Essays in Economic Geography and Finance

by

Ludwig Boris Chincarini

B.A., Economics, University of California at Berkeley, 1991

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1995

Author ....
Department of Economics
May 24, 1995

Certified by 5/19/95
Paul Krugman
Professor of Economics
Thesis Supervisor

Certified by ....
Rudiger Dornbusch
Ford Professor of International Economics
Thesis Supervisor

Accepted by ....
Richard Eckaus
Professor of Economics
Chairman, Departmental Committee on Graduate Students

Essays in Economic Geography and Finance

by

Ludwig Boris Chincarini

Submitted to the Department of Economics
on May 24, 1995, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis consists of three chapters which are quite distinct from each other. The first chapter concerns itself with the field of economic geography. The second chapter is concerned with issues in the literature of labor migration. Finally, the third chapter was co-written with Guillermo Llorente and is on the subject of finance, specifically on the informational content of volume and returns in the stock market.

Chapter one attempts to understand how cities evolve over time. A theoretical model, adapted from the ideas of Krugman (1991, 1994), is constructed to understand the formation of economic clusters over time. Specifically, the model concerns itself with reconstructing the movement of workers and industries to locations based on centripetal and centrifugal economic forces. In a dynamic context, this allows one to have a better understanding of the evolution of cities. Previous studies have relied on numerical methods in discrete space to analyze the formation of "cities". This paper reformulates the Krugman spatial economy in a more solidified manner which gives interesting analytical results without the need for simulations and can partially explain the emergence of Edge Cities in the last thirty years.

There has been a vast amount of research and speculation as to why people migrate from one region to another. Chapter two of this thesis uses individual data from the Public Use Micro Sample of 1990 in order to create a sub-sample of people that moved from one region to another between 1985-1990. The characteristics of the region chosen are compared to the other regions they could have chosen so as to get a "revealed preference" argument as to why people migrate. The conditional logit technique that is used in the estimation along with the careful geographical division of locations brings forth an array of interesting results, some of which corroborate previous studies and some of which are entirely new. The results have policy implications which local governments will be concerned with, especially when considering issues of future growth and prosperity of their region. For economic geographers, this study presents many results upon which theorists can base their models of economic geography.

Chapter three tests for the informational content of trading volume per se as postulated by Blume, Easley, and O'Hara (1994) and about its "identification" function in price reversals as postulated in Campbell, Grossman, and Wang (1993) using different trading strategy schemes. We find evidence of the information contained in volume and its difference from that of return information. This result is robust to the stock size and to different time horizons. We also find evidence of price reversal for highly traded stocks, consistent for different time horizons, and among stocks of different size. For low traded stocks we do not find price trending.

Thesis Supervisor: Paul Krugman
Title: Professor of Economics

Thesis Supervisor: Rudiger Dornbusch
Title: Ford Professor of International Economics

Questa tesi è dedicata alla memoria di mio nonno Ludovico.
Aprile 14, 1909-Gennaio 4, 1995.

Opportunities do not come with their values stamped upon them. Everyone must be challenged. A day dawns, quite like other days; in it a single hour comes, quite like other hours; but in that day and in that hour the chance of a lifetime faces us. To face every opportunity of life thoughtfully and ask its meaning bravely and earnestly, is the only way to meet the supreme opportunities when they come, whether open-faced or disguised.—Maltbie Babcock.

BIOGRAPHICAL NOTE

Ludwig Boris Chincarini was born in Daun, West Germany on April 29, 1969. He finished his "O" and "A" level examinations in maths, physics, and history at the Sir James Henderson British School of Milan in 1988. He graduated from the University of California at Berkeley in May, 1991 with a Bachelors of Arts degree in economics. While there he received numerous honors, including Phi Beta Kappa, Golden Key National Honor Society, Omicron Delta Epsilon Honor Society in Economics, the Earl Rolph prize for best senior thesis, Departmental Citation, and highest distinction in general scholarship.

Since 1991, the author has attended the Massachusetts Institute of Technology in the pursuit of his doctoral degree in economics. His fields of specialization are macroeconomics and international economics with minors in public finance and econometrics. While there he has acted as a teaching assistant for Paul Krugman and Rudiger Dornbusch, has lectured many students in a very popular macroeconomics course at M.I.T., and has worked as a research assistant for Professor Janet Currie, Professors Julio Rotemberg and Michael Woodford, and Professor Rudiger Dornbusch.

Contents

List of Figures

List of Tables

Acknowledgements

Family (in alphabetical order):

Angelo Chincarini ("Watermelon"): *He taught me how success and stature can go together with good humor and fun.*

Mom and Dad: *Without them, I wouldn't be who I am.*

Nonna Maria Luisa Chincarini: *Her example has always inspired me and given me hope for the future.*

Wolfgang Chincarini: *He has always been my best friend, my teacher, my support, my pride, and my inspiration. When I think of the most difficult moments in my life, they were the times that he was not present.*

Friends (in alphabetical order):

Neer Asherie: *We first met at Sir James Henderson British School of Milan. We were friends in a time when, for different reasons, others opposed us. When I was sixteen and he was fourteen, Neer stimulated me intellectually unlike anyone could do or has since done. We learned maths, physics, and life together in those years of high school. He departed for Cambridge to study physics and I left for Berkeley to study economics. For three years, we did not communicate at all. In August 1991, we met again at M.I.T. Our friendship began again where it had left off. We have helped each other through the good, the bad, and the ugly. In these four years, I have again been stimulated by him, both morally and intellectually. These last few years haven't been easy, but we managed to create ideas and laughter together; we even managed to act in the annual economics skit party together ("I am?"). He is one of the few people that directly aided the completion of my thesis. He gave me many suggestions and tutoring in the techniques that are used in the first chapter of this thesis; without him,* *chapter one would not have existed. He has also proofread everything I ever wrote, from grant applications to phone memos. He is one of the few people who is hard on me in my own best interest and has always had faith in my ability. For this I am very grateful. It is a pleasure to know Neer, an honor to be his friend, and I am lucky enough to be his "very best good friend."*

Oded Asherie: *We have had some fun nights out in Boston; "hmmmph!" He also proofread many of my written pages.*

J. Guillermo Alvarez-Llorente: *He has been a friend throughout graduate school. He offered to co-write a paper with me at a time when I really needed to put something together and finish. I am grateful to him for that.*

Jozsef Losonczy: *We spent many hours finding ways to make the graduate school experience bearable. He made me laugh when I was going through some difficult times.*

Ioanna Manousoyanaki: *Ioanna died on May 9, 1988. I really never had a chance to thank her for the positive influence she had on me. Her kindness and relaxed attitude kept the household sane. Academically, she introduced me to mathematical induction in the Christmas break of 1986. From then on, mathematics was fun.*

Paul Schulstad: *We met in the graduate dorm of my first year at M.I.T. He was important in providing support for me when I was going through my course work at M.I.T. We spent many hours outside of the classroom making the stressful days less so. His generosity and good spirit made life pleasant. On many occasions he patiently listened to my exhausting stories and on the late nights offered a couch to sleep on. We have taught each other a lot and it has been a pleasure being his friend.*

Craig Wheeler: *Often we would take trips to Portland, Maine to visit Grittys. It was a joyful diversion from M.I.T. Craig made the days interesting with his sensational cooking, his deep understanding of political issues, and his famous quotes that brought everyone hours of laughter: "You can't eat a two pound steak?"*

Rebecca Wimberly: *There is too much to say. Rebecca, more than anyone else, has occupied my thoughts in the last two and a half years. She used to say that love was all that mattered and that love was eternal. In her absence, my thoughts of her remain and my love for her has only grown stronger; and I have finally, perhaps too* *late, understood what she meant.*

Daniel Zanger ("Lysander"): *We spent many days of graduate school philosophizing about social issues and reciting revised phrases from Shakespeare.*

Teachers and Colleagues:

Mr. Addison: *Mr. Addison was my 10th grade biology teacher. I thank him for allowing everyone except me to choose their lab partners; the punishment served me well.*

Mr. Mason: *Mr. Mason was my "O" and "A" level physics teacher and the director of the school plays. He taught me many lessons of life as a theatrical director, a physics teacher, and a friend. He was one of the guys on my side when it really mattered. I will always remember him for giving me the chance to act and to think. His best lesson was that of "extremes" and his best phrase "Chincarini!".*

Mr. Murphry: *Mr. Murphry was my 8th and 10th grade social studies teacher. For kids in adolescence, school is the last thing on their minds. He made civics and world history exciting and entertaining. He also frequently gave me detention checks for talking in class. I am not sure exactly how, but I think by the 3000th detention check, I was a better person.*

Mrs. Norman: *Mrs. Norman was my mathematics teacher for my "O" and "A" levels. When I entered her class, I would turn in assignments that were sloppy and illegible and the solutions were frequently wrong. She taught me the elegant language of mathematics. She commanded respect and is the best teacher I have ever had in my life.*

Professor Rudi Dornbusch: *I came to M.I.T. to learn how to understand the world economy; to become an international economist. Rudiger Dornbusch formally taught me two classes of macro and international economics. He informally taught me every time I met him. In the international breakfast workshop, he helped form our character and our presentation style. He has been a father figure in many ways. His knowledge of the international economy is unrivalled. As his teaching assistant for an International Finance M.B.A. course, I realized how versatile he was; he could explain difficult concepts in such a simple, eloquent manner. I was fortunate enough to hear*

Rudi speak time and time again. I feel that I am closer to being an international economist, and a large part of the credit goes to him.

Professor Paul Krugman: *Paul lectured my first course in international economics. I got to know him better as his teaching assistant for international economics. It was he who introduced me to economic geography. Unfortunately, he left for Stanford in the middle of my third year, which left me to do a lot of my research without his guidance. I thank him for the initial encouragement.*

Professor Steve Pischke: *My second paper would never have been written if it had not been for the early encouragement of Steven Pischke. Thus, I thank him for the push when I was uncertain.*

Professor Kenneth Train: *There are very few people who are experts in the area of choice estimation; Ken Train is one of those. He gave me a great deal of advice on using the techniques of the second chapter of this thesis. He also provided encouragement when I was unsure as whether to continue or not. I am grateful for his help.*

In addition, **Professor Avinash Dixit** *while visiting M.I.T. in the Fall of 1994 took the time to read a draft of my paper and make helpful suggestions.* **Professor Steve Goldman, Professor Carmen Matutes, Professor Theodore Keeler,** and **Professor Roger Craine** *were all influential in my decision to attend graduate school; they exemplified what academics should be about.* **Professor Robert C. Merton** *taught me continuous time finance and advance capital market theory. He revitalized my enthusiasm for economics at a time when others had discouraged me.* **Professor Olivier Blanchard** *taught me my core macroeconomic courses at M.I.T. and was always friendly towards me. I had the pleasure to teach macroeconomics from the draft of his new book.* **Professor Julio Rotemberg** *gave me the chance to learn business cycle theory as his research assistant. Finally,* **Fischer Black** *provided non-academic encouragement and it was a pleasure interacting with him.*

There are numerous others who along the way made the depressing days seem less so, including **Kobsak Pootrakool, Ilan Goldfajn, Luis Herrera, Li-Lan Cheng, Steve Levitt, Robert Shimer, Kolman Strumpf** for making the economics department more entertaining, **Jack Porter**, **Leo Kropywiansky**, **Walter Baily**, **Kaivan Munshi**, **Moshe Ben-Akiva**, **Koen Vermeylen**, **Lorenza Martinez**, **Christine Meyer**, **Kornelia Krajnyak**, **Gary King**, **Theresa Benevento**, **William Chang** *for computer assistance*, **Charles Hadlock** *for reminding me that "...you got to get out of there dude...", and all the students of my macroeconomic courses who made it worth getting up in the morning.*

Introduction

This thesis consists of three chapters that are quite distinct from each other. The first chapter is concerned with understanding the evolution of economic clusters observed in the world. Chapter two focuses on issues of labor migration. Chapter three studies the informational content of volume and returns in the stock market.

What are the dynamics of city formation?

The dynamics of city formation describe the process by which cities form or evolve over time. Over the centuries, we continually observe the rise and fall of cities. An important question to ask is why are there clusters of economic activity, rather than uniform dispersion of firms and people. This chapter investigates what processes of human behavior could lead to this spatial order.

The study of the spatial structure of economic activity dates back to Von Thünen (1826) and Fetter, R.A. (1924). Most of the work in this area, primarily carried out by economists and geographers, has focused on static models. These models, by definition, fail to explain how migration, capital movements, and *catastrophes* affect the spatial distribution of economic activity over time. In addition, these static models assume the centripetal forces that cause agglomeration rather than deriving these forces from microeconomic foundations.[1] In other words, the models simply stated

[1]The terms centripetal and centrifugal forces are borrowed from the physics literature for forces that are generated when a body is rotating about a point. The centrifugal force is an imaginary force that people speak of when referring to the feeling of being pulled outward. In the context of this chapter, centripetal forces refer to forces that tend to cause the agglomeration of economic clusters and centrifugal forces are those that tend to disperse economic activity.

that people agglomerate because there are externalities to being close to other firms or people. In 1991, newer models were built that created the increasing returns to agglomeration (centripetal forces) from microeconomic foundations (Krugman, 1991). Some of these models also made the first attempts at addressing the dynamics of the spatial structure.

What is the general economic intuition? Cities or economic clusters form because there are scale economies in locating where the market is. Thus, firms and workers will locate near each other in order to capture these benefits. One may think of these benefits in terms of labor pooling or spillovers of knowledge. Of course, the centrifugal force (pulling) causing more cities to form is also present. One can think of these forces as land rents, or any negative impact of overcrowding for that matter, to be the cause of movement away from an economic cluster.

Chapter one of this thesis expands on the previous work by using microeconomic foundations for agglomeration and by modelling the economy in a well structured and realistic manner. Using techniques of perturbation theory, this chapter creates a framework for understanding the formation of cities. The hexagonal spatial structure, edge cities, and the existence of cities *per se* are all easily explained. The principle is quite simple: start from an equilibrium distribution of workers over space. Distort the distribution of workers over the landscape into some chaotic pattern of your choice. Over time, an ordered spatial structure will emerge as workers migrate to higher wage regions. The final stage of evolution of the spatial structure will be stable and consist of a known number of economic clusters, where the number is determined by the degree of scale economies, the cost of transporting goods to the hinterland, the size of the spatial structure, and the percentage of firms for which there are increasing returns to agglomeration.

This research makes a lunge forward in understanding how such complicated subsystems can evolve into an ordered system based on simple underlying economics. Previous research relied on computer simulations to analyze the spatial distribution over time. My research gives interesting analytical solutions, thus formalizing the ideas of previous authors.

Geographic Mobility

If I pay an Italian three times his or her wage in Milano to move to Bangladesh to teach the locals computing, will he or she go? The mobility of individuals has always been a concern for economists. It is one of the ways that market imbalances can be corrected. If there is a bad shock in city A and everyone lies in the grass and smokes weed, the entire economy will suffer. If instead, the unemployed in city A depart to city B where there is an excess demand for labor, the economy will be less severely affected.

There has been a vast amount of research on migration. Due to the limited data available, most studies have focused on flows of people in and out of areas. With the availability of micro data sets, giving information on individual people, more studies have attempted to control for the individual specific effects on the decision to move. The key phrase is *the decision to move* . All of the previous studies, to the best of my knowledge, have focused on what factors induce one to leave a location of residence. Chapter two of this thesis focuses on another aspect of migration. Given that one has already decided to move, what characteristics of a location will be important to him or her in choosing that residence? The motivation for focusing on movers rather than on the decision to move is twofold. For one, I believe that there is a certain percentage of people in every location that would not leave home even if a bomb blew up in their backyard. Thus, I want to eliminate this type of bias on the estimation of the importance of various attributes in the decision to choose a location. The other motivation is directed more towards policy issues for local governments. Given that someone has decided to move, what will attract him or her to one locality over another and what factors are revealed preferred to migrants? Everyone wants low crime (or so they claim), but how important is it compared to going to see the Celtics at the Boston Garden? These issues are carefully examined in chapter two.

In addition to the issues already mentioned above, chapter two focuses on the movement of migrants to very specific geographical areas in the hope of extracting as much information about the factors that are important in choosing a locale of residence. This geographical preciseness, combined with some elegant econometric machinery, sheds light on the relative importance of various characteristics on the decision to immigrate. The results are important for theorists, especially concerning the assumptions that they make about people's migration patterns in their models of economic geography. The results are also of concern for local governments, especially if one believes that the immigration of certain types of people induces growth.

The main findings of chapter two are the following: the characteristics people claim to be important in a place of residence are different from the revealed preferences as indicated by where they actually do move. The migration patterns of the elderly and the young are slightly different. As one would expect, the elderly weight non-economic characteristics of a location more heavily than the young. Different races differ in their migration patterns, but more significantly, races prefer to move where there are larger percentages of their own race and tend to move away from other races. There does not seem to be a preference for racial diversity in choosing a locale of residence. Migration behavior does not seem to vary that much with educational attainment. Finally, homosexuals and foreigners are more sensitive to non-economic factors in choosing a locale of residence.

Differentiating Between Volume and Return

Is there a psychological element to trader behavior on the stock market? Are Wall Street traders *trigger happy*? Recent research in the finance trading literature has found support for this *overreaction* hypothesis.[2] Traders seem to overreact to news events. Investment strategies that sell winner stocks (stocks with positive return in that period) and buy loser stocks (stocks with negative return in that period) provide profits on average. This can be understood as follows. If traders do indeed overreact to news events, then they will drive the price of an asset above its true value on the announcement of unexpected good news and drive it too low on the announcement

[2]The *overreaction* hypothesis is the hypothesis that traders overreact to news events and thus the prices of the underlying securities will overshoot their equilibrium prices.

of unexpected bad news. Eventually, we expect traders or the market to learn the true value of the asset and prices should move towards their true values. This kind of behavior will generate a negative autocorrelation of returns over time. Whatever the relevant time frame, selling winner or high return stocks and buying loser or low return stocks will generate profits over time. This type of investment approach is known in the literature as pursuing *contrarian* strategies.

The literature of volume and return analysis has matured further with two theoretical articles (Blume *et al.*(1994), Campbell *et al.*(1993)). One describes traders and their information processing, while the other models how volume signals the underlying movements in other variables. The theoretical models cited above are different in scope, but both conclude that trading volume should contain important information not contained in an asset's price. The model of Blume *et al.* (1994) states that volume can refine the information contained in price to traders. The model of Campbell *et al.* (1993) is not a model of trader behavior, but rather a model that explains how volume can contain valuable information about the risk aversion of investors that is not always contained in the price statistic. The idea is the following: There are two types of shocks to securities. One is a news type of shock that alters the price of the underlying security depending on whether it was bad or good news. Another type of shock is due to random shifts in stock demand of noninformational traders that are accommodated by risk averse market makers. This type of shock also affects the stock price. In the former scenario, there is no reason to expect price movements in the future. The latter case will cause the expected return to be higher in the case of an increase in the risk aversion of the marginal investor. This makes sense. The expected return will have to be higher for the market makers to hold the position, and this occurs by a price decrease in the current period. Volume plays a role by indicating which shock has taken place. The first type of shock need not be accompanied by trading volume, while the second shock causes certain agents to trade away their positions. Thus, low returns accompanied by high trading volume are likely to yield positive expected returns. More precisely, the autocorrelation of the stock return should decline with trading volume.

The research that we undertake in chapter three of this thesis is catered to test the hypothesis that "price changes accompanied by high volume will tend to be reversed. This will be less true of price changes on days with low volume." We also test for the informational role of trading volume *per se*. We develop trading strategies which behave in a fashion that would result in expected profits over time if the theoretical models summarized above are indeed correct. We use volume and return information of stocks listed on the CSRP over the last thirty years and implement zero investment strategies that make use of both volume and return information as suggested by these models.

The main findings of the chapter are the following: trading volume does contain information that is different from that of price, and that the price of high volume stocks as well as the price for low volume stocks tends to reverse. Thus, while volume does seem to carry information separate from that of returns, it is unclear how adequate the model of Campbell *et al.* (1993) is at describing the informational process of volume within the time horizons we consider. We find price reversal for both high and low volume stocks.

Chapter 1

The Dynamics of City Formation

1.1 Introduction

The spatial organization of cities is in constant flux, yet there seems to be a regular ordering of cities or towns throughout the landscape. This chapter attempts to explain the formation of such seemingly structured economic clusters. This chapter builds on ideas that Paul Krugman has brought to the study of economic geography (Krugman, 1991, 1994). The main approach is to consider an economy that consists of workers and farmers and is located in space. There are increasing returns in producing manufactured goods and constant returns to scale in producing farm products. Workers are the only mobile factor in the economy. There are two forces acting on the spatial structure, increasing returns to scale which pull workers together and costs to distributing goods to the immobile farmers of the economy which pull workers apart. Everyone has tastes for all goods, i.e. Dixit-Stiglitz preferences. In the analysis, the economy begins at a uniformly distributed spatial structure. This uniform distribution of workers is slightly perturbed into some random distribution of workers over the spatial structure. Eventually, as the dynamics work themselves through, the economy arrives at an ordered distribution of workers. In particular, the order is similar to that observed in reality. An evenly spaced distribution of cities is predicted. This final distribution depends on transport costs, economies of scale, and the relevant land area. The major contribution to the literature is that no simulations are needed with this model. An analytical solution is derived. Given transport costs, economies of scale, and relevant land area, the model predicts the number of cities that will form. This number is consistent with the simulations that Krugman has performed with a discrete spatial model (Krugman, 1991). The other contribution is that for special cases of parameter values, two types of cities may be forming. One being a major city, while the other is a surrounding smaller city. Thus, the model offers an explanation of Edge Cities that have formed within the last thirty years. The chapter is organized as follows: section 1.2 discusses the previous literature in economic geography; section 1.3 introduces the spatial model and the various techniques used in the analysis of city formation; section 1.4 discusses how various parameters affect the distribution of economic clusters; section 1.5 discusses the model's ability and limitations in explaining the formation of Edge Cities; and section 1.6 concludes the chapter with some directions for further research.

1.2 Previous Work

In "A Dynamic Spatial Model" Krugman (1991) discusses how various centripetal and centrifugal forces interact to form agglomerations of economic activity. Due to increasing returns, it is advantageous to concentrate production at a few locations. Because of transportation costs, the best locations are those with good access to markets (backward linkages) and suppliers (forward linkages). These locations will be precisely where manufactures have located. But not all factors are mobile, and this is the centrifugal force of the system.[1] Market potential analysis has a long tradition in the literature of economic geography. The main theoretical weakness of it is the lack of microeconomic foundations. There is also central place theory which speaks about the tradeoff between transport costs and economies of scale. Central place theory was originally developed by the work of Walter Christaller (1933). He built models of spatial marketing by incorporating locational considerations into neo-

[1] One could imagine a more realistic model where land rents would serve this function; as an area became too expensive, people would spread out.

classical economic theory. Settlements are viewed as primarily local trade centers serving local hinterlands. In his work, he proposed a relationship between the number, size, and distribution of towns. His model predicted that central places should be surrounded by hexagonal service areas and that service areas are of equal size and located equi-distant from each other. The larger central places are located further apart and provide services that are demanded less frequently. These are also distributed evenly and are equal size, but serve larger areas. He advanced the possibility that transportation concerns could be at times more important than marketing in the locations of central cities, but never formally modelled this. In his empirical work, he concluded that southern Germany was a mix of the two forces. As to where the central places actually originated, he alluded to the accidents of history. Following his work, August Lösch (1954) argued that the pattern of hexagon formation was due entirely to the economic conditions of monopolistic competition created by space itself. He stated that the hexagonal formation was the ideal form of space when goods differed in the market area they require. Different goods will have different size hexagonal market areas. A shortcoming of these models was that the modelling of the economic incentives was very unsophisticated. The empirical work which was done primarily on the United States, China, South America, India, and West Africa turns out to confirm the central place conclusions. The formation of cities and towns follows some hexagonal structure. However, it is still unclear if the process causing these formations in that described by central place theory. This discussion is termed the "form-process confusion" and is discussed by Szymanski and Agnew (1981). Krugman's contribution to the literature was to model increasing returns with Dixit-Stiglitz monopolistic competition. This turns out to be a mechanism to model the ideas that previous authors were alluding to. Krugman looked at a two-locational model and a multi-locational model, and observed the stability or instability of a given concentration of workers. Simulations are used to understand how the economy would evolve given a certain beginning share of workers divided up among 12 regions. The eventual outcome depends on the parameters of the system. The basic point was that catastrophic dynamics can arise from fully rational economic agents in a fully general equilibrium model with all the trimmings.

