los Angeles County and still keep the quality of location constant. I also do not know how much latitude was allowed for differences in the square footage and lot footage of the houses. (The Council does not make available to the public the micro data their averages are based upon.) It is doubtful they had exactly the same amount of these and I don't know the latitude allowed.

The second data base was obtained from The Pasadena Board of Realtors Multiple Listing Books. The data were compiled quarterly -- from the second quarter of 1968 to the third quarter of 1982 -- in volumes which list what houses sold.

There are six pieces of information collected on each transaction. These are the price for which the house sold, the number of square feet in the house, the number of square feet on the lot on which the house rested, a rental price at which the house would have rented, the year in which the house sold, the interest rate on the mortgage of the house when it was bought, and a dummy variable for the quality of the area in which the house was located.

The rental price was constructed for each house by regressing the rental price of houses that rented in each quarter in a given area upon a constant and the square feet of the houses. Rental prices for the houses that sold were then estimated by multiplying the number of square feet in their house times the rental coefficient of house square feet in that quarter and adding the constant of that quarter.

There were two dummy variables used as indices of quality of location. Area 4 consisted of area 4 in the Pasadena Multiple Listing Book. This area is a predominantly lower income area with a significantly higher crime rate than the other area which was used, i.e., area 7. The transactions which were used were those which came at the lower end of the price range in the MLS book; the MLS book is arranged with the houses which sold for the least amount of money in the beginning of the section and the transactions I compiled were always collected starting at the beginning of the section. Thus while areas 4 and 7 are not perfectly homogenous within themselves in so far as the quality of area goes, the fact that the transactions were taken from the lower priced range of each area in every quarter serves to make the intra-area quality more uniform than it otherwise would be.

One possible bias in the model exists because the houses which were selected in the Pasadena areas were always selected from the lowest prices which sold. The reason for doing this is to make the neighborhood effect in each of the areas more uniform. The higher priced houses in the southeast Pasadena area are in a significantly nicer area than the lower priced houses. The higher priced houses consist of estate style homes in areas with wider and better maintained streets which also have better street lights. Thus in order to keep the neighborhood effect as uniform as possible, houses were selected from the lower priced houses which sold in both Pasadena areas.

An adjusted price was constructed for each of the 1617 transactions which were recorded. This adjusted price was the cost in real terms (using the second quarter of 1968 as a base) of buying the house and paying the mortgage on it for a year plus the rental value of the house for a year. It was computed according to the following formula:

$$P_{it} = \frac{P_{it} - 12[.94rental_{it} - (.85P_{it} \ x \ ppt_{it})] + .02P_{it}}{d_t} \qquad (6)$$

Particularly to those unfamiliar with the business of buying, renting, and then selling houses, the above expression (6) may appear rather mysterious and ad hoc. In fact, however, it is the natural way to adjust the nominal selling price of a house to reflect the monetary effect of carrying the house for a year. If we take the actual price someone paid for a house to be his real economic cost of holding the house for a year, we would be making a mistake. For this cost of buying the home can be greater or smaller than the relevant economic cost depending upon whether or not the house yielded a positive or negative cash flow over the time period in question. The situation is analogous to determining the relevant price of a security on the stock market. If one stock yielded higher dividends than another stock even though both sold for the same prices over a one year period, they would not have the same relevant economic prices. The one with the higher dividend stream would have had a lower economic price at the beginning of the period because its net cost was lower than the other one (or alternatively one could say it had a higher economic price at the end of the period). Similarly, if two houses sold for the same price over a one year period but one house had a higher cash flow than the other one, we would want to say that the house with the higher cash flow had a lower economic cost than the other house. This is the motivation for adjusting the data according to the formula in (6).

Let us now go through expression (6) term by term. P_{it} is the price the ith house sold for in period t; $rental_{it}$ is the estimsted rental value of that house, ppt_{it} is the price per thousand dollars that an owner would have to pay per month on a loan amortized over thirty years, and d_t is the deflator used for the tth quarter to deflate the price into real terms with 1968 as the base year. It is assumed that an individual makes a down payment of 15% which is why ppt is multiplied by only .85; since the individual made a down payment of .15, he is only paying interest on .85 of the price. A management fee of 6% monthly rent is used which is why only 94% of the rent is allocated under positive cash flow. A property tax rate of 2% of the sales price per year is used for the years 1968 to 1978; that is why .02 is added to the sales price in figuring out the cash flow of the house for that year. (After the passage of propsition 13 in 1978 this figure is reduced to 1%.) Thus expression (6) is used to create for each house an adjusted price which is a price in constant dollars after taking into account the positive or negative cash flow of the house for a one year period.

IV. The Estimation Procedure and the Hypothesis to Test

In a data base involving both cross sections and time series (such as the one in this paper), it is necessary to test if the data can be pooled across cross sections and between time periods. The areas are quite distinct socioeconomically and casual observation of the coefficients from the different areas revealed they were significantly different. The following regressions were run which allowed the intercept and slope coefficients of each area to be different. These regressions enable us to determine whether it is permissible to pool the data across time.

$$P_{it} = a_{1t}D_{1it} + a_{2t}D_{2it} + B_{1t}D_{1it}HF_{it} + B_{2t}D_{2it}HF_{it} + B_{3t}D_{1it}LF_{it} + B_{4t}D_{2it}LF_{it} + u_{it} \qquad (7)$$

$i = 1, \ldots 30;\ t = 1, \ldots 54$

In the above expression, the net adjusted price which was calculated for each house from equation (6) is the dependent variable. The independent variables are respectively: the intercept term if the house is in area 1 at time t, the intercept term if the house is in area 2 at time t, the slope coefficient of house square feet if the house is in area 1 at time t, the slope coefficient of house square feet if the house is in area 2 at time t, the slope coefficient of lot square feet if the house is in area 1 at time t, the slope coefficient of lot square feet if the house is in area 2 at time t, and the error term. There are in general 30 observations in each quarter and so i runs from 1 to 30; there are 54 quarters and so t runs from 1 to 54.

The following regression was then run which pooled the data over time.

$$P_i = a_1 D_{1i} + a_2 D_{2i} + B_1 D_{1i} HF_i + B_2 D_{2i} HF_i + B_3 D_{1i} LF_i + B_4 D_{2i} LF_i + u_i \quad (8)$$

$$i = 1, \ldots 1617$$

The following F test was used to determine if the data could be pooled over time.

$$F = \frac{S_2 - S_1/r}{S_1/n - k} \quad (9)$$

In Equation (9), S_2 stands for the restricted residual sum of squares obtained from (8), S_1 stands for the unrestricted residual sum of squares obtained from (7), r stands for the number of restrictions and equals (53 x 6) = 318, n is the number of observations and equals 1617, k is the number of explanatory variables in the unrestricted model and equals (54 x 6) = 324. The value of this F ratio is 30.04. The critical value at the 95% level is 1.00 and so the data cannot be straightforwardly pooled.

It is not surprising that the data cannot be straightforwardly pooled because there has been an upward trend in real terms of California housing prices. To allow for this trend, I have run a regression (Equation 10) which allows time to interact with each of the independent explanatory variables given in (7).

$$P_{it} = a_1 D_{1it} + a_1' t D_{it} + a_2 D_{2it} + a_2' t D_{2it} + B_1 D_{1it} HF_{it} + \quad (10)$$

$$B_1' t D_{1it} + B_2 D_{2it} HF_{it} + B_2' t D_{2it} HF_{it} + B_3 D_{1it} LF_{it} +$$

$$B_3' t D_{1it} LF_{it} + B_4 D_{2it} LF_{it} + B_4' t D_{2it} LF_{it} + u_{it}$$

When (8) is so re-run, the r-squared jumps from .47 to .82. See Tables 1 and 2 for these results.