1.3 The Model

I consider an economy located on a circle of radius r.[2] The workers are distributed along the circumference of the circle. The farmers are also distributed along the circumference of the circle, but are immobile.[3] In particular we will let $\lambda(x)$ denote the density of workers along the circle and $\psi(x)$ denote the density of farmers along the circumference of the circle. The distribution of workers is a function of x, where x is some point around the circle.[4] It is more convenient to represent $x = r\theta$. Also below, one should think of my z=$r\phi$. These are just the natural distances around the circumference of a circle expressed in terms of the radius and the angle dividing the locations. One can now re-write Krugman's economy in terms of a continuum along the circumference of the circle of radius r. Before doing this, however, I make a departure from the earlier Krugman paper. The earlier paper has a form of transport costs of the "iceberg" type. In particular, if y goods are shipped only $e^{-\tau|x-z|}y$ goods arrive to location x from location z. The problem with this transportation cost factor on a circle is that it brings in "tedious" problems mathematically and although these problems are eliminated along an infinite line as in Krugman (1994), this geometry raises other complications. Instead the transportation cost factor will be written as:

$$D(\theta, \phi) = e^{-2\tau r \sin^2\left(\frac{\phi-\theta}{2}\right)}. \tag{1.1}$$

This measure of distance avoids the problem of having to deal with absolute distances which can be quite tedious in finite space. See figure B-1 in appendix B. The distance measure continues to incorporate the idea that the further away a

[2] A more realistic two-dimensional model on a sphere could have been analyzed, however the mathematics are exhausting and one gains no new insights.

[3] One can really think of the farmers as residents and the workers as firms, where all residents produce workers that go to these firms and work costlessly.

[4] For clarity, see B-1 in appendix B.

location is, the higher the transport costs. The advantage it has is that when one revolves around the circle by 180°, the distance of those locations are not further away because goods will be shipped via the shorter route. The distance factor takes this into account, and while it does not represent distances exactly, it does so proportionately.[5]

The equations (Krugman, 1991) are translated from the discrete space case to the continuous case along the circumference of the circle. The equations are:[6]

$$Y(\theta,t) = (1-\mu)\psi(\theta,t) + \mu\lambda(\theta,t)w(\theta,t), \tag{1.2}$$

$$T(\theta,t) = \left[\int_0^{2\pi} \lambda(\phi,t)w(\phi,t)^{1-\sigma}e^{-2\tau(\sigma-1)r\sin^2\left(\frac{\phi-\theta}{2}\right)}r d\phi\right]^{\frac{1}{1-\sigma}}, \tag{1.3}$$

$$w(\theta,t) = \left[\int_0^{2\pi} Y(\phi,t)T(\phi,t)^{\sigma-1}e^{-2\tau(\sigma-1)r\sin^2\left(\frac{\phi-\theta}{2}\right)}r d\phi\right]^{\frac{1}{\sigma}}, \tag{1.4}$$

$$\omega(\theta,t) = w(\theta,t)T(\theta,t)^{-\mu}, \tag{1.5}$$

$$\bar{\omega}(t) = \int_0^{2\pi} \lambda(\theta,t)\omega(\theta,t) r d\theta, \tag{1.6}$$

$$\frac{\partial\lambda(\theta,t)}{\partial t} = \gamma\left[\omega(\theta,t) - \bar{\omega}(t)\right]\lambda(\theta,t). \tag{1.7}$$

So far this is just a transformation of the discrete space problem onto the circle in continuous space with a modified transportation factor. The exogenous variables of the system are the distribution of farmers and workers. From this one can calculate the above equations which will now be defined. $Y(\theta,t)$ is the income at each location. Given the assumption of zero transport costs of agricultural goods, the wage rate

[5]This choice of distant factor allows me to find an analytic solution for the problem and removes the need for simulations.

[6]All the variables depend on θ (the variable for distance since r is a parameter) which is the location around the circumference of the circle and t which is time.

of farmers is the same at all locations. μ is the fraction of the population that are workers, and therefore $1-\mu$ is the fraction that are farmers. Measuring all prices and wages in terms of the agricultural good, one obtains equation 1.2. $T(\theta,t)$ is the true or ideal price index of the manufactures aggregate to consumers at each location. This is the price of goods from different locations after arrival, thus accounting for wages and distance. Given these true price indices, one can solve for wage rates and obtain equation 1.4.[7] Because workers care about real wages in terms of manufactured goods as well, we say that the real wage, $\omega(\theta,t)$, will depend on the wage in terms of the agricultural good and in terms of the manufactures' price index. This gives us equation 1.5. Equation 1.6 is the mean wage of workers over all locations. Finally, equation 1.7 is the assumed law of motion of the workers in the economy. It implies that workers move to or move away from a region at some rate which depends on the real wage at that location relative to the mean wage of the economy.

Now, I wish to study the evolution of the system from equilibrium. First, I will find an initial equilibrium and then perturb it slightly and observe the corresponding dynamics. In particular, let $\lambda_0(\theta)$ be equal to $\frac{1}{2\pi r}$. This implies that we are starting from an initial uniform distribution of workers. It does not matter much for the analysis, but I will also assume that farmers are distributed uniformly, such that $\psi_0(\theta)=\frac{1}{2\pi r}$. The linear approximations to these equations will consist of the initial equilibrium and the perturbed part of the system. Below I collect zero a· d first order terms and assume that higher order terms are of negligible importance.[8]

$$Y_0(\theta)=(1-\mu)\psi_0(\theta)+\mu\lambda_0(\theta)w_0(\theta), \tag{1.8}$$

$$T_0(\theta)=\left[\int_0^{2\pi}\lambda_0(\phi)w_0(\phi)^{1-\sigma}e^{-2\tau(\sigma-1)r\sin^2\left(\frac{\phi-\theta}{2}\right)}r\,d\phi\right]^{\frac{1}{1-\sigma}}, \tag{1.9}$$

[7]See the appendix of Krugman (1991).

[8]See the appendix A, section A.1 for more details.

$$w_0(\theta) = \left[\int_0^{2\pi} Y_0(\phi) T_0(\phi)^{\sigma-1} e^{-2\tau(\sigma-1)r\sin^2\left(\frac{\phi-\theta}{2}\right)} r d\phi\right]^{\frac{1}{\sigma}}, \tag{1.10}$$

$$\omega_0(\theta) = w_0(\theta) T_0(\theta)^{-\mu}, \tag{1.11}$$

$$\bar{\omega}_0 = \int_0^{2\pi} \lambda_0(\theta)\omega_0(\theta) r d\theta, \tag{1.12}$$

$$\frac{\partial \lambda_0(\theta)}{\partial t} = \gamma\left[\omega_0(\theta) - \bar{\omega}_0\right]\lambda_0(\theta) = 0. \tag{1.13}$$

The first order equations depend on the equilibrium or initial conditions, in particular:

$$Y_1(\theta, t) = \mu\left[\lambda_0(\theta) w_1(\theta, t) + \lambda_1(\theta, t) w_0(\theta)\right], \tag{1.14}$$

$$\begin{aligned} T_1(\theta, t) &= \frac{1}{1-\sigma}\int_0^{2\pi} \lambda_1(\phi, t) w_0(\phi)^{1-\sigma} e^{-2\tau(\sigma-1)r\sin^2\left(\frac{\phi-\theta}{2}\right)} r d\phi + \\ &\int_0^{2\pi} \lambda_0(\phi)\frac{w_1(\phi, t)}{w_0(\phi)} w_0^{1-\sigma}(\phi) e^{-2\tau(\sigma-1)r\sin^2\left(\frac{\phi-\theta}{2}\right)} r d\phi, \end{aligned} \tag{1.15}$$

$$\begin{aligned} w_1(\theta, t) &= \frac{1}{\sigma}\int_0^{2\pi} Y_1(\phi, t) T_0(\phi)^{\sigma-1} e^{-2\tau(\sigma-1)r\sin^2\left(\frac{\phi-\theta}{2}\right)} r d\phi + \\ &\frac{\sigma-1}{\sigma}\int_0^{2\pi} Y_0(\phi)\frac{T_1(\phi, t)}{T_0(\phi)} T_0(\phi)^{\sigma-1} e^{-2\tau(\sigma-1)r\sin^2\left(\frac{\phi-\theta}{2}\right)} r d\phi, \end{aligned} \tag{1.16}$$

$$\omega_1(\theta, t) = w_1(\theta) T_0(r\theta)^{-\mu} - \mu\frac{T_1(\theta, t)}{T_0(\theta)} T_0^{-\mu}(\theta), \tag{1.17}$$

$$\bar{\omega}_1(t) = \int_0^{2\pi} \left[\lambda_1(\theta, t)\omega_0(\theta) + \lambda_0(\theta)\omega_1(\theta, t)\right] r d\theta, \tag{1.18}$$

$$\frac{\partial\lambda_1}{\partial t} = \gamma\left[\omega_1(\theta,t) - \bar{\omega}_1(t)\right]\lambda_0(\theta) + \gamma\left[\omega_0(\theta) - \bar{\omega}_0\right]\lambda_1(\theta,t). \tag{1.19}$$

These equations for the perturbations require an unecessary amount of algebra. To make things simpler, I will find the equilibrium conditions and then normalize the equilibrium so that for various variables their equilibrium value is equal to 1. This is done by multiplying by the appropriate constants and re-defining the equations of the system.[9] Below I write the new equations for 1.2-1.7:

$$Y(\theta,t) = \alpha\left[(1-\mu)\psi(\theta,t) + \mu\lambda(\theta,t)w(\theta,t)\right], \tag{1.20}$$

$$T(\theta,t) = \left[\beta\int_0^{2\pi}\lambda(\phi,t)w(\phi,t)^{1-\sigma}e^{-2\tau(\sigma-1)r\sin^2\left(\frac{\phi-\theta}{2}\right)}rd\phi\right]^{\frac{1}{1-\sigma}}, \tag{1.21}$$

$$w(\theta,t) = \left[\delta\int_0^{2\pi}Y(\phi,t)T(\phi,t)^{\sigma-1}e^{-2\tau(\sigma-1)r\sin^2\left(\frac{\phi-\theta}{2}\right)}rd\phi\right]^{\frac{1}{\sigma}}, \tag{1.22}$$

$$\omega(\theta,t) = w(\theta,t)T(\theta,t)^{-\mu}, \tag{1.23}$$

$$\bar{\omega}(t) = \int_0^{2\pi}\lambda(\theta,t)\omega(\theta,t)rd\theta, \tag{1.24}$$

$$\frac{\partial\lambda(\theta,t)}{\partial t} = \gamma\left[\omega(\theta) - \bar{\omega}\right]\lambda(\theta), \tag{1.25}$$

where $\alpha = 2\pi r$, $\beta = \frac{1}{\Gamma_0(\tau r(\sigma-1))}$, and $\delta = \frac{1}{2\pi r\Gamma_0(\tau r(\sigma-1))}$.[10]

Given our normalized equations, we have simpler expressions for the first order terms.

[9] Krugman does this, but unfortunately in his system, the equilibrium does not satisfy the equilibrium conditions. The normalization is presented in appendix A, section A.2.

[10] The function Γ_0 is defined in appendix A.

$$Y_1(\theta,t) = \alpha\mu\frac{w_1(\theta,t)}{2\pi r} + \alpha\mu\lambda_1(\theta,t), \tag{1.26}$$

$$T_1(\theta,t) = \frac{\beta}{1-\sigma}\int_0^{2\pi}\left[\lambda_1(\phi,t) + \frac{(1-\sigma)}{2\pi r}w_1(\phi,t)\right]e^{-2\tau(\sigma-1)r\sin^2\left(\frac{\phi-\theta}{2}\right)}r\,d\phi, \tag{1.27}$$

$$w_1(\theta,t) = \frac{\delta}{\sigma}\int_0^{2\pi}\left[Y_1(\phi,t) + (\sigma-1)T_1(\phi,t)\right]e^{-2\tau(\sigma-1)r\sin^2\left(\frac{\phi-\theta}{2}\right)}r\,d\phi, \tag{1.28}$$

$$\omega_1(\theta,t) = w_1(\theta,t) - \mu T_1(\theta,t), \tag{1.29}$$

$$\bar{\omega}_1(t) = \int_0^{2\pi}\left[\lambda_1(\theta,t) + \frac{\omega_1(\theta,t)}{2\pi r}\right]r\,d\theta, \tag{1.30}$$

$$\frac{\partial\lambda_1(\theta,t)}{\partial t} = \gamma\left[\omega_1(\theta,t) - \bar{\omega}_1(t)\right]\lambda_0(\theta). \tag{1.31}$$

The last equation comes from the definition of an equilibrium which requires that $\frac{\partial\lambda_0(\theta,t)}{\partial t}$ is equal to zero. It thus implies that $\omega_0(\theta) = \bar{\omega}_0$. As expected, the equilibrium wage rate is equal at all locations.

Now, I will propose a perturbation of the density of workers of the form:[11]

$$\lambda_1(\theta,t) = \sum_{m=-\infty}^{m=\infty} c_m(t)e^{im\theta} = \frac{a_0}{2} + \sum_{m=1}^{\infty}\left[a_m(t)\cos(m\theta) + b_m(t)\sin(m\theta)\right], \tag{1.32}$$

where

$$c_m(t) = \frac{1}{2\pi}\int_0^{2\pi} e^{-im\theta}\lambda_1(\theta,t)\,d\theta, \tag{1.33}$$

$$a_m(t) = \frac{1}{\pi}\int_0^{2\pi}\lambda_1(\theta,t)\cos(m\theta)\,d\theta, \tag{1.34}$$

[11] One could also have assumed a different form of perturbation, for instance $\lambda_1(\theta) = \rho\cos(m\theta)$, but the one chosen is more general. Also, one will note that ρ dictates the amplitude of the perturbation of workers, while m dictates the frequency.

$$b_m(t) = \frac{1}{\pi}\int_0^{2\pi} \lambda_1(\theta,t)\sin(m\theta)d\theta. \tag{1.35}$$

This perturbation is chosen specifically because it is the most general and realistic type of dis-equilibrium. It is the most general because any smooth function can be represented as a Fourier series expansion and it is the most realistic, because one would think that the perturbed spatial structure might actually be quite chaotic and complicated. Given this form of perturbation for the workers, the other variables of the system can be represented as multiples of the perturbation of workers. Specifically, I write the other variables as linear combinations of the perturbation of $\lambda(\theta,t)$.[12] Shortly, it will be clear that this is possible given the transportation factor chosen and the linear relation between the variables. Thus:

$$Y_1(\theta,t) = a_Y\lambda_1(\theta,t), \tag{1.36}$$

$$T_1(\theta,t) = a_T\lambda_1(\theta,t), \tag{1.37}$$

$$w_1(\theta,t) = a_w\lambda_1(\theta,t), \tag{1.38}$$

$$\omega_1(\theta,t) = a_\omega\lambda_1(\theta,t). \tag{1.39}$$

Using this form for the solution, one can substitute these expressions back into the first order equations and solve. One will repeatedly obtain an expression of the form:

$$K(\theta,t) = \int_0^{2\pi} c_m(t)e^{im\phi}e^{-2\tau(\sigma-1)r\sin^2\left(\frac{\phi-\theta}{2}\right)}d\phi. \tag{1.40}$$

[12]It would be interesting to perturb the distribution of farmers and allow for their mobility, in particular that they move away from high density areas. Due to the nature of the system, this cannot be done. Quadratic terms of the perturbed parts cancel from the equations. The only perturbations that could be allowed are multiples of $\lambda_1(\theta)$. These kind of perturbations would only enforce the agglomeration effect.

It is shown in appendix A, section A.4 that this equals:

$$K(\theta, t) = \underbrace{2\pi\Gamma_m\left(\tau r(\sigma-1)\right)}_{H(m,r,\tau,\sigma)} c_m(t)e^{im\theta}. \tag{1.41}$$

H is a function of the frequency of the wave, in other words it depends upon m. I will now define a new variable:

$$\eta = \frac{\Gamma_m(\tau r(\sigma-1))}{\Gamma_0(\tau r(\sigma-1))}. \tag{1.42}$$

Thus, $H(m, r, \tau, \sigma) = \frac{\eta}{\beta}$ and I will re-write the first order equations as:

$$a_Y = \frac{\alpha\mu}{2\pi r}a_w + \alpha\mu, \tag{1.43}$$

$$a_T = \frac{r\beta}{1-\sigma}\frac{\eta}{\beta} + \frac{r\beta a_w}{2\pi r}\frac{\eta}{\beta}, \tag{1.44}$$

$$a_w = \frac{r\delta a_Y}{\sigma}\frac{\eta}{\beta} + \frac{r\delta(\sigma-1)a_T}{\sigma}\frac{\eta}{\beta}, \tag{1.45}$$

$$a_\omega = a_w - \mu a_T. \tag{1.46}$$

One can solve these equations to give the final expression for a_ω:

$$a_\omega = 2\pi r\left[\frac{(1-\mu\eta)(\eta\mu-\eta^2)}{\sigma-\mu\eta-(\sigma-1)\eta^2} - \frac{\mu\eta}{1-\sigma}\right]. \tag{1.47}$$

This enables one to form an expression for the changing distribution of workers over time, in particular:

$$\frac{\partial\lambda_1(\theta,t)}{\partial t} = \gamma\left[\omega_1(\theta,t) - \bar{\omega}_1(t)\right]\lambda_0(\theta) \tag{1.48}$$

$$= \quad \frac{\gamma}{2\pi r}\left[a_\omega \lambda_1(\theta,t)\right] \tag{1.49}$$
$$= \quad \left(\frac{\gamma a_\omega}{2\pi r}\right)\lambda_1(\theta,t) \tag{1.50}$$
$$= \quad g_m \lambda_1(\theta,t), \tag{1.51}$$

where g_m is the growth rate of the fluctuation at that frequency. Note that $\bar{\omega}_1(t) = 0$.[13] Clearly, for the perturbations to grow, $\gamma > 0$.

1.4 Analyzing the Growth Factor

The equation for growth tells us that given certain parameters σ, μ, r, and τ, we will be able to measure the growth rate for different wavelengths or alternatively different frequencies. Given our parameters, there will exist some wavelength for which the growth factor, g_m, is a maximum. This is just the "preferred" wavelength. This tells us what the economy will eventually converge to without the need for simulations. This "preferred" frequency will ultimately dominate the spatial structure. From above, $g_m = \gamma\left[\frac{(1-\mu\eta)(\eta\mu-\eta^2)}{\sigma-\mu\eta-(\sigma-1)\eta^2} - \frac{\mu\eta}{1-\sigma}\right]$. Thus, the "preferred" frequency will be the m with the highest value of g_m. The preferred wavelength depends on various parameters, such as τ, σ, and μ. A higher σ implies that the elasticity of substitution between manufacturing products is higher. This indicates that there are less "economies of scale." There is less of a tendency for the economy to concentrate since products are more similar and thus the gains in terms of greater variety from concentration are lower. Hence, the preferred wavelength is smaller, and so there will be more cities with less accumulated mass in each one. A higher μ implies that the economy cares more about manufactured goods in their utility. *Ceteris paribus*, this leads to a larger preferred wavelength. The higher τ, the greater are the transportation costs, which would give the economy a greater tendency to locate in many regions, thus reducing the preferred wavelength. This is just the centrifugal force of the system.

The growth factor depends on m, but in the formulation throughout, m has been

[13]This is shown in section A.5 of appendix A.

restricted to being a positive integer (m=1,2,...). Thus, the integer value of m that gives the highest growth rate is our preferred wavelength (e.g. for values $\sigma = 2$, $\mu = 0.2$, $\tau = 0.2$, and $r = 1$, $m = 1$ is the preferred frequency, implying that one city or economic cluster will form). The effect of varying the transport costs, the share of income in manufacturing, and the elasticity of substitution or degree of increasing returns has been analyzed above. The tables below provide a summary of the results.

Table 1.1: Varying r and σ. With $\mu = 0.2$ and $\tau = 0.2$.

♠	$\sigma = 2$	$\sigma = 4$	$\sigma = 6$
r=1	$\frac{g_m}{\gamma} = 0.0249$ $m^* = 1$	$\frac{g_m}{\gamma} = 0.0128$ $m^* = 1$	$\frac{g_m}{\gamma} = 0.0059$ $m^* = 2$
r=2	$\frac{g_m}{\gamma} = 0.0396$ $m^* = 1$	$\frac{g_m}{\gamma} = 0.0117$ $m^* = 2$	$\frac{g_m}{\gamma} = 0.0068$ $m^* = 2$
r=3	$\frac{g_m}{\gamma} = 0.0448$ $m^* = 1$	$\frac{g_m}{\gamma} = 0.0133$ $m^* = 2$	$\frac{g_m}{\gamma} = 0.0080$ $m^* = 3$
r=4	$\frac{g_m}{\gamma} = 0.0191$ $m^* = 2$	$\frac{g_m}{\gamma} = 0.0111$ $m^* = 3$	$\frac{g_m}{\gamma} = 0.0070$ $m^* = 3$

The growth rates are given for the dominant frequencies given the parameters. m^* is the dominant frequency.

Table 1.2: Varying r and μ. With $\sigma = 4$ and $\tau = 0.2$.

♠	$\mu = 0.2$	$\mu = 0.4$	$\mu = 0.6$
r=1	$\frac{g_m}{\gamma} = 0.0128$ $m^* = 1$	$\frac{g_m}{\gamma} = 0.0462$ $m^* = 1$	$\frac{g_m}{\gamma} = 0.0782$ $m^* = 1$
r=2	$\frac{g_m}{\gamma} = 0.0117$ $m^* = 2$	$\frac{g_m}{\gamma} = 0.0531$ $m^* = 1$	$\frac{g_m}{\gamma} = 0.1132$ $m^* = 1$
r=3	$\frac{g_m}{\gamma} = 0.0133$ $m^* = 2$	$\frac{g_m}{\gamma} = 0.0440$ $m^* = 2$	$\frac{g_m}{\gamma} = 0.1216$ $m^* = 1$
r=4	$\frac{g_m}{\gamma} = 0.0111$ $m^* = 3$	$\frac{g_m}{\gamma} = 0.0522$ $m^* = 2$	$\frac{g_m}{\gamma} = 0.1163$ $m^* = 1$

The growth rates are given for the dominant frequencies given the parameters. m^* is the dominant frequency.

Table 1.3: Varying r and τ. With $\mu = 0.2$ and $\sigma = 4$.