In this context, a test of efficiency will be to see whether one could have done better by incorporating the information yielded by lagged values of real interest rates given the general upward trend of the last fifteen years. I have defined the real rate of interest to be the mortgage rate minus the rate of inflation over the last year. This real rate of interest will vary as the difference between mortgage rates and the rate of inflation varies. To test efficiency, I re-estimated (10) after including the lagged real rate of interest as one of the explanatory variables as in

$$P_{it} = a_1 D_{1it} + a_1' t D_{it} + a_2 D_{2it} + a_2' t D_{2it} + B_1 D_{1it} HF_{it} + \quad (11)$$

$$B_1' t D_{1it} HF_{it} + B_2 D_{2it} HF_{it} + B_2' t D_{2it} HF_{it} + B_3 D_{1it} LF_{it} +$$

$$B_3' t D_{1it} LF_{it} + B_4 D_{2it} LF_{it} + B_4' t D_{2it} LF_{it} + B_5 LAM_{t-1} + u_{it}$$

If the coefficient on the lagged mortgage rate was not zero and was significant, the market would not be efficient. If the coefficient on the interest rate was not zero, then the unconditional expectation of housing prices would not equal the expectation given lagged mortgage rates, i.e.,

$$E(p_{it} \mid Z_t) \neq E(p_{it} \mid LAM_{t-1}, Z_t) \qquad (12)$$

where M_{t-1} stands for a value of lagged mortgage rates and Z_t for the right hand side explanatory variables of (10). Let

$$X_{h,t+1} = E(p_{h,t+1} \mid Z_t, LAM_t) - E(p_{h,t+1} \mid Z_t) \neq 0 \qquad (13)$$

which is a violation of the efficiency condition given in (5) that

$$X_{h,t+1} = 0. \qquad (14)$$

In order to re-run (11) using an ordinary least squares approach with one of the explanatory variables now being lagged mortgage rates, and to test the nullity of the coefficient of the lagged mortgage rates, it is necessary to insure that direction of causality runs from the explanatory variables to the dependent variable. One of the assumptions of ordinairy least squares estimation is that the direction of causality is from the explanatory variables to the dependent variable. In the case we are considering, we must test to see if the dependent variable of the price of a house influences the explanatory variable of home mortgage rates. This may not seem a priori very likely but we still must test for this independence in order to insure the validity of the use of ordinary least squares. It also might be the case that a high price of housing causes a high mortgage rate because there is a greater amount of needed financing when the price of housing rises. Thus although the likelihood that housing prices determine mortgage rates is not particularly strong a priori, it still is possible that they do so influence them and we must test for this in order to insure the validity of ordinary least squares.

It is thus necessary to find an instrumental variable for mortgage rates in the Los Angeles area which is definitely not correlated with the housing prices in our sample but is highly correlated with the movement of Los Angeles area mortgage rates. In this regard I have selected the mortgage rates in the entire United States as an instrumental variable. It is necessary to regress the Los Angeles rate upon the U.S. rate and then to use the fitted value of the Los Angeles area rate from this regression. Thus (15) is the regression

$$LAM_{i,t-k} = c_1 D_{1it} + c_1' t D_{it} + c_2 D_{2it} + c_2' t D_{2it} + F_2 D_{2it} HF_{it} + F_2' t D_{2it} HF_{it} + F_3 D_{1it} LF_{it} + F_3' t D_{1it} LF_{it} + F_4 D_{2it} LF_{it} + F_4' t D_{2it} LF_{it} + F_5 USM_{i,t-k} + w_{it} \qquad (15)$$

where $LAM_{it} = LAM_t$ and $USM_{it} = USM_t$ for all i.

(11) is now run using the fitted values of the L.A. rate obtained from (15) as in Equation (16)

$$P_{it} = a_1 D_{1it} + a_1' t D_{it} + a_2 D_{2it} + a_2' t D_{2it} + B_1 D_{1it} HF_{it} + B_1' t D_{1it} HF_{it} + B_2 D_{2it} HF_{it} + B_2' t D_{2it} HF_{it} + B_3 D_{1it} LF_{it} + B_3' t D_{1it} LF_{it} + B_4 D_{2it} LF_{it} + B_4' t D_{2it} LF_{it} + B_6 LAMF_{t-1} + u_{it} \qquad (16)$$

From (11) and (16) we obtain the estimated coefficients and standard errors of the L.A. rate and the fitted value of the L.A. rate. From these we compute the following statistic which has a chi-square distribution with 1 degree of freedom:

$$\frac{(B_6 - B_5)^2}{(VarB_6 - VarB_5)} \tag{17}$$

and where $VarB_6$ and $VarB_5$ stand for (respectively) the variances of the L.A. rate and the fitted value of the L.A. rate. If we accept the hypothesis that the dependent variable of housing prices does not influence the explanatory variables, then we can use the L.A. rate in an efficiency test.[12] Equation (17) is the statistic which allows us to determine if the direction of causality runs in the appropriate direction.

The estimation procedure used for the data base provided by the Southern California Real Estate Research Council is similar. To allow for a time trend, I first estimated the equation

$$P_t = a + B_1 P_{t-1} + u_{t-1}. \tag{18}$$

An efficiency test in this context is to then estimate

$$P_t = a + B_1 P_{t-1} + B_2 LAM_{t-1} + u_{t-1} \tag{19}$$

and to see if the coefficient B_2 is non-zero and significant. This is in accord with the criteria we are using for efficiency with a homogenous product (because since we are not given any of the particular characteristics of houses it is assumed they are the same) which is that

$$E(P_{t+1} \mid p_t) = E(p_{t+1} \mid p_t, R_t). \qquad (20)$$

Equation (19) is then re-run using the fitted value of B_2 which is obtained from the equation

$$LAM_{t-k} = c + dp_{t-k} + e\,USM_{t-k} + w_t. \qquad (21)$$

The final step is to compare the estimated coefficients and variances as in (17).

V. Results and Conclusion

The key results of my study are contained in Tables 3, 4, and5. Table 3 shows the coefficient of the real rate lagged six months to be negative and not significant at the 75% level. Tables 4 and 5 are of interest because they show the coefficient of the real rate lagged 12 and 18 months (respectively) to be both negative and significant at the 99% level. The coefficient on the real mortgage rate lagged 12 months is minus 421.7 and the t-statistic is -3.917. Thus for every one point decrease in the lagged 12 month mortgage rates, the price of a house jumped $421.70. When one considers that most of the houses in 1968 sold between $10,000.00 and $27,000.00 (depending on the area, size of the house and lot), an increase of $421.70 for every one point decrease in the real lagged 12 month mortgage rate is rather significant. (The prices of the other years were deflated into 1968 dollars and so this coefficient is also noteworthy for them too.) Table 5 shows the coefficient on the real rate lagged 18 months to be -360.1. This is also noteworthy.

The results of Tables 4 and 5 comprise a refutation of the efficient market hypothesis in the real estate market in that one could have done better than the market average by conditioning upon lagged real interest rates of 12 and 18 months; the market did not fully take this information into account. When these rates are low, it has been established that prices will be higher in the next twelve to eighteen months than they otherwise would be. If a speculator were trying to figure out when to buy houses, he would have been better able to better predict the prices of houses by looking at lagged interest rates than merely by looking at the past price of housing. The predictions of future prices of this speculator who was conditioning his predictions upon lagged real mortgage rates (as well as the characteristics of the houses and the past price of housing) would have been systematically better than the market's predictions which did not incorporate this information about lagged real mortgage rates into current housing prices. Since this information wasn't incorporated into the price of housing, the market wasn't fully efficient.[13]

In the case of this paper, I can think of two real world factors which might complicate and possibly change the inefficiency result of this paper. (as they did in Alexander's later work). The fact that my model does not deal with these factors could lead to possible misspecification in the model. One factor is that the transaction costs of selling of approximately 6% were not explicitly considered in this paper. My model does not explicitly say when (or how often) one should have bought and sold real estate. Thus it is not possible to say how much more often (if at all) the brokerage commission of 6% would have been assessed if one had conditioned one's purchases upon lagged real interest rates.