♠	$\tau = 0.2$	$\tau = 0.4$	$\tau = 0.6$
r=1	$\frac{g_m}{\gamma} = 0.0128$ $m^* = 1$	$\frac{g_m}{\gamma} = 0.0019$ $m^* = 2$	$\frac{g_m}{\gamma} = 0.0021$ $m^* = 2$
r=2	$\frac{g_m}{\gamma} = 0.0117$ $m^* = 2$	$\frac{g_m}{\gamma} = 0.0114$ $m^* = 3$	$\frac{g_m}{\gamma} = 0.0134$ $m^* = 3$
r=3	$\frac{g_m}{\gamma} = 0.0133$ $m^* = 2$	$\frac{g_m}{\gamma} = 0.0134$ $m^* = 3$	$\frac{g_m}{\gamma} = 0.0135$ $m^* = 4$
r=4	$\frac{g_m}{\gamma} = 0.0111$ $m^* = 3$	$\frac{g_m}{\gamma} = 0.0128$ $m^* = 4$	$\frac{g_m}{\gamma} = 0.0125$ $m^* = 5$

The growth rates are given for the dominant frequencies given the parameters. m^* is the dominant frequency.

Another twist to this model is varying r; the radius of the circle or alternatively the final distance between each partner and its furthest location. Figure B-2 in appendix B is more illustrative (the figure uses values of $\sigma = 4$, $\tau = 0.2$, $\mu = 0.2$). When r is increased, the m with the highest growth will be higher.[14] This is quite intuitive. As r increases, the spatial structure spreads apart. This makes it more likely that more "central economic structures" should be formed since the benefits from previous structure of agglomeration will be less efficient due to the increasing centrifugal forces arising from the increased costs of transportation; hence it becomes optimal to establish more centers. Taking limits of r, one obtains results about the growth. In particular, as $r \longrightarrow \infty$ or $\tau \longrightarrow \infty \Longrightarrow \eta \longrightarrow 1 \Longrightarrow \frac{g_m}{\gamma} \longrightarrow \frac{\sigma-1-\mu\sigma}{1-\sigma}$. In other words, as either the area of concern tends to infinity or transport costs tend to infinity, there will be no growth unless $\mu > \frac{\sigma-1}{\sigma}$. There may be no growth, if the costs to deliver goods outweigh any benefits in terms of economies of scale. As $r \longrightarrow 0$ or $\tau \longrightarrow 0 \Longrightarrow \eta \longrightarrow \frac{\left(\frac{r(\sigma-1)\tau}{2}\right)^m}{m!}$. One city will form since the centrifugal forces are not important. This term in general will be small, but even smaller for higher frequencies, thus in the limit one city will form. One will notice that the radius of the circle and transportation costs affect the model in the same manner.

There is a very interesting case when for certain r or certain τ, there are two "preferred" wavelengths. This case is very interesting, since it will predict that we can have a very complicated evolution of the spatial structure. Nothing definite can be said about the evolution, since our linear theory cannot say much about the effect of amplitude. However, it might possibly result in a landscape similar to those depicted in figures B-3, B-4, and in two-dimensions, figure B-5. This is the Edge City case.

1.5 The Emergence of Edge Cities

The model outlined above does not seem to explain the formation of Edge Cities entirely satisfactorily. The centripetal forces cause agglomeration of workers which, once the dynamics have worked themselves through, results in only a few selected

[14]In tables 1.1, 1.2, and 1.3, the maximum m is expressed as m^*.

sights of concentration and the rest of the area being inhabited by farmers. Even if one interprets these farmers as residents of some sort, there is only a vague resemblance to the Edge City phenomena. Edge Cities form primarily out of residents and firms eventually moving out of the central city towards the suburbs or fringes of the metropolitan area. The primary reasons for this movement are due to increased land rents in the central city and the attractiveness of all the features that come from a less densely populated area.[15] In the model presented, agglomerations are formed due to higher relative wages where there are more workers or firms. The ordered structure that emerges can be thought of as explaining the regular spacing pattern of cities. One could really only apply this model to Edge Cities, if one considered that the real wage incorporated various other factors that are offered in Edge Cities. Regardless, this is not the case. What is needed to make this model a more realistic dynamic portrait of Edge City formation is land rents. As workers or firms move into an area, there must be some function which enters negatively into the residents utility function and at some point even the firms'. One must also consider population growth in a region, which in this model is assumed to be zero. These factors would cause agglomerations to form, but then a saturation point would eventually be reached and correspondingly Edge Cities could form. The problem with investigating these ideas in the perturbation mode is that amplitudes and quadratic terms are not important. One would really have to investigate the non-linear system as it is.

The model, which is built upon the earlier models of Krugman, has an important weakness by not having a saturation point. The non-linear system itself never reaches a new equilibrium *per se.* The dynamics stop once there is an infinite amount of workers in an empty space, in other words when everyone is located in an infinitely dense space. In a discrete case, this might seem plausible. Everyone moves until the locations have divided up the people. Obviously, this is unsatisfactory and the model described in the previous paragraph is needed. One could explicitly introduce saturation by allowing for the land rents as a quadratic term. Thus a equation of motion

[15]See the glamorized description of Edge Cities in appendix C.

of the following form: $\frac{\partial\lambda(\theta,t)}{\partial t} = \gamma[\omega(\theta,t) - \bar{\omega}(t)][\lambda(\theta,t) - a(\theta)\lambda^2(\theta,t)]$.[16] In the linear model as it is, it will only effect the quantitative results, not the qualitative nature and thus I have not pursued this line of analysis. In the non-linear setting however, it would be one step in the right direction for economic geography models to allow for saturation. The basic point of this preliminary discussion is that the literature in economic geography misses a major ingredient to understand any dynamics of city formation.

There is one special case that would explain the emergence of Edge Cities in the spatial structure of various countries throughout the world. That is in the special case where two frequencies both have the dominant growth rate. In this case, two "preferred" wavelengths will grow and dominate the spatial structure. One cannot be certain about the emerging structure, however it is quite possible that a structure similar to those in figures B-3, B-4, and B-5 would emerge.[17] This case would correspond well with the Edge City phenomena. Very large cities and very small cities may coexist, both agglomerating for the increasing returns, yet trading off the cost of transport differently. This special case of this theory provides an explanation for Edge Cities. The theory is also consistent with the fact that there are less Edge Cities in Europe than in the United States. The process of labor movement is crucial is the dynamics of these agglomerations. It has been documented that the Europeans respond less to wage differentials that do Americans.[18] Hence, in terms of the model, $\gamma_{EC} < \gamma_{USA}$. Thus, Edge Cities will take a longer time to form in Europe. This same explanation might be used to explain why Edge Cities have only emerged in the last thirty years.

[16] $a(\theta)$ allows for a saturation at some other distribution of workers.

[17] For the graphs, the following plot was used: $\lambda_1(\theta) = \alpha\cos(m_1\theta) + \beta\cos(m_2\theta)$, where α and β are chosen amplitudes and m_1 and m_2 are the "preferred" frequencies.

[18] See Padoa-Schioppa (1990).

1.6 Conclusion

Non-linearities make closed form solutions difficult in dynamic models of economic geography. However, using linear approximations as presented in this chapter, a clearer pattern of city formation evolves. This chapter uses a Krugman style economy to bring understanding to the dynamics of city formation. It is the typical battle between increasing returns to agglomeration and the transport cost factors that tend to disperse economic activity. We see that even when the economic landscape is quite disordered, workers and firms will eventually relocate so as to maximize their welfare in terms of their real wage. This chapter explains the regular spacing of cities that was observed originally by Christaller and continues to exist. Using new techniques of perturbation theory and a new distance factor, one is able to determine the number of cities that form throughout the landscape without the need for simulations. The model is not inconsistent with the formation of Edge Cities within the last thirty years. In fact, there are cases where the model specifically predicts that these Edge Cities will form. The chapter also outlines some potential problems with the current models of economic geography and constructing better models that incorporate these real phenomena of saturation should be the exercise of future research.

Chapter 2

On the Causes of Geographic Mobility

2.1 Introduction

Recently, there has been an explosion of theoretical papers in the area of economic geography. Many of these papers focus on the push and pull factors which tend to concentrate economic activity or disperse economic activity. These papers make assumptions about which forces cause concentration and which cause dispersion. Most of them agree that there is some kind of increasing returns to scale in locating businesses next to each other. However there is no consensus as to what causes the spreading out of economic activity. Paul Krugman stresses transportation factors (Krugman, 1991) such as dispersing the goods to a agriculture hinterland. Heller highlights the role of pollution, crime, and other consequences of crowding that cause people to move out of cities. Von Thünen emphasizes the idea of land rents which increase as the density in a location increases, causing people to move out. There has been a lot of empirical work on what forces cause emigration and what forces cause immigration. This chapter takes a new approach to identify what factors are causing people to move from one location to another. I make use of a data set which asks the question if people have moved within the last five years of the questionnaire and from where. This combined with data on the attributes of the cities they moved to and from gives me ample information to infer what causes people to migrate from one place to another. This chapter relies on a *revealed preference* type argument. We do not know *exactly* why a person moved from city A to city B. However we do know the attributes of the two cities. Thus, in moving from city A to city B, the person has selected one over the other, and after controlling for various factors, we must infer that $U_i(A(Z)) \geq U_i(B(Z))$ where i indicates the person and Z represent attributes of the locations, such as crime, pollution, average wages, density, etc. This chapter takes a more general framework and considers more than just one city A and one city B, in fact it considers a set of J cities. Using an econometric technique developed by McFadden (1976), I estimate the importance of these various attributes, Z, causing people to move from one location to another. This chapter also uses the richness of the data set to access differences across people in the decision to choose a location of residence. The analysis presented has policy implications and reveals more about the causes of migration.

This chapter is organized as follows: section 2.2 discusses the previous research on migration and the motivation for the work that I pursue in this chapter; section 2.3 introduces the model and estimation methods to be used in the analysis of migration behavior; section 2.4 discusses the data that was used for the analysis and the conditional logit regression results; section 2.5 compares the revealed best places to live with those given in popular journal surveys; section 2.6 discusses the possible extension of the work in this chapter for future research; and finally section 2.7 concludes this chapter.

2.2 Previous Research and Motivation

The papers that exam the factors that cause people to migrate from one region to another can be divided into those that use flow data of migrants and those that make use of micro datasets. One of the most extensive reviews of the migration literature comes from Greenwood (1975). He reviews some of the important aspects of migration. These include the effects of distance, the effects of income, the psychic costs of migration, information about destinations and migration, and personal characteristics and the decision to migrate. He presents the general conclusions regarding the causes of migration as available in the literature.[1] The general results are that gross migration declines with distance, that unemployment rates have an ambiguous sign, and that Hick's notion that wage differences are the primary cause of migration has *not* been confirmed.[2] Many studies suggest significant differences between the migratory patterns of blacks and whites, although these studies *fail* to control for age, education, and employment factors. Greenwood also mentions the failure of many studies to include *end of the period* location variables in the regressions.[3] In addition, Greenwood discusses the consequences of migration, which will not be discussed here since they are not directly related to the purpose of this chapter.

Kahley (1991) provides an elegant summary of the research on population migration. Common personal characteristics have been shown to affect migration: people in their twenties tend to move more than other age groups, more men move than women, the unemployed tend to migrate more than the employed, and educated people migrate more than non-educated people. Some common destination characteristics that induce migration to a region are job opportunities and wages, a lower cost of living, low state and local taxes, and good public services, such as police and fire protection. Kahley cites some unsolved migration issues such as the asymmetries between immigration and emigration, the data limitations with flow data[4], problems with model specification.[5] For example many studies have found that unemployment rates in a region have insignificant and even unanticipated signs, suggesting perhaps model misspecification of some sort. State and local fiscal variables and climate variables also have had ambiguous statistical results.

[1] Some of these conclusions have been disputed by work done after 1975.

[2] Only when one considers the real wage to be more than just wage adjusted for living costs does it appear to be significant. In other words, one must consider amenities, such as attractive climates and friends and family.

[3] This causes endogeneity of the regressors and leads to bias in the estimates.

[4] These problems are now less important due to the availability of microdata sets, in particular the Panel Study of Income Dynamics (PSID) and the Public Use Microdata Samples (PUMS).

[5] For example, population tends to attract migrants, but is that an economies of scale argument or is population just an indicator of an area that has had past success in attracting migrants?

Lundborg (1991) studies the differences in migration behavior across the Nordic populations. He finds that Finnish migration to Sweden is mainly caused by real wage differentials. However, it does seem that the Finnish respond less to unemployment rates than do the Danish or Norwegian nationals.[6]

Most of the studies rely on flow data, but recently the availability of microdata on individuals has allowed researchers to examine the differences among people that decide to move. Fox, Herzog, and Schlottman (1988) study the effects of fiscal structure on the decision to migrate. They use data on metropolitan areas from the the public use microdata sample of 1980. They focus on the decision to migrate or not. They find that fiscal variables are more important for *pushing* people out from an area than for *pulling* them into an area.[7] Among these factors, sales, income, and property taxes all increase the likelihood of departing from a location, expenditures on education and parks and recreation decrease the probability of moving, and finally welfare expenditures by the local government increase the tendency to move from an area.[8]

Berger and Blomquist (1992) focus more specifically on the quality of life aspects of migration. They use microdata from the 1980 census. They restrict their sample to 250 counties of the United States and use a multilogit model to access the importance of various factors on the decision to move. They analyze how various locational factors affect the individual's decision to move. They find that the quality of life variables are not significant in deciding whether or not to move. However, once an individual has decided to move, they find that quality of life variables become important. They also find that moving costs are not important in the decision to migrate.

Spilimbergo and Ubeda (1993) use a panel data set of individuals to understand the causes of migration. They study the impact of various factors in one's decision to

[6]This will be part of the motivation to only work with movers in my chapter. Most migration studies concentrate on the issue of whether to migrate or not. This is another important concern, however in my study I want to eliminate the bias caused by people that will not move regardless of the situation. I choose to concentrate on movers so that the *home bias* effect is less prevalent.

[7]The authors believe this to be an informational asymmetry.

[8]They allude to the idea that local governments that use progressive taxes to transfer income will actually induce emigration by the sources of revenue, which tend to be higher educated people. They do not actually test this proposition. I shall examine it briefly in the work to follow.

move from a certain metropolitan area. They find that a higher degree of education increases the probability of moving and so does being younger than 24 years of age. On the other hand, being married or owning a house discourages moving. They also find that higher income tends to discourage migration and that people who have migrated before tend to migrate again. The main focus of their paper is on family attachment and they do find that there is inertia in moving from an area containing much of one's extended family. They find that this attachment is stronger for blacks than whites. Thus blacks tend to respond less to economic incentives and tend to remain in a *worse* economic state rather than move.[9]

A specific area of interest has been trying to understand the migration decisions of elderly people. Serow (1987) examines the interstate mobility of the young and the elderly between 1975 and 1980. Studying both immigration and emigration, he finds that there is a substantial difference in the factors that influence young and old migrants. His findings are that the elderly tend to move to areas where other elderly people already reside, that the real cost of living tends to discourage migration to a region, that the elderly tended to move towards higher areas of unemployment. Serow also found that the climate did not significantly affect the decision of elderly migrants, that the supply of health care was negatively related, and that the growth in real income of a region in previous years was a positive determinant for elderly movers. For people between five and fifty-four years of age, he finds that median age has no effect, real wages are insignificant, unemployment tends to discourage migration to a region, that the climate was insignificant, the supply of health care discouraged migration to a region, and that the growth in real income in previous years was a positive determinant of immigration. The main conclusion is that the elderly and the young have quite different reactions to observed locational characteristics.

Biggar, Longino, and Yeatts (1987) investigate the concept of distance and the willingness to move to a region. They conclude that the destination makes quite a

[9]They also present evidence that responses to polls asking people why they might move and the actual reasons why they do move are not strongly correlated. This castes some doubt on the usefulness of using surveys to decide on the best places to live.

difference as to where people move regardless of distance.[10] Essentially, this means that people care about the characteristics of the location they move to and whether it is very far from the original destination or not is much less relevant.

Litwak and Longino (1987) studied the differences among the elderly in moving from metropolitan areas to non-metropolitan areas and vice versa. They found that disabled and widowed people tend to move toward metropolitan areas, reflecting the greater health facilities there. They only present averages and it would be interesting to study the significance of their results.

Golant (1990) studies regional migration patterns of the elderly in the United States for the periods 1955-1960, 1965-1970, 1975-1980, and 1980-1985.[11] He found that the overall mobility of the elderly declined over this period and that the tendency for the elderly to move out of a certain region declined. However, of those elderly migrants that did leave their region, a higher percentage moved to the south. His findings are difficult to compare directly to other studies, since for the 1980-1985 time period he used Current Population Surveys which are smaller size samples of the total population as compared to the census. Future research should determine the robustness of these findings.

Serow and Sly (1991) present a survey of some of the existing literature on elderly migration. They mention the work of Ring (1988) who found that amenities are of greatest importance in the decision to move in geographically larger countries, such as the United States, Australia, and Canada. Several papers have examined survey data for elderly individuals within countries (Serow (1987b), Friedrich and Koch (1988), and Speare and Meyer (1988)). In the United States' survey, fourteen percent of the elderly move for employment related reasons, thirty-three percent for friends and family reasons, thirty-two percent for the environment, and twenty-one percent for other reasons.[12] These responses were quite similar across countries. However in

[10]This result enables one to neglect distant measures in future work.

[11]The periods are divided in five year intervals according to the availability of information on migrants as provided by the census.

[12]This survey was given from 1974-1976 and the questionnaires were given to people older than 55.

Japan, Germany, and Belgium people were less inclined to state that the environment had anything to do with their decision to move.[13]

Cebula (1993) focuses on the interstate migration of the elderly primarily so that he can avoid using labor market variables that might introduce multicollinearity. He finds that both young and older people tend to be discouraged from areas with high living costs as measured by the index of McMahon (1992). He also finds that the elderly tend to prefer areas with more sunlight, less crime, a higher concentration of people their own age, and less unemployment.

The motivation for the research that I conduct came from various sources. The initial inspiration came from the theoretical literature of economic geography. Various authors modelled population or worker movement by assuming certain processes for the centripetal and centrifugal forces (Chincarini (1995), Krugman (1991), and Heller (1993)) that govern the growth of economic clusters. This chapter tries to identify which forces can realistically be assumed in the economic geography models.

Every year *Business Week* publishes a survey entitled "The Best Places to Live in America." Such categories as crime, health care, housing, education, weather, and leisure are ranked in order of importance. *Ex ante* we can all describe paradise; it has the best of everything. In reality, consumers and residents must make choices. They must choose between having more pollution and higher wages or less crime but higher taxes. This chapter approaches the question of the best places to live from another angle, that of *revealed preference.* By observing where movers choose to live relative to where they could have chosen to live, we obtain their revealed preferences as to where are the "best" places to live.

Given the demographic transition of most world economies, the differences between the migration patterns of the elderly and those of the young will play a major role in the spatial structure. Issues such as these are investigated in the section on elderly migration. Many migration studies ignore the aspects of policy on the inflow of labor. This chapter tries to focus specifically on the migration patterns of the highly educated versus those of the less educated. Is it actually important, *ceteris*

[13]This also might be due to the lack of heterogeneity in the environment of these countries.

paribus, for the local governments to consider the long term effects of tax policies and to a lesser extent environmental policies?[14] Finally, the richness of the data set and the conditional logit estimation procedure allow me to adequately correct for some of the errors in previous studies, for example in studying the differences between black and white migrants.

2.3 The Model

Given that a certain individual decides to move, it is assumed that this person or household will look at the alternative locations available for the move and decide which one to choose based an evaluation of the characteristics of these different locations.[15] Thus, the probability of household i choosing a certain location or county j from a set J_n can be represented as:[16]

$$P_{ij} = Prob(U_{ij} > U_{ik}) \; \forall k \; in \; J_n \; where \; k \neq j. \tag{2.1}$$

A conditional logit model will be used to access the importance of various characteristics on location choice.[17] A set of dummy variables will be created to allow for individual specific effects. The log-likelihood function is:

[14]One major reason that city and county government are opposed to environmental regulation is that they fear it will drive businesses away (Ward(1990)). If the migration of the highly educated or highly skilled is to areas with lower pollution, this aspect may not be as negative as they believe. Rick Tagtow (1990) discusses the many advantages of trees and greenery on urban life.

[15]One might worry about informational problems due to giving equal weight to locations that are very far from the initial residence. However, work by Biggar, Longino, and Yeatts (1987) suggests that distance is not such an important factor.

[16]McFadden (1975) showed how this can be derived from utility theory. In particular, he proposes that the utility of individual i, $U_{ij} = \mu_{ij} + \epsilon_{ij}$, where μ_{ij} is a nonstochastic function of explanatory variables and unknown parameters and ϵ_{ij} is an unobservable random variable. Furthermore, $\mu_{ij} = \alpha + \mathbf{z}'_{ij}\beta + \mathbf{w}'_i\gamma$, where $\mathbf{z}_{ij}$ is a vector of county characteristics and $\mathbf{w}_i$ is a vector of the ith person's socioeconomic characteristics. The probability is expressed as: $P_{ij} = Prob(Y_{ij} = 1) = \frac{e^{\theta' W_{ij}}}{\sum_{k=1}^{m} e^{\theta' W_{ik}}}$.

[17]For a discussion on the appropriateness of using this kind of utility based theory, see Lin-Yuan, Yihua et. al. (1994).

$$\log L = \sum_{i=1}^{n}\sum_{j=1}^{m} Y_{ij} \log P_{ij} \tag{2.2}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} Y_{ij}(\theta' W_{ij}) - \sum_{i=1}^{n} \log\left(\sum_{k=1}^{m} e^{\theta' W_{ik}}\right), \tag{2.3}$$

where W_{ij} is $(k+m)$ dimensional vector which equals $[X_{ij}Z_{ij}]'$. X_{ij} is a $k \times 1$ vector of various attributes for location j, Z_{ij} is an m dimensional vector with Z_i as the jth element for each characteristic choice and all other elements being 0. θ is a $(k+m)$ dimensional vector of coefficients, which equals $[\beta\alpha]'$. β is a k dimensional vector representing the coefficients of the attributes of the locations and α is an m dimensional vector of coefficients for various individual characteristics, Y_{ij} is an indicator variable that takes on the value of 1 if the ith individual chooses the jth location and 0 otherwise, n is the number of observations or people moving in the sample, and m is the number of different locations people can choose from.

In order to find the coefficients, one chooses estimates so as to maximize the log-likelihood function. Taking first derivatives and setting them equal to zero:

$$\frac{\partial \log L}{\partial \theta} = \sum_{i=1}^{n} W_i(Y_i - P_i) = 0. \tag{2.4}$$

Taking second derivatives and expectations:

$$-E\left(\frac{\partial^2 \log L}{\partial\theta\partial\theta}\right) = E\left(\frac{\partial \log L}{\partial \theta}\frac{\partial \log L}{\partial \theta'}\right) = \sum_{i=1}^{n} X_i A_i X_i', \tag{2.5}$$

where $A_i = E(Y_i - P_i)(Y_i - P_i)' = D(P_i) - P_i P_i'$ and $D(P_i)$ is a diagonal matrix with its jth diagonal element equal to P_{ij}. Hence, the covariance matrix of $\hat{\beta}_{ML}$ is $(\sum_{i=1}^{n} X_i A_i X_i')^{-1}$. The maximum likelihood estimates are obtained by using the Newton-Raphson method.[18]

[18]This is the method that STATA uses to find the maximum of these nonlinear functions.

2.4 Empirical Analysis

2.4.1 The Data

The information on individuals or households and their mobility decisions was obtained from the *1990 Census of Population and Housing.*[19] This data set contains information on individuals in over 1000 geographic units across the United States.[20] The important feature of this data set is that it contains information on individuals' place of residence five years prior to 1990. This, combined with information about the PUMAs (public use micro areas) prior to 1985, can be used to estimate a conditional logit model of what characteristics of locations and individuals cause people to choose the locations they move to. This ordering of the data avoids any endogeneity problems that are sometimes present in studies of migration. A sub-sample of households was chosen that had moved from one PUMA to another.[21] I chose to observe the impact of locational characteristics only on movers to eliminate the problems of dealing with people who would never move regardless of the situation. Also, to avoid sampling individuals with obvious mobility constraints, I removed disabled individuals, students, those serving in the military, and those being hospitalized for mental reasons.[22] These PUMAs are usually combinations of various counties and/or cities in a particular region within a state. There are instances when a county that is very large is split into many different PUMAs.[23] The data for individual movers was quite large and due to computer limitations I sampled the data further by taking one percent of all the movers in the census.[24] The characteristics of various PUMAs were obtained from the *1988 County and City Databook* by aggregating the counties/cities

[19]I used the five percent sample data base.

[20]These units are at times several counties, a single county, or even a city or group of cities. These PUMAs will be more carefully defined further on in the text.

[21]Thus, I remove non-movers and movers within the same geographical area between 1985-1990.

[22]The summary statistics are presented in appendix E.

[23]There is another government document (not available yet) entitled the *County to County Migration File.* It has all the characteristics of people and their place of residence in 1985 and their current residence in 1990, and it might be available in the early Fall of 1994.

[24]The one percent sample was obtained from each state individually and then the individual movers to each state were collected into one larger data set.

to get the respective PUMA. The actual variables that are used from this data set are included in the appendix. Weather data was obtained from the *Climate Normals of the United States* (1991). Pollution data was obtained from Professor Matt Kahn.[25]

There are some limitations to the analysis. The PUMS (public use microdata sample) does not contain a question relating to why the person actually moved. Thus, this study really must rely on the *revealed preference* type of argument. One objection to this approach is that many people might move to a region only because they were offered a job there and for no other reason. I will not be able to detect this in my estimations. I claim, however, that this should only dampen any relationships between location characteristics and the decision to move to that region. Perhaps, unemployment could be used as an indicator to partially account for this.[26] One reason economists usually give for people's mobility decision is wage compensation. It would be valuable to have data on the actual wages people were paid, but they do not exist. Instead, I use various measures of average wage of a region to account for that motivation.[27]

Some of the locational features that might attract or retract individuals from different locations are discussed below. For a detailed description of the variables used in the empirical work, see appendix D. (i) The general economic argument is that people move to different regions when wage differentials exist. In fact, the basic models by Krugman make a simplistic assumption that people move to areas where the real wage is above the *real mean wage* of the entire region. In reality, one believes that people will move if they have a wage offer that is higher is one location. However, this chapter tries to explain location to location migration based upon the *generalized* characteristics of a location. The actual data on counties gives personal per capita income for 1984 and money income per capita for 1979 and 1985. It also has manufacturing wages for a representative manufacturing worker in 1982.

[25]Matt Kahn is an assistant professor at Columbia University. I gratefully acknowledge the access to his data set.

[26]Athanasios Orphanides suggested using vacancy rates in various areas. I was not able to find this data, but feel that unemployment might be a sufficient control.

[27]See appendix D.

To create a real wage variable, I divide personal income per capita by mean housing costs in the region.[28] (ii) *Taxes* must be a consideration when deciding where to live, which again really relate to real income. I also consider the possibility of *future taxes* by using the proxy of current government debt per capita. (iii) The *future growth* of a region or the future prospects would seem to be a motivation for one to locate there. Although there is no such variable, past growth might proxy for this if individuals are adaptive and use a geometric weighting of past growth to form their expectations about future growth. New residential construction might be another indicator of future growth. (iv) *Unemployment* can be thought of as a signal for the low probability of job opportunities in a region and thus discourage people from moving there. (v) Heller constructs his theoretical paper around the idea that various factors such as pollution, crime, and other sociological factors cause people to leave an area. Thus, *crime* and *pollution* would be natural variables to consider. Less of both would make a place more attractive to migrants.[29] (vi) *Education* might be an indication of the sophistication of an area. An area with a higher percentage of educated people might attract more educated people. The quality of schools is also important. A variable such as the number of teachers per pupil or average SAT scores could measure this. Unfortunately, these were not available for all of the areas I wanted to consider. I chose to use government expenditure on education as a proxy.[30] (vii) *Climate* must be a concern for migrants. This variable might be more important for elderly migrants, since they are less concerned with the employment aspects of migration.[31] (viii) The racial composition of an area might be quite important for

[28]Various other sources for cost of living indices exist, however they are not geographically specific enough for the purposes of this chapter.

[29]I will use violent crimes per 100,000 people as a measure of crime. One might also want to investigate the total number of crimes as well. It is an interesting issue whether people are discouraged more by total crime or per capita crime. There is a psychological aspect to this discussion that makes probability theory not as relevant. I actually find that total crimes are insignificant on the decision to move to a region.

[30]This variable will be endogenous and is not entirely satisfactory. There might be higher government expenditure on education precisely, because the school system is so bad. Expansion Management in Overland, Kansas has a very good data set from 1990-1994 for education quality in various regions. However, as this is not in the time period of my study, I am unable to use it.

[31]I use the mean weather and rain fall in various regions. A more in depth study would consider the variability in the weather and rather than temperature, the number of clear days, even though

people. One *race* may want to migrate to areas with a higher percentage of their own race. Also, there might be a specific race that all other races are trying to avoid. Thus, I use an indicator variable for the percent of various races in an area, such as Asian, black, Hispanic, Indian, and white. (ix) The percent of people below poverty level, the percent of single parent households, and the percent of single households might also be important for mobility decisions. These factors may reveal the economic and moral weakness of an area. (x) *Health* facilities will be an issue for people, more so for the elderly. I use physicians per 100,000 and hospital beds per 100,000 as measures of health care availability.[32] (xi) I would like to have some sort of measure of product variety or diversity which attracts migrants looking for an array of goods. Due to data limitations, I will use a central city indicator to proxy for these facilities.[33] There are other variables such as density that try to account for the spillovers that Marshall originally spoke about.

2.4.2 The Technique

In section 2.3, the conditional logit model to be used for estimation was discussed. Before using this estimation procedure some modifications to the traditional estimation are in order. There are 1018 PUMA choices. These are too many for the system limitations of the computer programs that I used. To overcome these problems, I chose to sub-sample among the data choosing randomly 509 of the available 1018 PUMAs. Therefore, for every move consider only 509 destinations. This is still too large a sample for computational purposes, thus I allow each individual his or her actual choice destination and a random selection of 4 of the 508 remaining choices that the person could have chosen.[34] Estimation will be consistent, but not efficient, provided that the choices satisfy the IIA (independence of irrelevant alternatives) as-

these are probably highly correlated.

[32]Health care quality would be hard to measure, however would probably be highest in major cities.

[33]Large cities, such as London, must in fact rely on the retailing and variety of arts and entertainment for their continued prosperity (The Economist, 1994).

[34]For a thorough discussion of different sampling techniques in multinomial logit models, one is referred to the book by Ben-Akiva, Moshe and Steven R. Lerman (1985).

sumption (Train(1986)).[35] Intuitively, the IIA assumption will in all likelihood hold since there are so many location choices and the restriction to four choices should capture most of the heterogeneity among choices. One may think, a priori, that there is some region specific variable which would change the implied probabilities if the region was removed, therefore the IIA test may be appropriate. Given that I have substantially controlled for differences, I doubt this will be a concern. For completeness, I have included the test procedure for IIA using the test of Hsiao and Small (1982).[36] The test is:

$$\frac{1}{1 - N_1/(\alpha N)} \left\{-2[\mathcal{L}_{\tilde{C}}(\hat{\beta}_C) - \mathcal{L}_{\tilde{C}}(\hat{\beta}_{\tilde{C}})]\right\} \rightsquigarrow \chi^2_{\tilde{K}}, \tag{2.6}$$

where N is the number of observations in the unrestricted choice set estimation, N_1 is the number of observations in the restricted choice set estimation, $\mathcal{L}_{\tilde{C}}(\hat{\beta}_C)$ is the log likelihood value from the unrestricted estimation, $\mathcal{L}_{\tilde{C}}(\hat{\beta}_{\tilde{C}})$ is the log likelihood value for the restricted regression, and $\alpha \geq 1$ is a scalar. This test statistic is distributed as a chi-square distribution with $\tilde{K}$ degrees of freedom making it simple to check for the null hypothesis that the IIA assumption holds.

2.4.3 General Conditional Logit Estimation Without Individual Specific Effects

In this section I estimate the effect of locational characteristics on the decision to migrate, treating all individuals as homogeneous. The first estimation procedure allows for the alternative choice of any other PUMA within the United States. The second estimation procedure is very similar in scope to the first, but it differs in

[35]One might be worried about tne actual random path chosen. However, the estimation results are robust to different random samples. I also ran the estimation assuming that people that moved to a certain state had only the option to choose among locations in that state. This might capture a more realistic view as to what choices these people actually had at their disposal. The IIA property states that for any two alternatives, the ratio of their choice probabilities is independent of the systematic utility of any other alternatives in the choice set.

[36]This test is a modification of the test of McFadden, Tye, and Train (1977). No violation of IIA was found. A more precise test developed by Ben-Akiva, McFadden, and Train could also have been used.

deciding what the relevant choices are. In the all choice estimation, the relevant choices are any of the 1018 PUMA located in the United States. One may believe that people really do have the choice among any of these, but in fact it may be more realistic to assume that people can only choose among the PUMAs within a certain state or, if they have moved to a major metropolitan area, that they could only have chosen among other metropolitan areas. In the state choice estimation, I consider the conditional logit estimation assuming that the person only had choices among alternative PUMAs in the state of final location.

The results for both the all choice estimation and the state choice estimation are presented in tables 2.2 and 2.3.[37] I will begin by analyzing the all choice estimates, which I feel are the more representative choice set. The coefficient on the real wage is positive and significant suggesting that people do choose locations with high average real wages.[38] This result is robust to other measures of the real wage as defined in the appendix. The coefficient on property taxes per capita (TAX2) indicates the result we would expect; higher property taxes discourage movement to a region. The income tax per capita (TAX1) has a negative coefficient and is opposite to what we would expect to find. There are ways to rationalize this result. It might be that movers are more knowledgeable of and sensitive to property taxes, but less so to income taxes. The lack of information might be greater for income taxes, since movers usually obtain the housing information before moving. Another explanation for the coefficient on income tax per capita is the endogeneity of the local income tax and the failure for my model to control for government services. Strumpf (1995) has noted that while virtually every locality has property taxes, very few have income taxes. His paper suggests that the income tax is endogenous, in the sense that only *good* local governments or areas can choose to use it. This is because these communities offer other attractive services or reduced property taxes that more than compensate for the increased income tax. These services may or may not be easy to measure. In

[37]All of the regression tables are presented at the end of this chapter.

[38]This might be due to the correlation between actual wage offered and the average real wage of a region.

this chapter I have not controlled for them and this could be causing the bias in the estimates. The coefficient on the unemployment rate (UNEMP1) has the expected sign. Unemployment strongly discourages migration to an area. This is consistent with the view that there will be less job opportunity in those regions. The coefficient on debt per capita (FUTTAX1) is surprising. It suggests that local governments with high debt per capita and thus likely candidates for higher future taxes, attract migrants. Given that the size of this effect is relatively small, it might not be such an alarming concern. Another explanation for this result is that my measure of FUTTAX1 is outdated, given that it is from 1981. The moving decisions for the people in the sample occurred between 1985-1990. It is possible that these debt per capita ratios changed between 1981 and 1985. The proxy I chose for future growth prospects (FUTGRO1) of a region has an unexpected negative effect on migration to a region. I used personal income growth per capita from 1980 to 1984 as this measure. My conclusion is that this is not a good measure of future growth prospects of a region. Other measures of FUTGRO listed in the appendix, such as the number of new private housing units authorized by permit between 1980 and 1986, had a positive and significant effect on immigration. The coefficient on net migration per capita between 1980 and 1986 (NETMIG) has the expected sign. This variable was used to control for the *domino effect*; if people are already pouring into a region for some exogenous reason, even a fad perhaps, this might induce herd behavior. The number of crimes per capita (CRIME3) has a negative effect on the decision to choose a region. The proxy for education (EDUC3) has a negative coefficient. This is not surprising given that this is a poor proxy for the quality of education in a region. Another reason may be that government expenditure on education might precisely be higher in areas where education is poor due to the redistributive role of government. The coefficients on weather are plausible. All else equal, people prefer to move to climates that are warmer and have less rainfall. The coefficient on percent of people below poverty level (SOCCOND4) has a negative sign. Thus, people tend to move away from economically poor areas. Health care seems to attract migrants to a region.[39] We would naturally expect health care to be positively related to the choice of a residence. The larger the percentage of blacks in a region, the less likely people will choose the location as a place of residence. Whether these results come from racism or other factors will be analyzed more closely in section 2.4.5. Pollution discourages the choice of a location quite strongly. This has policy implications for various communities. If growth is affected by immigration, then controlling pollution levels will be beneficial and it may even pay for the costs of cleanup in the long run. The central city indicator has a positive coefficient, indicating that there is a tendency to move towards metropolitan areas. The advantages of entertainment, opera, and jobs may all play a part in this metropolitan preference. Density measures also had a slight positive effect on immigration.

Due to possible sample bias, I performed the regressions on two random samples, labeled sample 1 and sample 2. The results were identical. To some degree these results describe the effect of locational characteristics on the decision to choose a region for the *typical* mover within the United States.

Individual specific traits of individual movers also provide some interesting results. I performed conditional logit estimations controlling for the marriage status of the migrant and for the number of children in the family of the migrant.[40] One finds, in addition to the effects already present above, that married people and families with children tend to be more averse to crime and pollution. The other location specific variables are not affected significantly.

The state choice sample results are similar except that the real wage coefficient is negative and the weather coefficient is larger in magnitude. These results seem counter intuitive. One explanation is that when people move within a state, they will choose areas of residence that are not necessarily where the highest wage is. This means that they may move to the suburb or some place near the area where they will work. That weather is more significant is more puzzling, however it may be that

[39]This effect was not found by others, such as Serow (1987), perhaps because they did not adequately control for other factors.

[40]These results are not included in the text.

some people move to every state. Thus, in the entire sample, there is a tendency to move to warmer climates but less so. When considering only within state choices, it seems that it is even more preferable to choose the location with the warmest climate. I do not think that the conditional logit estimation restricted to the state sample is the correct estimation method. However, I included the results for those interested. The more appropriate estimation procedure is using all of the choices; they essentially have different focuses.

2.4.4 The Determinants of Migration for the Elderly and the Spatial Life Cycle

This section attempts to understand the migration patterns of the elderly. The focus is interesting because it allows one to concentrate less on labor market considerations. The demographic changes due to longer life spans will have a major impact on our society in terms of social security insurance and savings. It may also affect the spatial structure in important ways. Specifically, if there are more older people and their locational preferences are quite distinct from the rest of the population, we should expect a predictably changing spatial structure. This will have effects on city and economic planning.

Table 2.4 presents the estimation results on the sample consisting of young and old migrants. The main results suggest what is intuitive. The real wage is not a concern for elderly migrants, while it is for young migrants. Taxes are not as important for elderly migrants as they are for younger migrants. Weather is significantly more important for the elderly. Pollution is also a more significant deterrent for the elderly's choice of a community. Finally, the percentage of old people already present in the region tends to attract elderly migrants.[41] These results suggest that with the changing demographic structure (i.e. there are increasingly more older people), we will observe a landscape that is more spread out in terms of population. Older people

[41] Other studies have used median age, which I find insignificant, and thus use a more precise measure which is the percentage of people over 55.

will continue to move to more fertile, less polluted, and warmer regions. Older people will aggregate together.

2.4.5 Whites and Blacks: Who's Running from Whom?

There has been published empirical work that documents a difference in the migration patterns between blacks and whites (Spilimbergo and Ubeda, 1994). The empirical work in this section focuses upon the locational characteristics that induce migration to a particular region across races. The analysis attempts to answer two questions: (i) Do different races care about different locational characteristics when choosing a location to reside in? (ii) Do races tend to flock together or do they have a taste for diversity?

The results are presented in tables 2.4 and 2.5. The differences among races seem to be small. I shall only briefly mention the characteristics that differ. For all groups, except Asians, the real wage seems to be an important consideration when choosing a location of residence. Blacks seems to be the most averse to the unemployment in a region, although all races are discouraged by it.[42] High crime in an area discourages all races to some degree, except Hispanics. Blacks and Hispanics especially are attracted to warmer climates.[43] The other location specific variables affect races similarly.

The more interesting results concern the racial characteristics of various regions. Whites are discouraged to move to areas with Asians and blacks, but almost four times as unlikely to move to an area where there are one percent more blacks than where there are one percent more Asians. There is no clear effect on the percentage of Hispanics in the region. Blacks move towards regions where there are a higher percentage of blacks and Hispanics, although there is a greater tendency for them to move to areas with a larger percentages of blacks. Blacks also tend to move away

[42]This slightly differs from previous studies which finds that blacks are not responsive to economic factors. The focus is different, this study does not ask, what will cause blacks to migrate, but given that they do migrate, they tend to choose places caring a great deal about economic factors.

[43]A 1 percent increase in the temperature of a region increases the probability of Hispanics moving there by 2.1 percent, blacks by 1.05 percent, whites by 0.57 percent, and for Asians weather is insignificant.

from Asians and whites.[44] Asians tend to move toward areas where there are other Asians, blacks, Hispanics, and whites.[45] Asians do not seem to avoid any race in their residential choice decision. This result may be due to my broad definition of Asians. Hispanics tend to move toward areas of high Hispanic concentration. The percentage of blacks has an insignificant effect on Hispanic immigration and the percentage of Asians and whites has a negative effect.

Some of these effects for racial clustering might be driven by income class. If one expects blacks to be poorer in general, then they will move to areas where other blacks are because of income reasons and not because of racial preference. One will observe from the summary statistics in appendix E for each race that this is not the case. For completeness, I performed regressions including dummies for income levels. The coefficients were statistically significant, but economically insignificant and there was still the racial bias.[46]

This chapter cannot distinguish whether it is the residents of a particular region that repel potential immigrants or whether it is a conscious choice of migrants to avoid that region. However, given freedom to choose residential location, the results must indicate some racial bias in choosing a residential location. All races tend to prefer to move to areas with higher concentrations of their own race. In addition, it seems that not only are whites avoiding blacks, but blacks are avoiding whites as well. Hispanics tend to avoid whites and Asians, but Asians do not seem to avoid any race. These results are in conflict with the recent notion that people prefer racial diversity.[47] There seems to be a natural tendency towards segregation among all races.

---

[44]Other estimations included the percentage of whites in a region and found the negative effect on the probability of moving to that region for blacks. These results are not presented in this chapter.

[45]This result may be due to the fact that my definition of Asians is very broad consisting of many types of people among the Asian race. The groups are Pacific Islander, Asian, Chinese, Taiwanese, Filipino, Japanese, Asian Indian, Korean, Vietnamese, Cambodian, Hmong, Laotian, Thai, Bangladeshi, Burmese, Indonesian, Malayan, Okinawan, Pakistani, Sri Lankan, Hawaiian, Samoan, Tahitian, Tongan, Guamanian, Northern Mariana Islander, Palauan, Fijian, and other Melanesian.

[46]The dummy variables in the table are written as INCRACB, INCRACA, and INCRACH.

[47]See Business Week (1994).

2.4.6 Trimming Trees and the Migration of the Highly Educated

The mobility of the highly skilled or highly educated is important for a variety of reasons, among which is local growth. If there is a place for public policy to alter the locational characteristics of any city so as to attract the highly educated individuals that will promote strong growth, then knowing the preferences of this group is of direct importance. In this section, I attempt to distinguish the mobility patterns of three groups of educated people; those with only high school education, those with bachelor's degrees, and those with masters or doctorate degrees. Most of the coefficients do not differ across groups and I will only discuss the ones that do differ (see table 2.8). The real wage is most significantly important to less educated people. This results might be due to the fact that less educated individuals' actual wages are more nearly correlated with the wage of a region. It might also be due to the fact that doctorates care less about materialistic rewards as opposed to other individuals.[48] More educated people are less likely to choose a region of residence with high unemployment. This might be due to the nature of their occupations.[49] More educated people seem to place a greater premium on climate than less educated people. Finally, more educated people are less likely to avoid areas with a higher concentration of black people. It appears that the racial bias is less severe for the more educated. This might indicate that racial bias decreases with education, but it exists nevertheless.

The data on the migration of different education classes seems to suggest that more educated people care more about non-economic characteristics than do other groups. The policy issues are less clear cut. It is difficult to distinguish any difference among education class where taxes or pollution are concerned. While pollution does discourage the immigration of the highly educated, it discourages the immigration of the less educated to a greater extent. These are probably the only variables that

[48]This might explain the doctorate result, but it is harder to believe the one for bachelors' candidates.

[49]Lower skilled jobs might be more volatile in terms of unemployment than high skilled jobs.

come under some local government influence.

2.4.7 Homos, Heteros, Natives, and Those Damn Foreigners

This section examines the impact of locational characteristics on the choice of residence for homosexuals and foreigners. The study of homosexual migration was pursued for no other reason than the *rare* availability of the data.[50] The study of foreign migration was pursued due to the claim of other authors, such as Spilimbergo and Ubeda (1994), that Europeans are more attached to family and move less for economic reasons. My sample is slightly biased, due to the fact that these foreigners have already moved to the United States and thus may actually not be *typical* foreign residents.

The results of table 2.9 suggest that homosexuals do not differ substantially in their migration patterns from heterosexuals.[51] The only significant differences are the homosexual negative reaction to the real wage, their stronger reaction to unemployment in a region, and their stronger preference for warmer climates. They seem to be less race biased, and most oddly, tend to move to areas with a higher percentage of poverty.

Foreigners seems to differ in two key respects from US citizens. The real wage has a negative aspect on their decision to choose a location and they tend to prefer warmer climates. This is a very crude indication that foreigners to the United States indeed care less about economic incentives in their locational choice.[52]

[50]The census data allows one to distinguish gay or lesbian couples. There is a questions that asks about non-married partner. By comparing the sexes it is possible to infer the sexual preference of the head of household.

[51]One can use any other of the tables to understand the migration of heterosexuals.

[52]A more in depth study would divide the foreigners into different countries and try to understand their migration pattern better.

2.5 Revealed Best Places to Live

Business Week and other magazines frequently publish "best places to live" articles. They are usually based on readers' polls that ask people what they value most in a place of residence. It is interesting to compare what people say with what they actually do and my research allows me to do that. The ranking of Business Week's (1994) latest poll is as follows (in descending order of importance): (i) low crime rate (ii) clean water/clean air (iii) plentiful doctors/many hospitals (iv) low income taxes/low property taxes (v) housing appreciation (vi) future job growth (vii) low risk of state tax rise (viii) inexpensive living (ix) recent job growth (x) near lakes or oceans (xi) low unemployment rate (xii) sunny weather (xiii) near national forests or parks (xiv) racially diverse population (xv) near a big city/symphony/near major league sports. Do these polls give a true indication of where are the best places to live?

Table 2.1: Comparisons of the Surveyed Best Places to Live and the Revealed Best Places to Live

Rank	Business Week Survey	Revealed Best Places
1	Low Crime	Sunny Weather
2	Clean Air and Water	Low Unemployment Rate
3	Plentiful Doctors/Many Hospitals	Clean Air
4	Low Income Taxes/Property Taxes	Income Taxes
5	Housing Appreciation	Prior Growth in Personal Income
6	Future Job Growth	Low Crime
7	Low Risk of State Tax Rise	High Real Wages
8	Inexpensive Living	Low Property Taxes
9	Recent Job Growth	Govt. Exp. on Education
10	Near Lakes or Oceans	Near Big City
11	Low Unemployment Rate	Lower Percentage Black
12	Sunny Weather	High Risk of Tax Rise
13	Near National Forests/Parks	Plentiful Doctors/Many Hospitals
14	Racially Diverse Population	Less Poverty
15	Near Big City	—

By observing table 2.2 for the estimation, ignoring individual specific effects, one will notice from the aggregate elasticities that the characteristics of the *revealed* best places to live are slightly different from those in the *Business Week* survey. See table

2.1 for a direct comparison.