The second factor is that the effects of the tax system have not been considered in this paper. It was not possible to obtain data on the income tax brackets of the owners of the houses and so the after tax prices of buying a house and holding it for a year were not computed; only the before tax price was computed. It is possible that if this different set of prices was used, market efficiency could possibly result.

In the second data set, the coefficient on lagged real interest rates was never significant at the 90% level. In the Riverside-San Bernadino area, the coefficient on lagged real rates of 18 months was -94.17 and the t-statistic was -1.27 when the required critical value at the 90% level is 1.31. That was the closest the t-statistic came to being significant.

The results of the first micro data set should be viewed with more confidence than the results of the aggregated data set for two reasons. First, the micro-data set explicitly controlled for variations in the quality of the houses in the sample. We can hope that by controlling the sample of houses in the aggregate data set, the variation in house quality was controlled for but since there must have been some variation in the sample (in terms of location, house square feet, and lot square feet), this control will not have been perfect. Second, the prices in the aggregated data set do not take into account the positive or negative cash flow of the house since there were no estimated rental prices in the aggregated data of the Southern California Real Estate Research Council. These two reasons lead one to have more confidence in the micro data set.

It is of interest that the test for the exogeneity of the L.A. rate showed that the L.A. rate was rightly considered exogenous in all of the 16 regressions that were run on this data. The values of the test specified in Equation (20) are given in Table 6. The critical value for rejecting the exogeneity of the L.A. rate at the 95% level is 3.84. As can be seen by examining the table, it was permissable to accept the L.A. rate as exogenous in every one of the tests; the largest value achieved was .05. This provides some support for taking the L.A. rate as exogenous in our first data base. It also seems empirically plausible that the prices of the group of 1617 houses in our survey did not affect the L.A. interest rate. The exogeneity of the L.A. rate thus seems to be a rather uncontroversial assumption.[14]

Another possibility which deserves mention is that the market we have studied here may have been in temporary disequiibrium during the period we studied. The period we studied was one in which inflation was at relatively low levels at the beginning of the period and at relatively high levels at the end of the period. This resulted in moderate real rates of interest at the beginning of the period in question and then resulted in rather low or negative rates of real interest in the period of the mid to late 1970's. Real rates of interest then rose to rather high levels after inflation subsided in 1981 and 1982. It may be that the market was in temporary disequilibrium during this period and that the perticipants in the market had not had time to learn about the effect of real rates of inflation (rather than the case that long run, repeated profit opportunities were being missed.)

FOOTNOTES

This paper has benefited significantly from the comments of Burt Klein, Bruce Cain, Jeff Dubin, and especially Quang Vuong.

(1) This is the notion of an efficient market characterized as a fair game given in Fama's "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 1970, pp. 385-386.

(2) Paul Samuelson, "Proof that Properly Anticipated Prices Fluctuate Randomly," *Industrial Management Review*, Spring, 1965.

(3) Benoit Mandelbrot, "Forecasts of Future Prices, Unbiased Markets, and Martingale Models," *Journal of Business*, January 1966.

(4) Louis Bachelier, "Theorie de la Speculation," Paris, 1900.

(5) Maurice G. Kendall, "The Analysis of Economic Time-Series, Part 1: Prices," *Journal of the Royal Statistical Society*, 1953.

(6) Sidney Alexander, "Price Movements in Speculative Markets: Trends or Random Walks," *Industrial Management Review*, May 1961.

(7) Ibid., pp. 338-372.

(8) Eugene Fama and Marshall Blume, "Filter Rules and Stock Market Trading Profits," *Journal of Business*, January 1966.

(9) Keith Kube, "Buy Now Because..." *Real Estate Today*, April 1982.

(10) See also, however, the possibility of misspecification discussed in the conclusion.

(11) In future test of the real estate market, it might make more sense to investigate if the dependent variable of house square feet can rightly be taken as exogenous. Many say that the current high price of housing has caused a number of square feet in a house to decrease; thus the exogeneity of this variable might also be questioned. Fortunately no one has speculated that this trend has taken place over the last 15 years and so its future possibility would not seem to affect this study.

(12) See "Specification Tests in Econometrics," by J. A. Hausman, *Econometrica*, November 1978, pp. 1251-1273.

(13) There are other interesting tests that one could have done for efficiency. In particular, Jeff Dubin has pointed out that one could test for the joint effect of lagged real rates of six, twelve, and eighteen months. For our purposes, however, testing these in isolation has been sufficient because even when tested in isolation the lagged rates of twelve and eighteen months reveal the market to be inefficient.

(14) Given the low a priori probability that the L.A. rate was not exogenous, and the results of the exogeneity test mentioned in the paper, an exogeneity test was not run on the sample of the

Pasadena Board of Realtors.

REFERENCES

Alexander, Sidney. "Price Movements in Speculative Markets: Trends or Random Walks." *Industrial Managment Review*, May 1961.

Bachelier, Louis. "Theorie de la Speculation." Paris, 1900.

Fama, Eugene. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance*, 1970, pp. 385-386.

Fama, Eugene and Blume Marshall. "Filter Rules and Stock Market Trading Profits." *Journal of Business*, January 1966.

Hausman, J. A. "Specification Tests in Econometrics." *Econometrica*, November 1978, pp. 1251-1273.

Kendall, Maurice G. "The Analysis of Economic Time-Series, Part 1: Prices." *Journal of the Royal Statistical Society*, 1953.

Kube, Keith. "Buy Now Because..." *Real Estate Today*, April 1982.

Lowry, Albert J. *How You Can Become Financially Independent by Investing in Real Estate* (New York: Simon and Schuster, 1972).

Mandelbrot, Benoit. "Forecasts of Future Prices, Unbiased Markets, and Martingale Models." *Journal of Business*, January 1966.

Samuelson, Paul. "Proof that Properly Anticipated Prices Fluctuate Randomly." *Industrial Management Review*, Spring 1965.

TABLE 1

Dependent variable is adjp

Right-hand variables: houseft1 houseft2 lotft1 lotft2
area1 area2

Variable Name	Estimated Coefficient	Standard Error	T-Statistic
houseft1	1.525	1.680	0.9076
houseft2	22.62	1.025	22.06
lotft1	0.03321	0.1150	0.2889
lotft2	2.981	0.2084	14.30
area1	12250.0	1890.0	6.483
area2	-22910.0	1606.0	-14.26

r-squared = 0.4714

number of observations = 1617

sum of squared residuals = 0.3869×10^{12}

standard error of the regression = 15500

TABLE 2

Dependent variable is adjp

Right-hand variables: houseft1 houseft1t houseft2 houseft2t

lotft1 lotft1t lotft2 lotft2t

area1 area1t area2 area2t

Variable Name	Estimated Coefficient	Standard Error	T-Statistic
houseft1	7.257	2.084	3.482
houseft1t	-0.1298	0.06441	-2.016
houseft2	-14.05	1.252	-11.22
houseft2t	1.117	0.03569	31.29
lotft1	0.02470	0.1688	0.1463
lotft1t	0.000945	0.005841	0.1663
lotft2	0.5442	0.2472	2.202
lotft2t	0.0207	0.006785	3.055
area1	-1393.0	2405.0	-0.5793
area1t	382.0	69.34	5.509
area2	21220.0	2156.0	9.842
area2t	-911.6	52.41	-17.38