The survey apparently overrates the importance of crime, health, and the low risk of state tax rise. The results actually show that although crime and health are important considerations, they are less important than other considerations such as pollution or the real wage. My study indicates that the low risk of a state tax rise actually is not a concern, and might in fact be a deterrent.[53]

The survey apparently underrates the importance of unemployment and weather. These are listed as last on the list of ranking of locational attributes, yet in my study they are among the most important considerations for movers in the decision of residential choice. In addition, racial diversity, contrary to the results of the survey, is not a preference among migrants, but actually a deterrent.

The survey is correct in its assessment of the importance of pollution relative to the other variables. The importance of forests and parks was not accessed in this chapter due to data limitations.

Overall it seems that the *revealed* best places to live are different than the ones produced in the survey where the relative importance of various attributes of the locations are concerned.

2.6 Suggestions for Further Research

The possibilities for further research in this area are large. It would be interesting to observe the choice between metro and non-metro regions and attempt to identify the individual or location specific characteristics that cause people to move towards Edge Cities.[54] Using information on worker type, one could identify the features that distinguish different classes of workers in the decision to move. The separation of blue collar workers and white collar workers might also allow one to control for people moving to certain areas just because their industry is located there. Using informa-

[53]The problems present in this measure were mentioned in section 2.4.3.

[54]For a discussion of Edge Cities, one is referred to Chincarini (1995) or Garreau (1991).

tion on the percentage of certain industries in the areas could capture this effect.[55] One might also be able to identify whether certain factors affect the migration of the highly skilled and lower skilled.[56] Any notable differences would have direct policy implications. Concerning the "best places to live", one of the items mentioned was racial diversity. My chapter suggests that racial diversity does not attract migrants. However, one could pursue a more in depth study by creating a type of Herfindahl index for racial diversity and observing the effect this variable has on migration decisions. It was also suggested that observing a certain class of movers' response to AFDC policies might be interesting.[57] One measure for quality of life variables that I did not use in this study is the amount of *greenery* in a location. By this, I mean the number of parks, trees, and wildlife. This might prove an interesting area to investigate.

2.7 Conclusion

The determinants of migration are important for policy reasons and for structuring economic geography models. This chapter has focused on understanding these determinants by empirically examining the migration behavior of people within the United States between 1985-1990. The major contribution to the literature is in using a newer technique with a very rich and geographical precise data set. By using specific geographical units and controlling for many factors, some of the ambiguous results of the literature have been corrected. The work presented here can be extended in many ways already mentioned. This chapter has briefly, but more correctly, investigated the fascinating behavior of individuals in choosing localities of residence.

[55]The County and City Databook has variables such as the percentage of manufacturing, retail services, banking, etc. in a region.

[56]This might also remove the problems with my measure of average real wage, since one might expect blue collar workers actual real wages to be more correlated to the average real wage.

[57]Professor Dora Costa of M.I.T. made this suggestion.

Table 2.2: Conditional Logit Model for the United States Sample

Regressors	Sample 1	Sample 1	Sample 2	Sample 2
	Clogit Coeff.	Aggregate Elasticity	Clogit Coeff.	Aggregate Elasticity
REALWAG1	1.94***	0.334	1.034**	0.141
	(0.402)		(.457)	
TAX1	0.002***	0.443	0.002***	0.374
	(.0003)		(.000)	
TAX2	−0.002***	-0.349	−0.002***	-0.319
	(.0003)		(.000)	
UNEMP1	−0.103***	-0.489	−0.089***	-0.360
	(.012)		(.013)	
FUTTAX1	0.0003***	0.160	0.0002***	0.136
	(.000)		(.000)	
FUTGRO1	−0.015***	-0.384	−0.013***	-0.263
	(.003)		(.003)	
NETMIG	—	—	3.292***	0.025
			(.552)	
CRIME3	−0.0001***	-0.276	−0.0002***	-0.353
	(.000)		(.000)	
EDUC3	−0.001***	-0.331	−0.002***	-0.335
	(.000)		(.000)	
ANN	0.018***	0.617	0.026***	0.740
	(.005)		(.005)	
RAIN	−0.001	0.023	−0.017***	-0.305
	(.002)		(.001)	
SOCCOND4	−0.015**	-0.136	−0.041***	-0.265
	(.007)		(.009)	
HEALTH1	0.001***	0.140	0.003***	0.243
	(.000)		(.000)	
RACEB	−0.025***	-0.168	−0.009***	-0.047
	(.003)		(.003)	
POLLUTE	−0.844***	-0.484	−1.48***	-0.049
	(.129)		(.209)	
CENCTY	0.621***	0.191	0.403***	0.095
	(.056)		(.063)	
$PseudoR^2$	0.123		0.177	
χ^2	1205.86		1517.31	
Obs.	18475		16571	
$\mathcal{L}(\hat{\beta}_C)$	-4282.02		-3540.64	

* *t* test significant at the 0.10 level. ** *t* test significant at the 0.05 level. *** *t* test significant at the 0.01 level.

Table 2.3: Conditional Logit Model for the State Choice Sample

Regressors	Sample 1	Sample 1
	Clogit Coeff.	Aggregate Elasticity
REALWAG1	−2.40***	-0.435
	(.713)	
TAX1	0.008***	1.995
	(.001)	
TAX2	−0.011***	-2.22
	(.001)	
UNEMP1	−0.079***	-0.378
	(.018)	
FUTTAX1	0.0003***	0.188
	(.000)	
FUTGRO1	0.011**	0.289
	(.005)	
NETMIG	-0.722	-0.005
	(.869)	
CRIME3	−0.0001***	-0.376
	(.000)	
EDUC3	−0.0003***	-0.086
	(.0001)	
ANN	0.148***	5.04
	(.011)	
RAIN	0.039	0.919
	(.006)	
SOCCOND4	−0.09***	-0.693
	(.011)	
HEALTH1	0.002***	0.204
	(.000)	
RACEB	−0.001	-0.007
	(.004)	
POLLUTE	−0.542***	-0.025
	(.141)	
CENCTY	0.505***	0.149
	(.072)	
$PseudoR^2$	0.158	
χ^2	2053.96	
Obs.	18202	
$\mathcal{L}(\hat{\beta}_C)$	-5490.44	

* t test significant at the 0.10 level. ** t test significant at the 0.05 level. *** t test significant at the 0.01 level.

Table 2.4: Conditional Logit Estimates for the Elderly and Young

Regressors	Elderly (> 55) Clogit Coeff.	Elderly (> 55) Aggregate Elasticity	Young (< 55) Clogit Coeff.	Young (< 55) Aggregate Elasticity
REALWAG1	0.904	0.124	3.36***	0.474
	(0.571)		(.441)	
TAX1	0.002***	0.346	0.002***	0.399
	(.000)		(.000)	
TAX2	−0.002***	-0.242	−0.002***	-0.338
	(.001)		(.000)	
UNEMP1	−0.068***	-0.271	−0.099***	-0.496
	(.016)		(.012)	
FUTTAX1	0.0001***	0.055	0.0002***	0.155
	(.000)		(.000)	
FUTGRO1	−0.018**	-0.361	−0.025***	-0.632
	(.004)		(.003)	
NETMIG	4.36***	0.039	3.85***	0.034
	(.767)		(.559)	
CRIME3	−0.0001***	-0.309	−0.0001***	-0.361
	(.000)		(.000)	
EDUC3	−0.002***	-0.547	−0.001***	-0.242
	(.001)		(.000)	
ANN	0.035***	0.989	0.002	0.017
	(.006)		(.005)	
RAIN	−0.021***	-0.377	−0.002	-0.045
	(.003)		(.002)	
SOCCOND4	−0.034***	-0.218	−0.013*	-0.116
	(.011)		(.008)	
HEALTH1	0.002***	-0.147	0.001***	0.120
	(.001)		(.000)	
RACEB	−0.009**	-0.048	−0.018***	-0.103
	(.004)		(.003)	
POLLUTE	−1.16***	-0.053	−0.753***	-0.042
	(.209)		(0.128)	
CENCTY	−0.104	0.124	0.557***	0.177
	(.081)		(.059)	
OLDPER	0.048***	0.448	0.0001***	0.252
	(.006)		(.000)	
$Pseudo R^2$	0.152		0.131	
χ^2	795.96		1282.23	
Obs.	9891		18475	
$\mathcal{L}(\hat{\beta}_C)$	-2225.69		-4245.84	

* t test significant at the 0.10 level. ** t test significant at the 0.05 level. *** t test significant at the 0.01 level.

Table 2.5: Conditional Logit Estimates for Whites and Blacks

Regressors	Whites Clogit Coeff.	Whites Aggregate Elasticity	Blacks Clogit Coeff.	Blacks Aggregate Elasticity
REALWAG1	2.66***	0.379	2.59**	0.361
	(.893)		(1.14)	
TAX1	0.001**	0.261	0.0004	0.092
	(.001)		(.0005)	
TAX2	−0.001	-0.166	−0.003***	-0.397
	(.001)		(.001)	
UNEMP1	−0.123***	-0.519	−0.211***	-0.875
	(.023)		(.029)	
FUTTAX1	−0.0001***	0.076	0.000	0.020
	(.000)		(.000)	
FUTGRO1	−0.016***	-0.340	−0.029***	-0.596
	(.005)		(.008)	
NETMIG	3.48***	0.031	3.24**	0.027
	(.973)		(1.33)	
CRIME3	−0.0001***	-0.338	−0.0001**	-0.183
	(.000)		(.000)	
EDUC3	−0.001**	-0.343	−0.003***	-0.766
	(.000)		(.001)	
ANN	0.02**	0.571	0.036***	1.05
	(.009)		(.012)	
RAIN	−0.022***	-0.419	-0.004	-0.077
	(.005)		(.005)	
SOCCOND4	−0.006	-0.040	−0.092***	-0.607
	(.016)		(.017)	
HEALTH1	0.002***	0.197	0.001*	0.099
	(.001)		(.001)	
RACEB	−0.035***	-0.222	0.060***	0.342
	(.008)		(.008)	
RACEA	−0.124**	-0.064	−0.226***	-0.118
	(.051)		(.049)	
RACEH	−0.011	-0.026	0.038***	0.100
	(.011)		(.013)	
POLLUTE	−0.678***	-0.032	−0.399**	-0.017
	(.221)		(.190)	
CENCTY	0.736***	0.187	0.326**	0.080
	(.111)		(.135)	
INCRACB	0.000	0.030	0.000	0.029
	(.001)		(.000)	
INCRACA	0.000***	0.042	0.000***	0.050
	(.000)		(.000)	
INCRACH	0.000***	0.032	0.000***	0.042
	(.000)		(.000)	
$PseudoR^2$	0.195		0.336	
χ^2	537.81		888.18	
Obs.	4961		5005	
$\mathcal{L}(\hat{\beta}_C)$	-1108.12		-876.61	

* t test significant at the 0.10 level. ** t test significant at the 0.05 level. *** t test significant at the 0.01 level.

Table 2.6: Conditional Logit Estimates for Asians and Hispanics

Regressors	Asians Clogit Coeff.	Asians Aggregate Elasticity	Hispanics Clogit Coeff.	Hispanics Aggregate Elasticity
REALWAG1	−1.62	-0.214	−5.23***	-0.756
	(1.45)		(1.54)	
TAX1	0.0002	0.043	0.000	0.073
	(.001)		(.001)	
TAX2	0.0005	0.079	-0.000	-0.015
	(.001)		(.001)	
UNEMP1	−0.215***	-0.856	−0.167***	-0.698
	(.036)		(.035)	
FUTTAX1	-0.0001	-0.059	-0.000	-0.048
	(.000)		(.000)	
FUTGRO1	−0.031***	-0.602	0.006	0.125
	(.009)		(.007)	
NETMIG	4.12***	0.033	−3.03*	-0.025
	(1.46)		(1.59)	
CRIME3	−0.0001**	-0.186	0.0002***	0.520
	(.000)		(.000)	
EDUC3	−0.003***	-0.620	0.002*	0.524
	(.001)		(.001)	
ANN	0.004	0.099	0.071***	2.08
	(.014)		(.014)	
RAIN	−0.027***	-0.479	−0.059***	-1.16
	(.006)		(.006)	
SOCCOND4	−0.084***	-0.536	−0.054*	-0.368
	(.024)		(.028)	
HEALTH1	0.002**	0.138	0.002***	0.195
	(.001)		(.001)	
RACEB	0.036***	0.201	-0.004	-0.021
	(.010)		(.013)	
RACEA	0.104**	0.047	−0.121**	-0.059
	(.046)		(.049)	
RACEH	0.024*	0.056	0.031***	0.075
	(.015)		(.011)	
POLLUTE	−1.59***	-0.070	−0.895***	-0.042
	(.322)		(.344)	
CENCTY	0.422**	0.094	0.077	0.020
	(.165)		(.207)	
INCRACB	0.000	0.016	0.000**	0.087
	(.000)		(.000)	
INCRACA	0.000	0.004	0.000**	0.026
	(.000)		(.000)	
INCRACH	0.000	0.030	0.000	0.007
	(.000)		(.000)	
$PseudoR^2$	0.472		0.641	
χ^2	1059.43		1854.49	
Obs.	5129		5247	
$\mathcal{L}(\hat{\beta}_C)$	-592.99		-519.48	

* t test significant at the 0.10 level. ** t test significant at the 0.05 level. *** t test significant at the 0.01 level.

Table 2.7: Conditional Logit Estimates for Different Education Levels

Regressors	H.S.	H.S.	B.A.	B.A.	Master/Ph.D.	Master/Ph.D.
	Clogit Coeff.	A.E.	Clogit Coeff.	A.E.	Clogit Coeff.	A.E.
REALWAG1	4.059***	0.695	−1.40**	-0.198	0.796	0.111
	(0.701)		(.581)		(.595)	
TAX1	0.002***	0.543	0.0007**	0.148	0.003***	0.530
	(.000)		(.000)		(.000)	
TAX2	−0.002***	-0.401	−0.0003	-0.052	−0.003***	-0.398
	(.000)		(.000)		(.001)	
UNEMP1	−0.081***	-0.382	−0.133***	-0.535	−0.172***	-0.701
	(.019)		(.017)		(.018)	
FUTTAX1	0.0002***	0.154	0.0002***	0.120	0.0001***	0.064
	(.000)		(.000)		(.000)	
FUTGRO1	−0.025***	-0.627	−0.017***	-0.348	−0.019***	-0.408
	(.005)		(.004)		(.004)	
NETMIG	4.23***	0.041	0.605	0.006	1.67***	0.015
	(.842)		(.710)		(.726)	
CRIME3	−0.0001***	-0.331	−0.0002***	-0.359	−0.0001***	-0.212
	(.000)		(.000)		(.000)	
EDUC3	−0.003***	-0.907	−0.0002	-0.048	−0.002***	-0.353
	(.001)		(.001)		(.001)	
ANN	0.000	0.002	0.044***	0.1.29	0.026***	0.753
	(.007)		(.006)		(.006)	
RAIN	−0.007**	-0.168	−0.009***	-0.163	−0.016***	-0.302
	(.003)		(.002)		(.003)	
SOCCOND4	−0.002	-0.017	−0.025**	-0.161	−0.022*	-0.143
	(.011)		(.011)		(.011)	
HEALTH1	0.001*	0.084	0.003***	0.208	0.003***	0.236
	(.000)		(.000)		(.000)	
RACEB	−0.023***	-0.155	−0.01**	-0.059	−0.007*	-0.043
	(.005)		(.004)		(.004)	
POLLUTE	−1.43***	-0.086	−0.394***	-0.018	−0.748***	-0.035
	(.313)		(.105)		(.126)	
CENCTY	0.587***	0.178	0.699***	0.182	0.236***	0.062
	(.088)		(.078)		(.080)	
$Pseudo R^2$	0.121		0.177		0.185	
χ^2	482.19		961.78		965.28	
Obs.	7820		9883		9846	
$\mathcal{L}(\hat{\beta}_C)$	-1749.89		-2240.31		-2121.44	

* t test significant at the 0.10 level. ** t test significant at the 0.05 level. *** t test significant at the 0.01 level.

Table 2.8: Conditional Logit Estimates for Homosexuals and Foreigners

Regressors	Gay Clogit Coeff.	Gay A.E.	Lesbian Clogit Coeff.	Lesbian A.E.	Foreign Born Clogit Coeff.	Foreign Born A.E.
REALWAG1	−6.70***	-0.967	−13.63***	-1.94	−6.27***	-0.817
	(1.40)		(.1.80)		(.798)	
TAX1	0.002***	0.468	0.002**	0.313	−0.0002	-0.033
	(.001)		(.001)		(.000)	
TAX2	0.0003	0.049	−0.005***	-0.816	0.003***	0.453
	(.001)		(.001)		(.001)	
UNEMP1	−0.249***	-1.05	−0.258***	-1.05	−0.144***	-0.554
	(.041)		(.048)		(.020)	
FUTTAX1	0.0001*	0.068	0.0001**	0.079	0.000	0.003
	(.000)		(.000)		(.000)	
FUTGRO1	−0.052***	-1.13	−0.029**	-0.621	−0.013**	-0.253
	(.011)		(.011)		(.005)	
NETMIG	1.25	0.015	-0.159	-0.002	−1.50*	-0.015
	(1.61)		(1.78)		(.815)	
CRIME3	−0.0001	-0.127	0.0001***	0.486	0.000**	0.147
	(.000)		(.000)		(.000)	
EDUC3	−0.006***	-1.51	−0.0000	0.003	−0.001*	-0.231
	(.001)		(.001)		(.001)	
ANN	0.096***	2.90	0.019	0.578	0.089***	2.43
	(.013)		(.014)		(.007)	
RAIN	−0.007	-0.135	0.005	0.101	−0.028***	-0.495
	(.005)		(.006)		(.003)	
SOCCOND4	0.127***	0.847	0.171	1.11	−0.025*	-0.147
	(.025)		(.029)		(.015)	
HEALTH1	−0.0006	-0.057	0.003***	0.245	0.003***	0.208
	(.001)		(.001)		(.000)	
RACEB	−0.009	-0.054	−0.073***	-0.448	0.006	0.032
	(.008)		(.009)		(.005)	
POLLUTE	−0.048	-0.002	0.075	0.003	−0.995***	-0.042
	(.176)		(.368)		(.137)	
CENCTY	0.444**	0.127	0.014	0.004	0.524***	0.131
	(.197)		(.225)		(.111)	
$PseudoR^2$	0.338		0.369		0.455	
χ^2	531.99		486.76		2533.07	
Obs.	2616		2252		11401	
$\mathcal{L}(\hat{\beta}_C)$	-522.19		-415.13		-1518.35	

* t test significant at the 0.10 level. ** t test significant at the 0.05 level. *** t test significant at the 0.01 level.

Chapter 3

Differentiating Between Volume and Return Information on Individual Stocks

CO-WRITTEN WITH GUILLERMO LLORENTE

3.1 Introduction

The accepted version of time varying risk premium and the increasing agreement about the non-linear relationships on the stock market, make it difficult to test for the general relationships among the variables in the market, such as volume and returns. The non-linearities present difficulties for testing all specifications of the market (LeBaron (1991)). Given these limitations, this chapter attempts to add more understanding to the issues without using any structural specification, and instead using what has been phrased *trading contrarian strategies.*

Many studies have found that *contrarian* strategies, that is strategies that sell winner stocks and buy loser stocks, can provide profits on average. There are two types of studies within this domain of research: those that consider short horizons (Lehmann (1991), Lo and MacKinley (1991), and Conrad *et al.* (1994)), and those that consider longer horizons ((DeBondt and Thaler (1985)). Most of the studies related to trading strategies concern themselves with the stock market *overreaction* hypothesis (Lehmann (1991), Lo and MacKinley (1991), and DeBondt and Thaler (1985)). The common empirical hypothesis of these studies is that overreaction implies that price changes of securities must be negatively autocorrelated for some holding period.

Within the price-volume literature, the theoretical models of Blume *et al.* (1994) and Campbell *et al.* (1993) have examined the relationship between trading volume and returns. Blume *et al.* (1994) present a model in which volume has informational content by its own and is used by the investors in their decisions. In Campbell *et al.* (1993), volume is a variable of interest because of its correlation with other variables, but in itself is unimportant: investors do not learn from volume nor use it in any decision making process. One of the main implications of Campbell et al. (1993) is that "price changes accompanied by high volume will tend to be reversed. This will be less true of price changes on days with low volume."

Conrad *et al.* (1994) use a particular specification of trading strategies that combine the volume and price information and conclude that the theoretical work of Campbell *et al.* (1993) is supported empirically. Although the results of Conrad *et al.* (1994) are extremely interesting, we find that there is room for further investigation.

Though Blume *et al.* (1994) do not specify a particular trading rule, Campbell *et al.* (1994) do specify a particular relationship between volume and returns through their autocovariances. In this chapter, we will take the results of price reversal for high volume stocks that Campbell *et al.* (1993) have noticed and the conclusions about informational content of volume that Blume *et al.* (1994) have brought forth to test whether or not volume does contain information *per se* (that is apart from its identification function) in addition to that present in returns. We will also test whether the price reversals predicted by Campbell *et al.* (1994) do indeed exist.

We will develop four kinds of trading strategies for investment in order to test for these hypotheses. The first strategy, what we call *volume lead*, is implemented creating portfolios with volume weights given by the deviation of individual security volume from market volume. The name, *volume lead*, comes from the fact that trading volume is the variable that signals the action to take given a change in price and is supposed to "identify", as Campbell *et al.* point out, the reason for the price change. The second strategy, similar to the "return weights" in Conrad *et al.* (1994) uses volume as an indicator variable but the weights are entirely based on returns. That is the stocks are classified in the same way as in the volume lead strategy, but the weights are different. The comparison of the results between these two strategies will give us important insights into the informational content of the trading volume. The third and fourth strategies combine both the return and volume information, keeping the same portfolio division. One can think in terms of overreaction or contrarian strategies when examining either strategy, however as we will see, of the four trading portfolios, only two of them, those related to high volume, can be considered to be contrarian, the other two are not. The reason for this classification is be able to make straightforward comparisons with the theory of Campbell *et al.* (1993). As a benchmark, we will also give the results obtained if a contrarian strategy for all portfolios is followed. The variable of study will be the profit of portfolios of securities formed following the dynamics imposed by each strategy.

One important point in all of the studies that deal with trading strategies is the impossibility of accounting for the risk of the position in each period. Because we focus only on the expected profits of these trading strategies and not on risk, this study, as others in the past, should be considered as a convenient tool for exploring the covariance properties of the variables of concern. Therefore, the results have theoretical implications only to the extent that they provide restrictions on economic models that must be consistent with the empirical results.

The first step is to specify some strategies for agents to follow based on a theoretical framework. This will allow us to avoid problems with the structural specification of the non-linear relation, though in fact the strategies are also non-linear. Second, to test the strategies we will use portfolios of individual securities, portfolios that can be constructed either by firm size, by volume of trade, or a combination of both. With this agenda we hope to shed some light, not only on the general relation between price and volume, but also to the differences and/or similarities when more specific characteristics about the stocks are considered.

One point we find important, but lacking in the work of Conrad *et al.*, is some measure of the aggregate portfolio. We think that the weights should consider the possible interrelations between returns and volume among securities due to the way agents choose and rebalance portfolios in the market. Thus, while the inclusion of an aggregate measure can cause cross-correlation effects and is thus difficult to interpret, we think that it can also add some valuable information.