r-squared = 0.8199

number of observations = 1617

sum of squared residuals = 0.1318×10^{12}

standard error of the regressions = 9063.0

TABLE 3

Dependent variable is adjp

Right-hand variables: houseft1 houseft1t houseft2 houseft2t

lotft1 lotft1t lotft2 lotft2t

area1 area1t area2 area2t

lagrate

Variable Name	Estimated Coefficient	Standard Error	T-Statistic
houseft1	7.260	2.085	3.483
houseft1t	-0.1302	0.06444	-2.021
houseft2	-14.07	1.254	-11.22
houseft2t	1.118	0.03575	31.26
lotft1	0.02450	0.1688	0.1451
lotft1t	0.0009493	0.005843	0.1625
lotft2	0.5446	0.2472	2.203
lotft2t	0.02076	0.006788	3.059
area1	-1282.0	2424.0	-0.5288
area1t	380.6	69.46	5.479
area2	21360.0	2168.0	9.761
area2t	-914.1	53.05	-17.23
lagrate	-34.32	91.99	-0.3731

r-squared = 0.8199

number of observations = 1617

sum of squared residuals = 0.1318×10^{12}

standard error of the regression = 9066.0

TABLE 4

Dependent variable is adjp

Right-hand variables: houseft1 houseft1t houseft2 houseft2t

lotft1 lotft1t lotft2 lotft2t

area1 area1t area2 area2t

lagrate4

Variable Name	Estimated Coefficient	Standard Error	T-Statistic
houseft1	7.246	2.075	3.493
houseft1t	-0.1351	0.06414	-2.122
houseft2	-14.27	1.248	-11.43
houseft2t	1.120	0.03554	31.52
lotft1	0.04189	0.1681	0.2493
lotft1t	0.00003098	0.005820	0.005323
lotft2	0.5319	0.2461	2.162
lotft2t	0.02102	0.006755	3.112
area1	269.3	2432.0	0.1108
area1t	359.3	69.27	5.188
area2	23340.0	2214.0	10.54
area2t	-951.9	53.21	-17.89
lagrate4	-421.7	107.7	-3.917

r-squared = 0.8216

number of observations = 1617

sum of squared residuals = 0.1306×10^{12}

standard error of the regressions = 9023.0

TABLE 5

Dependent variable is adjp

Right-hand variables: houseft1 houseft1t houseft2 houseft2t

lotft1 lotft1t lotft2 lotft2t

area1 area1t area2 area2t

lagrate6

Variable Name	Estimated Coefficient	Standard Error	T-Statistic
houseft1	7.171	2.079	3.450
houseft1t	-0.1319	0.06424	-2.053
houseft2	-14.08	1.249	-11.27
houseft2t	1.113	0.03562	31.24
lotft1	0.04645	0.1684	0.2758
lotft1t	-0.0001657	0.005836	-0.02839
lotft2	0.5368	0.2465	2.173
lotft2t	0.02064	0.006767	3.050
area1	142.8	2448.0	0.05832
area1t	356.7	69.62	5.123
area2	22940.0	2219.0	10.34
area2t	-939.9	53.07	-17.71
lagrate6	-360.1	115.0	-3.132

r-squared = 0.8210

number of observations = 1617

sum of squared residuals = 0.1310×10^{12}

standard error of the regressions = 9038.0

TABLE 6

Seven County Area

RLA(-2)	.002
RLA(-4)	.015
RLA(-6)	.006

Los Angeles County

RLA(-2)	.001
RLA(-4)	.003
RLA(-6)	.033

Orange County

RLA(-2)	.002
RLA(-4)	.050
RLA(-6)	.015

Riverside and San Bernadino

RLA(-2)	.004
RLA(-4)	.026
RLA(-6)	.042