The main results of the chapter are as follows. Trading volume has important informational content *per se* apart from its identification role and is different from the informational content of returns. The profits from the strategies that consider only volume as the weighting variable have the same qualitative but different quantitative results as those in which the weights are based solely on returns or a combination of returns and volume. Second, as Campbell *et al.* (1993) point out, we find price reversal for those stocks that experience a decline in price and have high trading volume with respect to the market. We cannot conclude anything about stocks that have a price increase and have higher than market volume. Finally, contrary to what one would expect, stocks with low volume also experience price reversal, independently of the direction of the change in price. These results are robust to the classification of stocks by size, particularly for smaller size stocks. From the results of following contrarian strategies, we can assert the dominance of this action to the one executed before. We can also state that by using volume and price information, the investors can significantly do better than using only volume and slightly better than using only return as is pointed out by Blume *et al.* (1994).

Though our work is not directly comparable to the Conrad *et al.* (1994) paper, where possible to make a comparison, we have found that our results are only different for those stocks which experience a decline in price and have a low trading volume with respect to the market. We think that this difference comes mainly from our definition of trading volume which is different that the one used by Conrad *et al.* (1994) and will be explained in the next section.

The rest of this chapter is organized as follows. Section 3.2 provides the basic setup of the hypothesis to be tested and a description of the previous research. Section 3.3 illustrates the trading strategies that are used in the analysis. Section 3.4 describes the data sources, the sample selection, and discusses the results of this chapter. Section 3.5 concludes the chapter.

3.2 Motivation for the Analysis

The theoretical literature that studies asset prices and trading volume can be divided into two branches: rational expectations asset pricing models and models based on differences of opinion. Rational expectation models are motivated by differences in information. They focus on the differences between privately informed traders, uninformed traders, and noise traders (Grossman and Stiglitz (1976, 1980), Wang (1992), He and Wang (1993), Pfleiderer (1984), Kyle (1985), and Blume *et al.* (1994)[1]). The literature concerning differences of opinion generally bases itself on different reactions from traders to stock announcements (Harris and Raviv (1993), Kandel and Pearson (1992), and Varian (1985, 1989)).

In recent articles, Blume *et al.* (1994) and Campbell *et al.* (1993) have examined the theoretical relationships between trading volume and returns. Blume et al. present an equilibrium model in which "volume provides information on information quality that cannot be deduced from the price statistics"; agents include the volume in their learning process, and the model suggests that there is a relationship between lagged volume and current returns on individual securities.

Campbell *et al.* explore the relationship between volume and returns by modelling the interactions between liquidity investors and risk-averse expected maximizers (that act as market makers). The market makers have to be compensated for satisfying the demands of liquidity traders in such a way, that an actual change in price, because of a selling pressure by liquidity traders, should be compensated by a change in the expected return. Volume information can help to distinguish between price movements

[1]In this model, the agents have myopic behavior, because of the relevation of the information each period.

which are due to the release of public information and those which reflect changes in expected returns. One of the main implications of their paper is that "... price changes accompanied by high volume will tend to be reversed; this will be less true of price changes on days with low volume...".

Following some of the theoretical articles on volume trading and returns, empirical research has attempted to verify the relationships between market volume and returns. Within this field, two directions have been taken. One set of studies has focused on the effects of new stock announcements, typically earnings announcements, on price and trading volume (Karpoff (1987), Morse (1980), Stickel and Verrechia (1992), and Kandel and Pearson 1993)). Another set of studies has focused on some structural specification, generally non-linear in nature (Campbell *et al.* (1993), LeBaron (1991), Gallan et al. (1992), Conrad *et al.* (1994), and Llorente (1995)). Most of these articles work with aggregate measures of price (return) and volume. In working with individual stocks, Llorente (1995) finds that volume (individual, as well as market) can add some information to the autocorrelations of individual stock returns in forecasting returns. The importance of volume is greater for small stocks. Conrad *et al.* (1994), using a variant of Lehmann's (1990) contrarian trading strategy, find evidence of a relationship between trading activity and subsequent autocovariances for individual weekly stock returns. "High-transaction securities experience price reversals, while low transaction securities have positive autocorrelated returns."

In this chapter we are going to test for the informational content of lagged volume and price as suggested by Blume *et al.* (1994) by developing some trading rules based on the results of Campbell *et al.* (1993). As a by-product of the design of the strategies, we will be able to test for the different behavior in prices corresponding to volume as postulated by Campbell et al. (1993).

We will use a variant of the Lehmann (1990) and Lo and MacKinlay (1990) contrarian portfolio strategies to measure the strength of the volume/return relationship. In a given period of time, t, we will classify the stocks into two portfolios according to their return in $t-k$ (positive or negative) and conditional on this we further divide them by volume (high or low with respect to the market) at time $t-k$ periods prior to the current period. Thus, we will have four portfolios. Depending on the classification, we will either short sell or buy the stocks in each portfolio by a dollar amount which will be specified in section 3.3. These weights will be determined by volume, return, or a combination of volume and return. For each portfolio, the weights are normalized so that they sum to one. This enables the formation of combined portfolios that are zero investment.

The method proposed has several important advantages over previous studies. To begin with, we measure profits and can provide statistics and economic information about the relationship between volume and return. Second, the profits from the first proposed strategy are directly related to the covariances between lagged volume and return, which are important in Blume *et al.* Third, the comparison of the profits from this strategy and the other two will allow us to address the information contained in the trading volume more accurately. Finally, these strategies are all examples of technical analysis.[2]

There are several important characteristics that differentiate our work from Conrad *et al.* (1994). First, our method to test for the importance of trading volume is completely different. Trading volume enters in all of their weight schemes as a modification of the information already transmitted by the returns. The profits from our *volume lead* strategy are directly related to the covariance between lagged volume and returns. This will avoid any influence of the bid-ask spread on the autocorrelations of returns. Second, using turnover as a variable for trading activity instead of the number of transactions or dollar-value volume we avoid differences across firms because of the different number of shares outstanding, prices, and any distortions due to block transactions. Moreover, as Cready and Mynatt (1991) point out, turnover can better reflect the informational content of the trade than the number of transactions. Lastly, we define high or low volume as the difference between individual turnover and market turnover. In this way, we consider individual stocks within the market and the idea of diversification when forming portfolios. This might produce

[2]Technical analysis is the study of volume and price information with the belief that it can provide information about underlying fundamentals not present in the price statistic alone.

cross effects among various stocks, which will add complication to the interpretation of the results, but we feel that it is important to consider the presence of the market.

3.3 Trading Strategies

3.3.1 A Volume Lead Strategy

The name for this strategy comes from the role that trading volume plays in it. Given a change in price, thus a high or low return stock, whether to buy or sell and how much to buy or sell will depend on the volume variable. The basic idea, following Campbell *et al.* (1993), is that price changes accompanied by high volume of a security in period $t-k$, indicate a high probability of price reversal in the near future, whereas, this is not true for price changes accompanied by low volume.

The difficult part in this set up is to define over what horizon will there be a reversal or when is the optimal timing between the pressure in the volume and the expected return. It could be argued that they are simultaneous, though the results of Campbell, *et al.* (1993), Conrad et al. (1994), and Llorente (1995) show that (as Stickel and Verrechia (1992) have argued) in a context of rational expectations "it takes at least one period for investors to compile and assimilate information about volume". We will assume that it takes at least one period for the market to incorporate information about volume, so that the change of direction in the change in price will occur at least one period later, and thus the strategy will be defined to follow this rule.[3]

The specification of the strategy is as follows: Consider a given set of N securities over T time periods. At the beginning of period t, divide the securities into two groups, those with a positive return in period $t-k$ and those with a negative return in period $t-k$.[4] Within the positive return or *winner* portfolio, short sell $u_{it}^{W,H}(k)$ dollars of the high volume securities and go long $u_{it}^{W,L}(k)$ dollars of low volume stocks.

[3]In this chapter, we will also study the behavior of the strategies for different lag periods.

[4]The value of k will be a constant that is determined *ex ante*, in our study k is equal to 1,2,4,and 26 weeks.

That is, short sell only those securities with a higher than market volume of trade in period $t-k$ and go long those securities with a lower trade volume than the market in period $t-k$. One must pursue the opposite (i.e. short sell low volume stocks and buy high volume stocks) for negative return securities. The intuition for these strategies comes from Campbell *et al.* (1993). They state that high volume securities will experience price reversal, but low volume securities will not. Therefore, as mentioned above, the trading strategies behave only in contrarian fashion for those securities that have a higher than market volume due to the expected negative autocorrelation. This will create four different portfolios of stocks. In each of these portfolios we will invest one dollar ($\sum u_{it}(k) = 1$).[5] The strategies are summarized below:

$$R_{it-k} > 0 \Rightarrow \begin{cases} V_{it-k} > V_{mt-k} \Rightarrow & u_{it}^{W,H}(k) = -\frac{(V_{it-k}-V_{mt-k})}{\sum_{i=1}^{N_{WH}}(V_{it-k}-V_{mt-k})} \\ \\ V_{it-k} \leq V_{mt-k} \Rightarrow & u_{it}^{W,L}(k) = \frac{(V_{it-k}-V_{mt-k})}{\sum_{i=1}^{N_{WL}}(V_{it-k}-V_{mt-k})}, \end{cases} \tag{3.1}$$

$$R_{it-k} \leq 0 \Rightarrow \begin{cases} V_{it-k} > V_{mt-k} \Rightarrow & u_{it}^{L,H}(k) = \frac{(V_{it-k}-V_{mt-k})}{\sum_{i=1}^{N_{LH}}(V_{it-k}-V_{mt-k})} \\ \\ V_{it-k} \leq V_{mt-k} \Rightarrow & u_{it}^{L,L}(k) = -\frac{(V_{it-k}-V_{mt-k})}{\sum_{i=1}^{N_{LL}}(V_{it-k}-V_{mt-k})}, \end{cases} \tag{3.2}$$

where R_{it-k} is the return for each security at time $t-k$, V_{it-k} is the volume traded of security i in period $t-k$ and V_{mt-k} is the average volume traded in all securities in period $t-k$; that is, $V_{mt-k} = \frac{1}{N}\sum_{i=1}^{N} V_{it-k}$. In this chapter we shall define volume as the ratio of the number of shares traded to the number of shares outstanding (referred to as turnover). This will make V_{it-k} bounded on the interval $[0,1]$. Consequently, u_{ipt} will be bounded on the interval $[-1,1]$.

These strategies give us a combination of four portfolios. Two of these portfolios will be sold short and two will be bought. The profits from the combined strategies at time t will be given by:

[5] It turns out, that in two portfolios, we will short sell one dollar and in two of them we will invest one dollar, making the overall strategy a zero-investment strategy. The reader should also realize that we can scale the one dollar by any positive amount.

$$\pi_t(k) = \sum_{i=1}^{N_{WH}} u_{it}^{W,H}(k) R_{it}^{W,H} + \sum_{i=1}^{N_{WL}} u_{it}^{W,L}(k) R_{it}^{W,L} + \sum_{i=1}^{N_{LH}} u_{it}^{L,H}(k) R_{it}^{L,H} + \sum_{i=1}^{N_{LL}} u_{it}^{L,L}(k) R_{it}^{L,L}. \tag{3.3}$$

We shall now analyze the profits for any generic portfolio. To begin with, we define N_p as the number of securities included in the portfolio of a certain strategy at any time t, where $N_p \subset N$. The weights in this generic portfolio are:

$$u_{it}(k) = \frac{(V_{it-k} - V_{mt-k})}{\sum_{i=1}^{N_p}(V_{it-k} - V_{mt-k})}, \tag{3.4}$$

where $V_{mt-k} = \frac{1}{N}\sum_{i=1}^{N} V_{it-k}$. The normalization factor assures that we will invest one dollar in the portfolio ($\sum_{i=1}^{N_p} u_{it}(k) = 1$). The profits in period t are:

$$\pi_t(k) = \sum_{i=1}^{N_p} \frac{(V_{it-k} - V_{mt-k})}{\sum_{i=1}^{N_p}(V_{it-k} - V_{mt-k})} R_{it}. \tag{3.5}$$

To make the notation simpler, define $\gamma_{it-k} = \frac{V_{it-k}}{\sum_{i=1}^{N_p}(V_{it-k} - V_{mt-k})}$ and $\gamma_{mt-k} = \frac{V_{mt-k}}{\sum_{i=1}^{N_p}(V_{it-k} - V_{mt-k})}$.

The mean profits over T periods of this k period ahead portfolio strategy will be:[6]

$$\bar{\pi}_t(k) = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \gamma_{mt-k})R_{it} \tag{3.6}$$

$$= -\frac{1}{T}\sum_{t=1}^{T} N_p(\gamma_{t-k}^p - \bar{\gamma}^p)(R_t^p - \bar{R}^p) + \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \bar{\gamma}^p)(R_{it} - \bar{R}^p) + \frac{1}{T}\sum_{t=1}^{T} R_t^p \sum_{i=1}^{N_p}(\gamma_{it-k} - \gamma_{mt-k}), \tag{3.7}$$

where $\gamma_{t-k}^p = \frac{1}{N_p}\sum_{i=1}^{N_p}\gamma_{it-k}$, $R_t^p = \frac{1}{N_p}\sum_{i=1}^{N_p} R_{it}$, $\bar{\gamma}^p = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{N_p}\sum_{i=1}^{N_p}\gamma_{it-k}$, and $\bar{R}^p = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{N_p}\sum_{i=1}^{N_p} R_{it}$.[7]

[6] The proof is included in the appendix for completeness.

[7] N_p depends on time the way our strategies have been defined, thus we cannot remove it from

Thus, average portfolio profits depend on the cross-covariances of the means of an equally weighted portfolio, the covariances between the volume and the return of the individual securities in the portfolio, and the arithmetic mean of the returns of the securities in the portfolio multiplied by the sum of all the weights. Therefore, the consideration of a market measure for trading volume will not produce direct important cross-sectional effects, besides those among securities in the same portfolio.

3.3.2 The Volume Lead Strategy with Return Only Based Weights

This strategy is very similar in spirit to the one above. The only modification is that the weights from above are scaled *only* by the actual returns in period $t-k$. Following the results of Lo and MacKinlay (1990), we have avoided using any measure of market return to prevent the cross-autocovariances among stocks to influence our conclusions. For the market volume, this is not a major problem as we explained above. This modification will allow us to access the information contained in our measure of volume. The strategies are summarized below:

$$R_{it-k} > 0 \Rightarrow \begin{cases} V_{it-k} > V_{mt-k} \Rightarrow & u_{it}^{W,H} = -\frac{R_{it-k}}{\sum_{i=1}^{N_{WH}} R_{it-k}} \\ \\ V_{it-k} \leq V_{mt-k} \Rightarrow & u_{it}^{W,L} = \frac{R_{it-k}}{\sum_{i=1}^{N_{WL}} R_{it-k}}, \end{cases} \tag{3.8}$$

$$R_{it-k} \leq 0 \Rightarrow \begin{cases} V_{it-k} > V_{mt-k} \Rightarrow & u_{it}^{L,H} = \frac{R_{it-k}}{\sum_{i=1}^{N_{LH}} R_{it-k}} \\ \\ V_{it-k} \leq V_{mt-k} \Rightarrow & u_{it}^{L,L} = -\frac{R_{it-k}}{\sum_{i=1}^{N_{LL}} R_{it-k}}, \end{cases} \tag{3.9}$$

where R_{it-k} is the return for each security at time $t-k$, V_{it-k} is the volume traded of security i in period $t-k$ and V_{mt-k} is the average volume traded in all securities in period $t-k$; that is, $V_{mt-k} = \frac{1}{N}\sum_{i=1}^{N} V_{it-k}$.

the summation over time.

3.3.3 The Volume Lead Strategy with Volume and Return Based Weights

Once the difference between the covariances of lagged volume and lagged return to the present returns is known independently, the next logical step is to evaluate these covariances considering all three variables together, that is lagged volume, lagged returns, and current returns. We will deal with this by proposing two alternative weighting schemes. The first weight, called VLR1, is a combination of lagged volume and lagged returns. The second weight, called VLR2, follows closely to the weights proposed by Conrad *et al.* (1994). In VLR2, volume only alters the autocovariance between returns and lagged returns. The main difference between VLR1 and VLR2 is the way in which the information compounded in volume enters in the autocorrelation in returns.

The strategy with more direct influence of volume information (VLR1) is given by:

$$R_{it-k} > 0 \Rightarrow \begin{cases} V_{it-k} > V_{mt-k} \Rightarrow & u_{it}^{W,H} = -\frac{(V_{it-k}-V_{mt-k})R_{it-k}}{\sum_{i=1}^{N_{WH}}(V_{it-k}-V_{mt-k})R_{it-k}} \\ V_{it-k} \leq V_{mt-k} \Rightarrow & u_{it}^{W,L} = \frac{(V_{it-k}-V_{mt-k})R_{it-k}}{\sum_{i=1}^{N_{WL}}(V_{it-k}-V_{mt-k})R_{it-k}}, \end{cases} \tag{3.10}$$

$$R_{it-k} \leq 0 \Rightarrow \begin{cases} V_{it-k} > V_{mt-k} \Rightarrow & u_{it}^{L,H} = \frac{(V_{it-k}-V_{mt-k})R_{it-k}}{\sum_{i=1}^{N_{LH}}(V_{it-k}-V_{mt-k})R_{it-k}} \\ V_{it-k} \leq V_{mt-k} \Rightarrow & u_{it}^{L,L} = -\frac{(V_{it-k}-V_{mt-k})R_{it-k}}{\sum_{i=1}^{N_{LL}}(V_{it-k}-V_{mt-k})R_{it-k}}, \end{cases} \tag{3.11}$$

where R_{it-k} is the return for each security at time $t-k$, V_{it-k} is the volume traded of security i in period $t-k$ and V_{mt-k} is the average volume traded in all securities in period $t-k$; that is, $V_{mt-k} = \frac{1}{N}\sum_{i=1}^{N} V_{it-k}$.

The strategy in which volume only modifies the autocovariance between lagged return and current return (VLR2) is given by:

$$R_{it-k} > 0 \Rightarrow \begin{cases} V_{it-k} > V_{mt-k} \Rightarrow & u_{it}^{W,H} = -\frac{(1+[V_{it-k}-V_{mt-k}])R_{it-k}}{\sum_{i=1}^{N_{WH}}(1+[V_{it-k}-V_{mt-k}])R_{it-k}} \\ V_{it-k} \leq V_{mt-k} \Rightarrow & u_{it}^{W,L} = \frac{(1+[V_{it-k}-V_{mt-k}])R_{it-k}}{\sum_{i=1}^{N_{WL}}(1+[V_{it-k}-V_{mt-k}])R_{it-k}}, \end{cases} \tag{3.12}$$

$$R_{it-k} \leq 0 \Rightarrow \begin{cases} V_{it-k} > V_{mt-k} \Rightarrow & u_{it}^{L,H} = \frac{(1+[V_{it-k}-V_{mt-k}])R_{it-k}}{\sum_{i=1}^{N_{LH}}(1+[V_{it-k}-V_{mt-k}])R_{it-k}} \\ V_{it-k} \leq V_{mt-k} \Rightarrow & u_{it}^{L,L} = -\frac{(1+[V_{it-k}-V_{mt-k}])R_{it-k}}{\sum_{i=1}^{N_{LL}}(1+[V_{it-k}-V_{mt-k}])R_{it-k}}, \end{cases} \tag{3.13}$$

where R_{it-k} is the return for each security at time $t-k$, V_{it-k} is the volume traded of security i in period $t-k$ and V_{mt-k} is the average volume traded in all securities in period $t-k$; that is, $V_{mt-k} = \frac{1}{N}\sum_{i=1}^{N} V_{it-k}$.

The second strategy follows more closely the spirit of the structural specification of Campbell *et al.* (1993). The autocovariance in returns is always the predominant factor in this strategy. Given the bounds of the volume measure discussed above, these weights only increase or decrease the weighting based on return information.

3.4 Empirical Analysis

To determine whether volume has informational content and whether it has different informational content than returns, we examine a sequence of trading rules. First, we examine profits from a portfolio strategy with weights based on volume information only. Next, we examine the differences in profits of these volume lead strategies with weights based solely on return information. Finally, we use both returns and volume information to calculate our weights.

3.4.1 The Data

The data used in this study came from the Center of Research in Security Prices (CRSP). We use the weekly (Wednesday to Wednesday) series of returns and volume of all individual stocks that have been continually listed on the CRSP from July 2, 1962 to December 31, 1992 and do not have more than twenty consecutive days of missing data or more than twenty consecutive days of no trading. There are 474 stocks that satisfy these requirements. In order to study the results by stock size, we divide the data into five quintiles determined by their market capitalization value in the middle of the sample period. Weekly sampling was chosen in order to avoid problems due to return and volume characteristics that typically differ by day of the week and to follow suit of other papers in the field. The measure used for market turnover is the one produced in Llorente (1995).[8] Market turnover is defined as the arithmetic mean of all the turnovers for all of the firms with valid records on CRSP on that considered day.

3.4.2 Volume Lead Results

This section evaluates the profitability of the portfolios described in the previous section. The weights were based on the previous volume variables, as defined above. The profits for the four basic portfolios (combinations of positive and negative returns and high and low volume) are returns to the invested dollar. The profits for the combined portfolios (positive return, negative return, high volume, and low volume, and total profits) are *real* profits because these portfolios are zero net investment (costless) portfolios.

Table G.1 (panel A) in appendix G presents the main results for the first strategy. The table reports the profits for one week lags. It considers the profits for all of the stocks in the sample and for different size stocks labeled as smallest, medium, and largest quintiles. The t statistics are presented in parenthesis.[9] All of these

[8]The reader is suggested to refer to that paper for a complete explanation of the methodology used to calculate this volume measure.

[9]The t statistics are computed using Newey-West corrected standard errors.

calculations ignore transaction costs, which will be considered later.

The results in table G.1 (panel A) sharply reject the price reversal hypothesis for those stocks with lower than market volume (independently of the return). The returns on these two portfolios are negative and significantly different from zero. Those stocks with higher than market volume and negative returns in period $t - k$ experience price reversal and for those with positive return we cannot reject the null hypothesis of *no* price reversal. When looking at the division of stocks by size, the results are quite similar, with striking persistence of stocks in the smallest quintile.

Thus, it seems that as Campbell *et al.* (1993) postulate, stocks with high volume and declines in price experience price reversals, but contrary to their results, stocks with low volume also experience price reversals. By observing the profits of any of the costless portfolios, they are positive for negative returns and for high volume. As we mentioned above, this result comes from the high volume negative return stocks. Even if some of the portfolios produce negative profits, the total combined portfolio produces a positive profit primarily derived from the low return, high volume stocks. Thus, the low return, high volume stocks profits are very high compared to other losing portfolios and outweigh the negative profits of those portfolios.

The analysis of the total profit by stock size reveals that these patterns are more important for smaller and medium size stocks than for larger stocks, as Blume *et al.* (1994) conjecture and Llorente (1995) demonstrates in another context.

Panel A of tables G.2, G.3, G.4, and G.5 produce the results for the strategies using different lag lengths ($k = 2, 4$, and 26). All of the calculations where performed with non-overlapping periods. The longer lagged periods are provided for interest, although one does not expect this to be testing for volume's informational content, since this is a relatively short-lived phenomenon. However, one can see that the results seem to hold even over a relatively longer time period. The persistence in the structure of profits for negative return stocks (regardless of high or low volume with respect to the market) shows up even in the lags of 26 weeks. Another striking result is that the total profit of the whole costless strategy is positive and significant for lags of two weeks and four weeks. These profits are higher than for one week lags. Analyzing the individual portfolios that induce this result, we observe that it comes again primarily from the stocks in the negative return and high volume. The return from this particular portfolio is higher than it was with a one week lag, and it offsets the portfolios that produce negative returns. The profits of the portfolios with positive returns ($R > 0$)in period $t - k$ are not significantly different from zero.

We observe the same patterns when concentrating only on stocks of a certain size, and continue to witness the strong persistence of smaller quintile stocks.

Tables G.6, G.7, and G.8 present other interesting statistics. Tables G.6 and G.7 present the mean investment per stock and period for the different portfolios. That is, how much of our one dollar investment we allocate to each stock within the particular portfolio. Table 8 presents the mean number of stocks per portfolio and over time.[10] Analyzing both tables together, we notice that a major proportion of our dollar investments are in stocks with higher than market volume, although evenly distributed among high and low return stocks. These results are artifacts of our data, which has more stocks with low or equal to market volume. We do not believe that the data will produce bias in our final results, because the investment per stock in a particular portfolio depends on the number of stocks in that portfolio (normalization factor) and on the deviation of its turnover from the market turnover.

A major conclusion from this purely volume lead strategy is that there exists price reversal for high volume stocks as postulated by Campbell *et al.* (1993) but there is inconclusive evidence about the "less probable" price reversal for low volume stocks. In fact, with our strategies and our sample, we find that there are price reversals with low volume stocks as well.

Another major conclusion is related to the central theme of this chapter: the informational content of trading volume. It is clear that for high volume stocks, the information present in volume is able to signal the reversal of prices and produce positive profits for the individual portfolios and the in the global portfolio.

[10]The distribution of the stocks among the portfolios for different lags are similar to the ones presented in table 8. This is an artifact of the non-overlapping construction of the strategies for different lags.

3.4.3 The Volume Lead Strategy with Return Only Based Weights

This section evaluates the profitability of the portfolios described in section 3.3. The main difference of this strategy from the volume lead strategy are the weights. In this strategy, as carefully documented in the previous section, the weights are based solely on the returns of the previous period, but still keeping the same classification of stocks in each portfolio. As before, the results represent returns to our dollar investments in the separate portfolios and profits to our combined portfolios.

This weight scheme has two important characteristics. First, contrary to the volume lead weights, the profits from the strategies followed here are based only on the autocorrelations of the individual returns. Volume is only an indicator variable; we assign it an "identification" function as in Campbell *et al.* (1993). Second, the comparisons of the results from these strategies with the volume lead strategies allow us to access the main hypothesis concerning the informational content of volume, separately from its "identification" role. As the reader will notice, the definitions of positive and negative returns and high and low volume for all securities are independent of the weight scheme. Thus, comparisons of profits among different weight schemes are perfectly valid.

Table G.1 (panel B) presents the main results from this strategy. The table reports profits for one week lags for different portfolios as before. Panel B of tables G.2, G.3, G.4 and G.5 present the same information as Table G.1 (panel B) but for different lag lengths.

The major results about the price reversal hypothesis found in the former strategies apply here as well. Low volume stocks experience price reversal, contrary to what the theory would predict, and high volume stocks, as theory predicts, experiences price reversals as well. It is interesting to note that the profits for the positive return and high volume stocks are significant for all stocks and for the smallest stocks for a lag length of one week, whereas in the volume lead case they were not. One could argue that because the strategy followed in this portfolio is contrarian in returns, the results really reflect the negative autocorrelation produced by the bid-ask for the smaller sized stocks. Even if this argument is credible, we do not think that the bid-ask spread is so important. First, the results of Lo and MacKinley (1991) about the bid-ask spread suggest that this spread does not have any major influence on the final results.[11] Second, with that line of reasoning, the strategy would only produce significant profits for the smallest stocks and not for the others in the case of negative returns and low volume, unless the information produced by the "identification" function assigned to volume is so important as to offset that fact. In any case, the importance of volume information for the high volume stocks is present.

The informational content contained in the volume variable, besides that of "identification", is clear when comparing the results in this section with those of the volume lead section. The profits for the individual portfolios have the same qualitative patterns and very similar quantitative results, independently of the stocks size or lag. The major differences are related to the magnitude of the profits for the individual portfolios, when positive they are greater with the weight scheme used in this section, and when negative, they are also more negative. In other words, this strategy produces profits that are of greater magnitude than in the volume lead strategy case. This also causes the profits for the global portfolios to be smaller with this weight scheme than with the weight scheme based on the volume lead strategy.

3.4.4 The Volume Lead Strategy with Volume and Return Based Weights

This section evaluates the profitability of the other portfolios described in section 3.3, specifically VLR1 and VLR2. The main difference of these strategies from the two mentioned above are again in the weight calculations. In these strategies, the weights are based on a combination of returns and volume of the previous period. Again the results represent returns to our dollar investments in the four basic portfolios and profits to our combined portfolios. Panel C and D of tables G.1, G.2, G.3, G.4, and

[11]The reader should realize the strong similarity between our sample and their sample.

G.5 present the main results for these weight schemes for different lags and stock sizes. The comparison of the results from these weighting schemes to the other two weighting schemes used above make it possible to address simultaneously the joint informational content of both volume and returns and their individual informational content.

The major results about the price reversal hypothesis found before in the other strategies apply here as well. High volume stocks experience reversal as do low volume stocks.

The most important result concerns the comparisons of profits with the other weight schemes. Observing panels A, B, C, and D of the mentioned tables we see that the general pattern structure is similar for all of them, independently of the lag period or stock size. The only two major differences worth noting is the change in high volume stocks with positive returns for lag length one. For this weighting scheme we find clear evidence of price reversal and positive profits. The other important difference is the profit for the global portfolio. For the first weighting scheme (the volume lead strategy), global profits are mainly positive and when significant, produce the highest profit among the four strategies. It seems that information impounded in the volume weight scheme is able to produce global profits even when the other two weighting schemes do not produce profits.

Comparing the profits from the four weighting schemes that include returns, we observe how the joint consideration of volume and returns matters. If volume only modifies the return information (see strategy VLR2, panel D), it does not matter too much. That is, volume does not add much information to that already presented in returns (see Panel B). Contrarily, if volume has its own presence in the covariance structure (strategy VLR1, panel C), it matters a lot and the profits are greater than those with returns alone (see panel B), as postulated by Blume *et al.* (1994).

3.4.5 Other Issues

In this section we would like to briefly address several issues that have already been mentioned in the exposition but not adequately explained. The issues are the following: the possible influence of the bid-ask spread, the transaction costs, the connection of the results with those in Campbell *et al.* (1993) and Llorente (1995), and the results of the strategies when they are contrarian based.

Though the bid-ask spread can be responsible for some of the results using the last two strategies we have presented, we, following the arguments of Lo and Mackinley (1990) and Lehmann (1990), do not think it is a major problem in our study. The results for the first strategy and their comparison with those for the two others support these arguments.

As in all the studies of this nature, our profits are calculated without considering the transaction costs due to the rebalancing of portfolios every period. In order to avoid this problem, and following Lehmann (1990), we calculated the same profits subtracting from the portfolios five percent of the absolute difference between weights in two consecutive periods ($\pi_{it}(k) - 0.05|u_{it}(k) - u_{it-1}(k)|$). All of these profits for every portfolio were negative. This result has been documented in other papers as well. Few of the returns to the portfolio strategies in this chapter exceed five percent, those that do tend to be for longer horizons and for smaller stock sizes. Therefore these profits, while being statistically and economically significant, may not be true profit measures for an individual investor.

There is an issue concerning the relationship between our results with those of Campbell *et al.* (1993) and Llorente (1995). Though these other authors' scenarios are not directly comparable to ours, because of the specification and risk problems mentioned above, the profits for our second, third, and particularly fourth strategies are directly related to the first autocorrelation of individual returns. For those stocks with positive return, we find evidence of the results from the above mentioned studies: the first autocorrelation is lower on high volume days than on low volume days, and its significance varies with the stock size.

The last issue concerns the comparisons of the profits of our strategies to those obtained with the same weighting schemes but behaving contrarian based on returns in period $t-k$. This essentially removes the identification function assigned to volume, but still considers the information that volume can carry. From this comparison we are able to to analyze the volume information from a different perspective and to compare our results with some other studies in the literature. Table G.9 presents the results for different weight schemes used by several authors. Though the results are qualitatively similar, one will notice that the actual weight scheme used has a great deal of influence on the quantitative results.

When comparing the profits of the pure contrarian for all strategies to those in which contrarian, is only executed in the high volume situation (see the last two columns of tables G.1 to G.5), we find what could have been inferred from the previous results: the price reversal pattern found in all volume situations (high or low) produces higher profits for the contrarian based weights. Among the different weight schemes, the higher profits come from considering only returns or returns and volume combined (strategy VLR1). Once again, we see how the information in trading volume can improve that already contained in returns.

3.5 Conclusion

In this chapter we have documented the importance of the information carried by trading volume and how different it is from that carried in the return variable, as conjectured by Blume *et al.* (1994). In order to test for this hypothesis and to avoid any structural specification a priori, we address the problem by implementing four trading strategies in the spirit of the contrarian trading strategy literature, and following the results of Campbell *et al.* (1993). As a by-product of the design in the strategies, we are able to test for the "identification" function of the volume variable as postulated in Campbell *et al.* (1993). With this specification we lose control of the risk of our portfolios at every period of time, but gain flexibility on the functional form.

The main results of the chapter are as follows. Trading volume has information different from that carried by returns. The portfolio profits from the strategies based on volume or returns alone have the same qualitative and but different quantitative patterns, which are in turn different to the patterns from a strategy consisting of both return and volume information. These results are independent of stock size and lag length in the information transition.

The second major result of the chapter is that the price of high volume stocks tend to reverse as well as those for low volume stocks. Once again, the results are robust to stock size and lag length.

Appendix A

Chapter 1: Mathematical Derivations

A.1 Deriving Zero Order and First Order Equations

One may use the perturbation theory to understand the motion of a non-linear system near some equilibrium by looking at a small perturbation of the equilibrium. Taking equations 1.2-1.7, one can express each variable in terms of an equilibrium value and a perturbed part, which is assumed to be small compared to the equilibrium, otherwise the approximations do not hold.[1] Thus:

$$Y(\theta,t) = Y_0(\theta) + Y_1(\theta,t), \tag{A.1}$$
$$T(\theta,t) = T_0(\theta) + T_1(\theta,t), \tag{A.2}$$
$$w(\theta,t) = w_0(\theta) + w_1(\theta,t), \tag{A.3}$$
$$\omega(\theta,t) = \omega_0(\theta) + \omega_1(\theta,t), \tag{A.4}$$
$$\bar{\omega}(\theta,t) = \bar{\omega}_0(\theta) + \bar{\omega}_1(\theta,t), \tag{A.5}$$
$$\frac{\partial\lambda(\theta,t)}{\partial t} = \frac{\partial\lambda_0(\theta)}{\partial t} + \frac{\partial\lambda_1(\theta,t)}{\partial t}. \tag{A.6}$$

[1] Higher order terms are neglected. This restricts the analysis to examining the initial (linear) evolution of the system away from equilibrium.

One then simply places these expressions into the equations and multiplies through, collecting zero and first order terms. The zero order term expressions are quite straightforward, however in calculating the first order terms, one will find it convenient to use the binomial expansion at a crucial juncture shown below.

The expression for $T(\theta)$ and the expression for $w(\theta)$ are expanded in the following convenient way:[2]

$$T(\theta,t) = [T_0(\theta) + T_1(\theta,t)]^{1-\sigma} \tag{A.7}$$
$$= T_0(\theta)^{1-\sigma} + (1-\sigma)T_0(\theta)^{-\sigma}T_1(\theta,t) + ... \tag{A.8}$$
$$\approx T_0(\theta)^{1-\sigma} + (1-\sigma)T_0(\theta)^{-\sigma}T_1(\theta,t). \tag{A.9}$$

Thus, higher order terms are neglected and one obtains a manageable expression for zero and first order terms of $T(\theta)$. The same steps may be applied in order to derive the zero and first order terms of $w(\theta)$.

A.2 Normalizing the Equilibrium

Given our equilibrium equations 1.8-1.13 we now normalize $w_0 = 1$. From this it follows that $Y_0 = \frac{1}{2\pi r}$, so we simply re-define $Y_0^* = \alpha Y_0$. From this it follows that $T_0 = \int_0^{2\pi} \frac{1}{2\pi r} e^{-2z\sin^2(\frac{\phi-\theta}{2})} r d\phi = \Gamma_0(\tau r(\sigma-1))$. Thus, we re-define $T_0^* = \beta T_0$. From this, we have equation (10) which with the normalized values for Y_0 and T_0 implies that $w_0 = \int_0^{2\pi} e^{-2z\sin^2(\frac{\phi-\theta}{2})} r d\phi = 2\pi r\Gamma_0(\tau r(\sigma-1))$. Thus, we re-define $w_0^* = \delta w_0$. Which also makes $\omega_0 = 1$. This completes the normalization of the equilibrium which is observed in equations 1.20-1.25.

[2] As a reminder, the expansion is of the form: $[a+b]^n = a^n + \binom{n}{1} a^{n-1}b + \binom{n}{2} a^{n-2}b^2 + ...$

A.3 Deriving The Constants of the Fourier Series Expansion

The perturbation is of the form:

$$\lambda_1(\theta,t) = \sum_{m=-\infty}^{m=\infty} c_m(t)e^{im\theta}, \tag{A.10}$$

$$\lambda_1(\theta,t)e^{-in\theta} = \sum_{m=-\infty}^{m=\infty} c_m(t)e^{i(m-n)\theta}, \tag{A.11}$$

$$\frac{1}{2\pi}\int_0^{2\pi} \lambda_1(\theta,t)e^{-in\theta}d\theta = \frac{1}{2\pi}\sum_{m=-\infty}^{m=\infty} c_m(t)\underbrace{\int_0^{2\pi} e^{i(m-n)\theta}}_{2\pi\delta_{m,n}}, \tag{A.12}$$

$$\frac{1}{2\pi}\int_0^{2\pi} \lambda_1(\theta,t)e^{-in\theta}d\theta = c_n(t), \tag{A.13}$$

where $\delta_{m,n} = 1$ for $m = n$ and $\delta_{m,n} = 0$ for $m \neq n$. Using the fact that $\cos(m\theta) = \frac{e^{im\theta}+e^{-im\theta}}{2}$ and $\sin(m\theta) = \frac{e^{im\theta}-e^{-im\theta}}{2i}$ one can derive the other constants using a similar logic.

A.4 Derivation of K(z)

$$\int_0^{2\pi} e^{p\cos\phi+q\sin\phi}\cos(m\phi)d\phi \tag{A.14}$$

$$= \frac{\pi}{(p^2+q^2)^{m/2}}\left[(p^2+2ipq-q^2)^{m/2}I_m\left(\sqrt{p^2+q^2}\right)\right] + \frac{\pi}{(p^2+q^2)^{m/2}}\left[(p^2-2ipq-q^2)^{m/2}I_m\left(\sqrt{p^2+q^2}\right)\right], \tag{A.15}$$

$$\int_0^{2\pi} e^{p\cos\phi+q\sin\phi}\sin(m\phi)d\phi \tag{A.16}$$

$$= \frac{i\pi}{(p^2+q^2)^{m/2}}\left[(p^2+2ipq-q^2)^{m/2}I_m\left(\sqrt{p^2+q^2}\right)\right] -$$

$$\frac{i\pi}{(p^2+q^2)^{m/2}}\left[(p^2-2ipq-q^2)^{m/2}I_m\left(\sqrt{p^2+q^2}\right)\right], \tag{A.17}$$

where $p^2+q^2>0$, $m=0,1,2,...$, $I_m(z)=(z/2)^m\sum_{k=0}^{\infty}\frac{(z^2/4)^k}{k!(m+k)!}$. One may also recognize that $I_m(z)=i^{-m}J_m(iz)$.[3]

Let R be given by:

$$\begin{aligned}
R &= \int_0^{2\pi} e^{p\cos\phi+q\sin\phi}e^{im\phi}d\phi && \text{(A.18)}\\
&= \int_0^{2\pi} e^{p\cos\phi+q\sin\phi}\left[\cos(m\phi)+i\sin(m\phi)\right]d\phi && \text{(A.19)}\\
&= \int_0^{2\pi} e^{p\cos\phi+q\sin\phi}\cos(m\phi)d\phi+i\int_0^{2\pi} e^{p\cos\phi+q\sin\phi}\sin(m\phi)d\phi && \text{(A.20)}\\
&= \frac{2\pi}{(p^2+q^2)^{m/2}}\left[(p^2+2ipq-q^2)^{m/2}I_m\left(\sqrt{p^2+q^2}\right)\right]. && \text{(A.21)}
\end{aligned}$$

Now, I am going to let $p=z\cos(\theta)$ and $q=z\sin(\theta)$. Thus, R is given by:

$$\begin{aligned}
R &= \int_0^{2\pi} e^{z\cos(\theta)\cos\phi+z\sin(\theta)\sin\phi}e^{im\phi}d\phi && \text{(A.22)}\\
&= \int_0^{2\pi} e^{z\cos(\phi-\theta)}e^{im\phi}d\phi, && \text{(A.23)}
\end{aligned}$$

which comes from $\cos(\phi-\theta)=\cos(\theta)\cos(\phi)+\sin(\theta)\sin(\phi)$. Also, for later use, it is helpful to note that $p^2+q^2=z^2$, $p^2-q^2=z^2\cos(2\theta)$, $2ipq=iz^2\sin(2\theta)$, and consequently $p^2+2ipq-q^2=z^2e^{i2\theta}$.

So using the formula from above, one has an expression for R:

$$R=2\pi e^{im\theta}I_m(z). \tag{A.24}$$

But, we have that:

$$\begin{aligned}
K(z) &= \int_0^{2\pi} e^{-2z\sin^2\left(\frac{\phi-\theta}{2}\right)}e^{im\phi}d\phi && \text{(A.25)}\\
&= \int_0^{2\pi} e^{-z(1-\cos(\phi-\theta))}e^{im\phi}d\phi && \text{(A.26)}
\end{aligned}$$

[3] $I_m(z)$ is a modified Bessel function of the second kind or more commonly known as the imaginary Bessel function.

$$= e^{-z} \underbrace{\int_0^{2\pi} e^{z\cos(\phi-\theta)} e^{im\phi} d\phi}_{R} \quad (A.27)$$

$$= 2\pi e^{-z} I_m(z) e^{im\theta} \quad (A.28)$$

$$= 2\pi \Gamma_m(z) e^{im\theta}. \quad (A.29)$$

where $z = \tau(\sigma - 1)r$ and $\Gamma_m(z) = e^{-z} I_m(z)$. In getting from equation A.25 to equation A.26, one uses the following steps:

$$\cos(2x) = \cos^2(x) - \sin^2(x) \quad (A.30)$$

$$= 1 - \sin^2(x) - \sin^2(x) \quad (A.31)$$

$$= 1 - 2\sin^2(x) \quad (A.32)$$

$$\Rightarrow 2\sin^2(x) = 1 - \cos(2x), \quad (A.33)$$

where in this instance $x = \frac{\phi - \theta}{2}$.

A.5 Proof that $\bar{\omega}_1 = 0$

$$\bar{\omega}_1(t) = \int_0^{2\pi} \left[\lambda_1(\theta, t) + \frac{\omega_1(\theta, t)}{2\pi r} \right] r d\theta \quad (A.34)$$

$$= \int_0^{2\pi} \rho e^{im\theta} r d\theta + \int_0^{2\pi} \frac{a_\omega \rho}{2\pi r} e^{im\theta} r d\theta \quad (A.35)$$

$$= r\rho \underbrace{\int_0^{2\pi} \cos(m\theta) + i\sin(m\theta) d\theta}_{=0} + \frac{a_\omega \rho}{2\pi} \underbrace{\int_0^{2\pi} \cos(m\theta) + i\sin(m\theta) d\theta}_{=0} \quad (A.36)$$

$$= 0. \quad (A.37)$$

This follows immediately from:[4]

$$e^{im\theta} = \int_0^{2\pi} \cos(m\theta) + i\sin(m\theta) r d\theta \quad (A.38)$$

$$= \int_0^{2\pi} \cos(m\theta) r d\theta + i \int_0^{2\pi} \sin(m\theta) r d\theta \quad (A.39)$$

[4]This proof is only valid for $m \neq 0$, which is the case for this entire chapter. In other words, a perturbation with $m = 0$ would increase the total population of workers which has no intuitive appeal.

$$= r\left([\sin(m\theta)]_0^{2\pi} + i\left[-\cos(m\theta)\right]_0^{2\pi}\right) \quad (A.40)$$

$$= 0. \quad (A.41)$$

A.6 Explicitly Dealing with Labor Mobility and Growth

From equation 1.48 we have the following relation:

$$\frac{\partial \lambda_1(\theta,t)}{\partial t} = \left(\frac{a_\omega \gamma}{2\pi r}\right)\lambda_1(\theta,t) \Rightarrow \quad (A.42)$$

$$\int \frac{1}{\lambda_1(\theta,t)} d\lambda_1 = \int \left(\frac{a_\omega \gamma}{2\pi r}\right) dt \Rightarrow \quad (A.43)$$

$$\ln \lambda_1(\theta,t) = \left(\frac{a_\omega \gamma}{2\pi r}\right) t + c \Rightarrow \quad (A.44)$$

$$\lambda_1(\theta,t) = e^{\left(\frac{a_\omega \gamma}{2\pi r}\right)t} e^c \quad (A.45)$$

$$= \lambda_1(0) e^{\left(\frac{a_\omega \gamma}{2\pi r}\right)t}, \quad (A.46)$$

where $\lambda_1(0)$ refers to the original perturbed state.

Thus, the growth rate of certain areas also depends on the sensitivity of labor mobility to the real wage differential, which is given by γ. So if γ is quite low, then it will take longer for this "preferred wavelength" to dominate the spatial pattern. And thus, one will observe many different density economic structures for a very long time. To take some illustrative examples, if $\gamma = 0$ then there would be no wavelengths growing. Another example: the time it would take for a region to double in density would be $\frac{\ln(2)}{a_\omega \gamma}$.

A.7 More about the Distance Factor

Observing figure B-1, one will notice that a proxy for the actual distance between locations x and z is the line drawn which connects them. If one splits this line into two parts, one is essentially dividing the angle between x and z. The length of half of the line connecting x and z is simply $r \sin\left(\frac{\phi-\theta}{2}\right)$. So the entire distance is two times this. This would be a measuring device that would be proportional to the distance between x and z. The problem with this measure is that when ϕ is less than θ, the value of the distance, although the same in magnitude, has an opposite sign. In order to obtain a distance measure that has the same sign, we simply square our sine term. Thus, we have the distance factor. Finally, we want to indicate that the further a location is, the more it costs to ship goods to that place. Thus, the exponential discounting is added. The τ is added as actual transport costs, so that later, transport costs can be varied.

A.8 Explicitly Demonstrating that Amplitudes Are Irrelevant

Because we know that the different frequencies do not interact with each other, we can express the growth of each frequency or wavelength individually.

$$\lambda_1(\theta, t) = c_m(t) e^{g_m t} e^{im\theta}. \tag{A.47}$$

$$\begin{aligned} \frac{\partial \lambda_1(\theta, t)}{\partial t} &= c_m g_m e^{g_m t} e^{im\theta} && \text{(A.48)} \\ &= \gamma a_\omega \lambda_0(\theta) \lambda_1(\theta, t) && \text{(A.49)} \\ &= \gamma a_\omega \lambda_0 c_m e^{g_m t} e^{im\theta}, && \text{(A.50)} \\ & && \text{(A.51)} \end{aligned}$$

which implies that $g_m = \gamma \lambda_0 a_\omega$. The amplitudes have no effect on the growth rate. This result comes from the fact that different frequencies do not interact with each other. Intuitively, even if one frequency has a greater amplitude, the fastest growing wavelength will dominate and overtake this wavelength.

Appendix B

Chapter 1: Figures for the Geographical Landscape

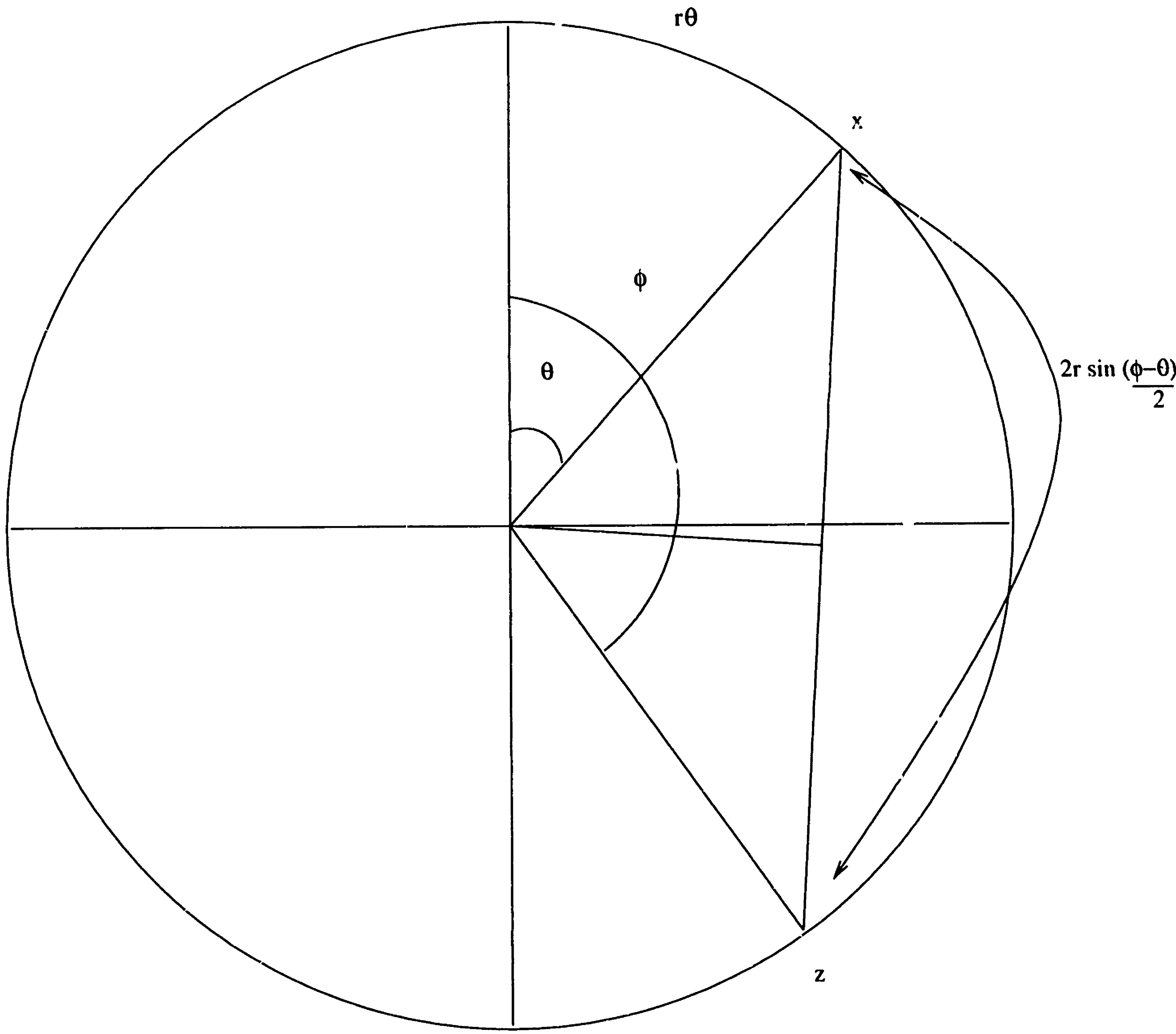

Figure B-1: The Geographical Economy Located on a Circle

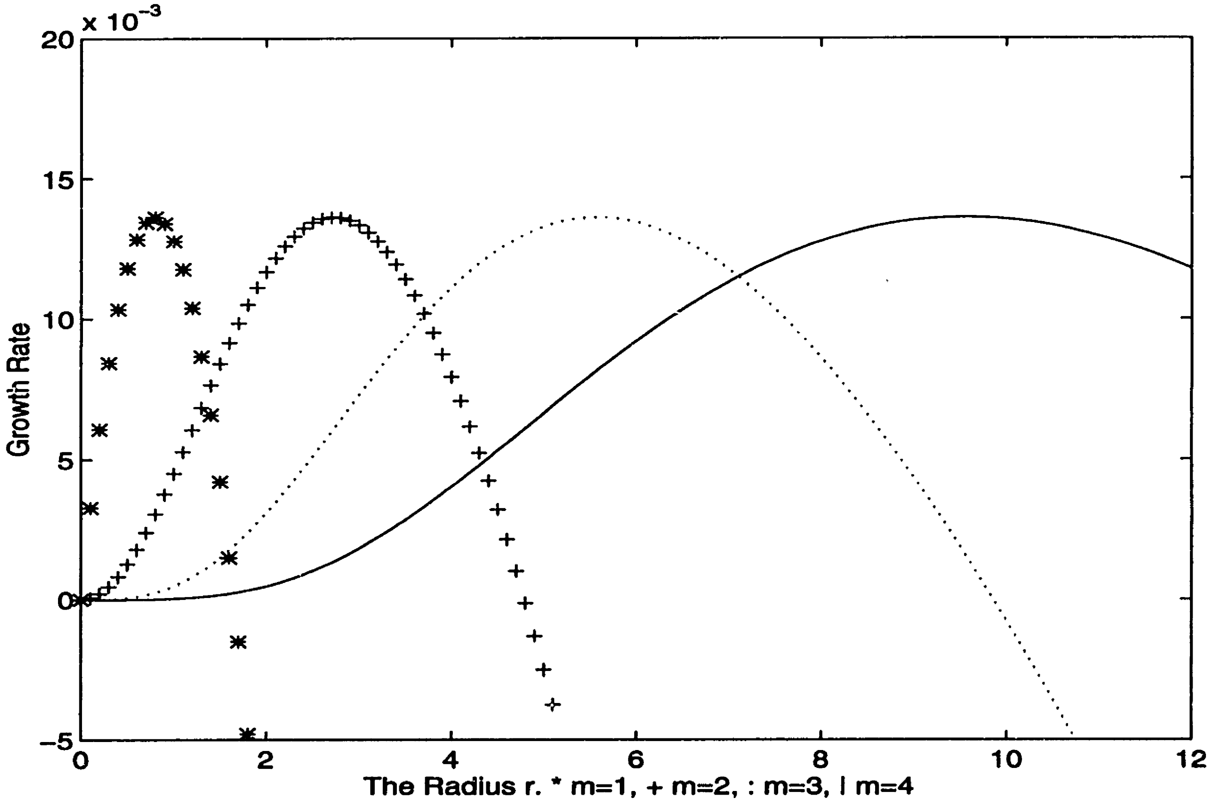

Figure B-2: The Growth Rate as the Radius, r, Varies

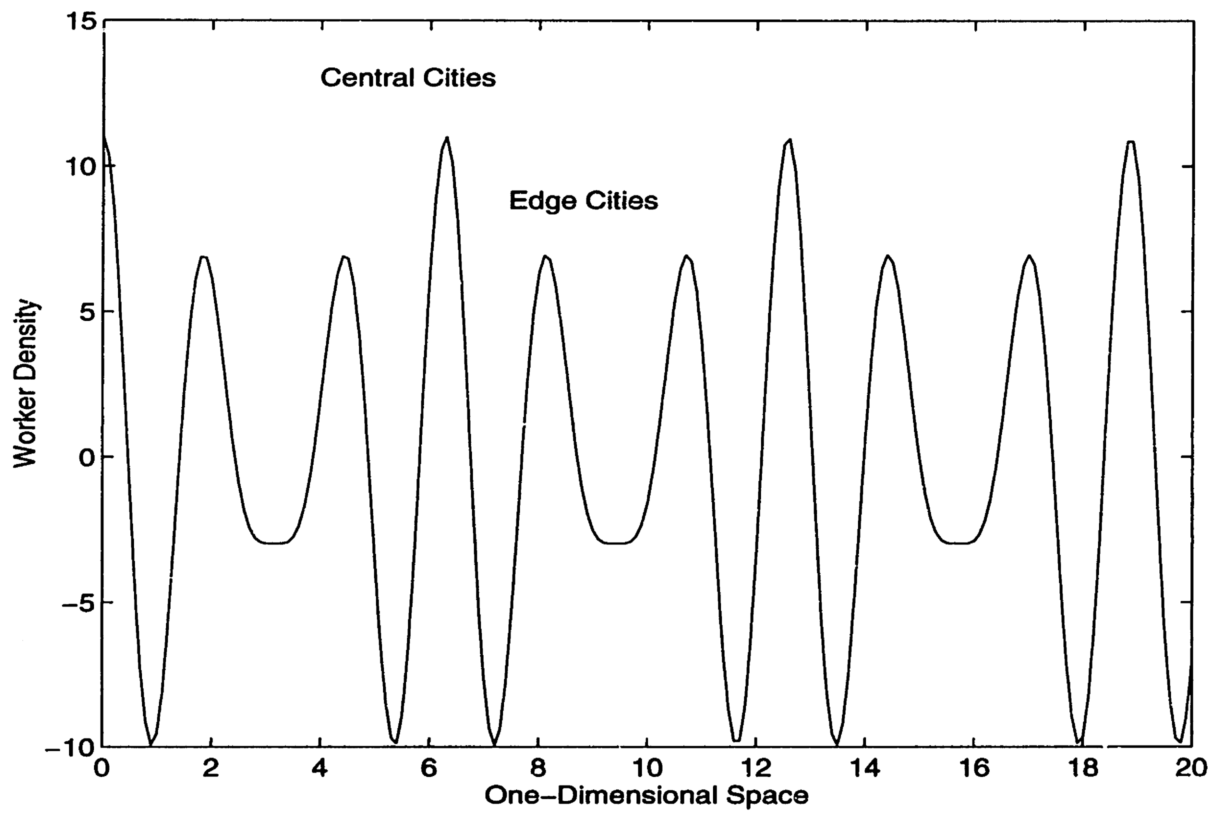

Figure B-3: Possible Spatial Structure when there Exist Two Dominant Growth Frequencies: $\alpha = 7$, $\beta = 4$, $m_1 = 3$, and $m_2 = 4$

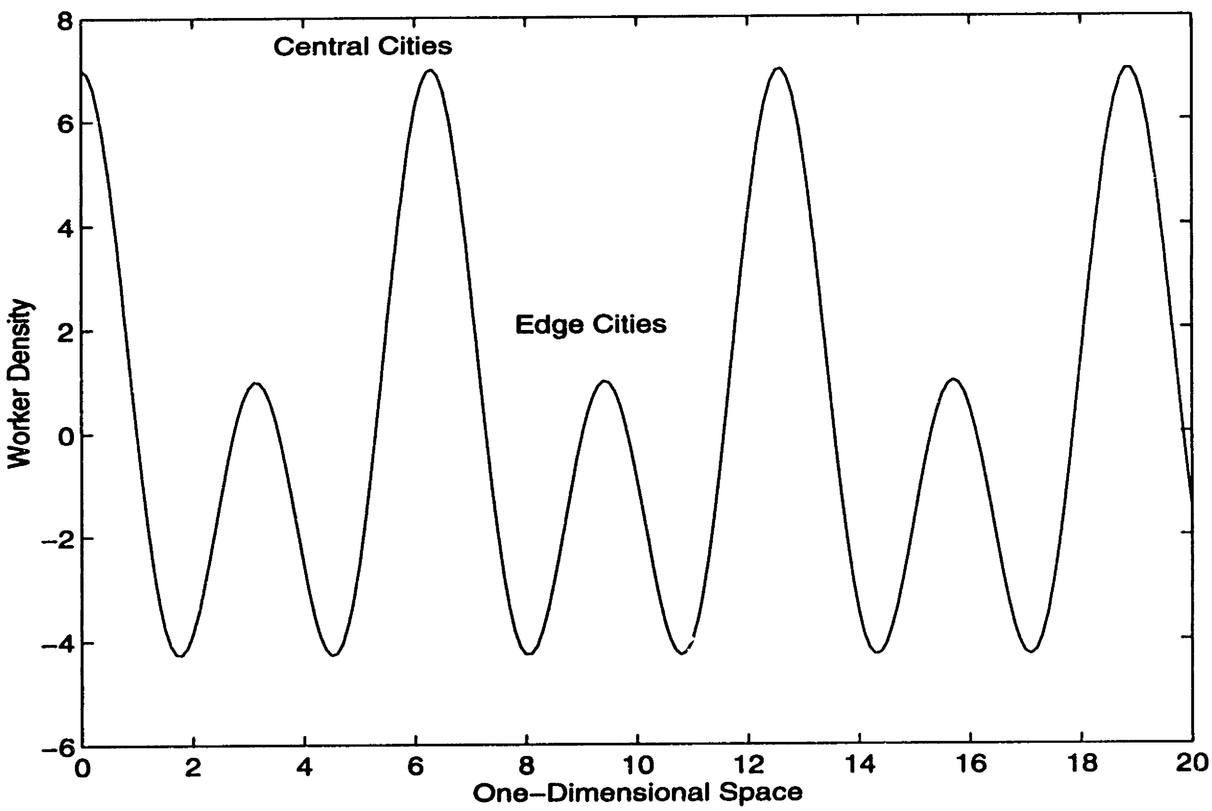

Figure B-4: Possible Spatial Structure when there Exist Two Dominant Growth Frequencies: $\alpha = 3$, $\beta = 2$, $m_1 = 1$, and $m_2 = 2$

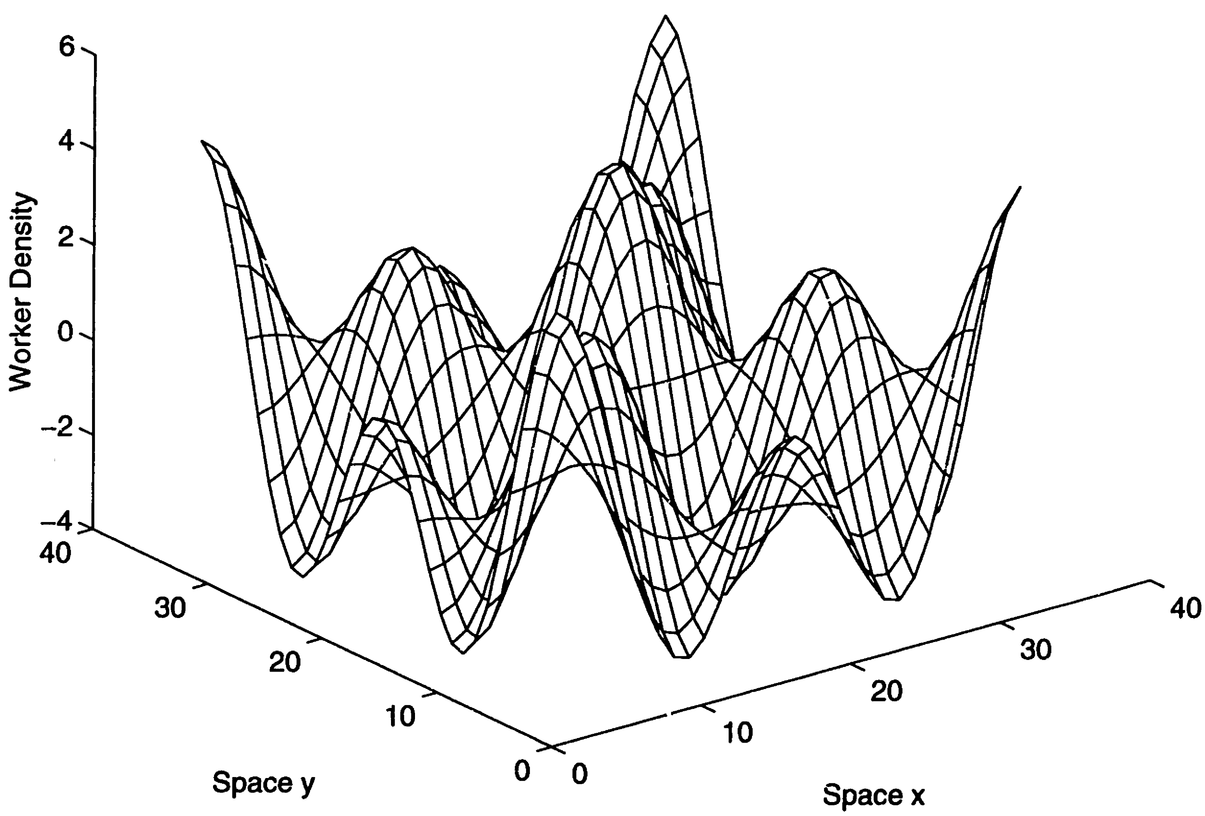

Figure B-5: Possible Two-Dimensional Landscape when there Exist Two Dominant Growth Frequencies

Appendix C

Chapter 1: Edge City Descriptions

Imagine an area that has been formed by the laws of commerce more than governance. A place that prides itself on being safe for women and children; where kids stroll far from their parents sight without having reason to fear being kidnapped or killed. A place where the store owners provide the kind of hospitality that one might find at home; not too extravagant, yet warm and clean. A place where glass elevators are the style, not so much for aesthetics, as much as for promoting a safe, open surrounding. Where one finds beautiful spacious housing, little commuting to the workplace, cosmopolitan attitudes, wildlife intertwined with the array of housing and office buildings, and employment opportunities. A location that contains two-thirds of American office facilities and where eighty percent of this change has occurred within the last two decades. A clustering of activity where even the graffiti provides encouraging statements such as " Words of wisdom: Be excellent to each other." A place that brings jobs, community, and nature together; a sort of Eden, a galactic city, a superburbia, a tommorrowland, or just simply an Edge City.

The brief introduction to this section gave one a glamorized vision of Edge Cities. Perhaps this was an exaggeration, but there still remain many other details about Edge Cities that need to be explained. Edge Cities are so termed because these cities form around major metropolitan cities. An Edge City is an area that brings together jobs, market places, and residential areas. This definition certainly sounds like that of a city. The precise definition of an Edge City is that (a) it has five million square feet or more of leasable office space (b) has 600,000 square feet of leasable retail space (c) has more jobs than bedrooms, thus again indicating that it is primarily a city whose population rises at 9:00 A.M. and decreases at 5:00 P.M. (d) is perceived by the population as one place and (e) was nothing like a city thirty years ago. Edge cities, believes David Birch, were formed by the desire of companies to move to a place where people could get to easily. It was small and medium firms that created Edge Cities. In the Boston area, there is one of the oldest edge cities at 128 and the Mass Pike, west of Boston. This Edge City dates back 30 years.[1] Edge City growth has patterns and limits to its growth. The limits are: (a) insurmountability (b) affordability (c) mobility (d) accessibility, and (e) nice. Most of these limits are quite obvious. Insurmountability refers to physical constraints of the region, e.g. you do not want to build an Edge City in a marsh. Affordability refers to the fact that the region must be affordable so that you have workers to fill jobs. Workers will only move to the region if they have affordable living. Mobility is the ease with which one gets around the place, especially as far as commuting costs are concerned. Accessibility is just a measure of how hard it is to get to the area from outside. David Shulman, director of research at Salomon brothers said "Southern New Hampshire is just too far from Logan.", which was his explanation for why it did not form an Edge City despite its strong growth and apparent advantages. Nice refers to what people want to live in, e.g. schools, cultural events, country clubs, abundant parks, athletic clubs, etc. A rule of thumb for Edge Cities is that when a company moves its headquarters, the move always reduces the commute of the CEO, and presumably he lives in a nice area. So those are all the limits in a brief look. As long ago as 1900, H.G. Wells noted in "The Probable Diffusion of Great Cities," that as technology and specifically transportation improved, there would be less need for central location. In fact, many Edge Cities are formed at highway intersections.

There are different types of Edge Cities. They are: (a) uptowns (b) boomers, and (c) greenfields. The first, uptowns, are essentially places that were old style towns

[1]For a list of Edge Cities across the United States and in Europe one is referred to the book by Joel Garreau.

beforehand, such as Alexandria, Virginia. Most of these areas were shaped when walking was the primary mode of transportation. Boomers, are the classic Edge City. These are usually located at the intersection of freeways and almost always centered on a mall. The boomer category can be further subdivided into (a) strips (b) nodes and (c) pigs in the python. The strip is just an urban form that goes along forever, miles long and only hundreds of yards wide, on either side of the freeway. A node is a relatively dense and contained area, like Tysons Corner, Virginia. Think of being able to draw a circle around these type of Edge Cities. The pig in the python is a cross between the two. The final form is the greenfield. Accordingly named because it forms at the intersection of several thousand acres of farmland and some developer's ego. It is really the extreme notion of marketplace formation. Examples include, Las Colinas, near the Dallas-Fort Worth airport, and Bishop Ranch, east of San Fransisco. Commutes to Edge Cities are typically two-thirds the typical commute to old downtown. Since there are problems with public transportation (people don't like buses and trains are expensive), cars are the primary mode of transportation. Edge Cities also happen to be too small to justify the use of railway. Computer industry analysts think that no more than three to five percent of workers will ever be home based. It is generally believed that the benefits from face-to-face contact will always keep these agglomerations together. Thus, Edge Cities will continue because people gain much by just everyday encounter with one another. Edge Cities in Atlanta contain mainly black people, which some believe supports the idea that Edge Cities are not a race phenomena or a result of "white flight". Pheonix is a great example of the other dimensions of Edge Cities such as shadow governments. Essentially, in certain Edge Cities office parks are in the child rearing business, parking lot officials run police forces, private enterprise builds public freeways, and subdivisions have a say in who lives there. Because Edge Cities really have no political boundaries they must form their own laws. In fact, these groups are very powerful and serve the needs of the residents perhaps much more efficiently than any local government could. A quick review of these shadow governments might help. They are both similar and different to traditional governments. They are similar because (a) they can assess mandatory fees to support themselves, i.e. can tax, (b) they can create rules and regulations, and (c) they have the power to coerce, i.e. police power. They are different because (a) the leaders are not directly accountable to all the people in a general election, (b) usually voting takes place as one dollar, one vote, and (c) these leaders are not usually subject to the constraints that the constitution imposes. For instance, if one wants to buy a home, it is very difficult to avoid the homeowners' association which has certain rules. But the beauty is that people voluntarily choose to subjugate themselves to the rules and laws of this shadow government.[2] For example, in Estrella, an Edge City in the Pheonix area, pornographic material was banned from households themselves. In fact, Pfister, began his shadow government when he decided that regular governments had no idea of what they were doing. A scholar at Yale, Ellickson, claims that shadow governments are nowhere as oppressive as conventional governments, which might have been a fear of some people.

Edge Cities are not just a United States phenomena. Edge Cities are forming even in countries whose social and economic policies and conditions are different from those of the United States. One can find Edge Cities around the area surrounding Toronto, Canada, or in the area surrounding Paris, France; in particular an area known as La Defense. And there are many other greenfield type Edge Cities around Paris. In Sydney, Australia there are Edge Cities forming. In London, England they are forming twenty to thirty miles from the center. In Seoul, Bangkok, and Mexico the same phenomena is being witnessed. European cities, especially, are finding that they are not structured for this new system of production, that is their cities are not well prepared for it, so Edge Cities form.

[2]This is not always the case. In a recent article in the Wall Street Journal, some resident complained that he did not agree with the restrictions on building fences.

Appendix D

Chapter 2: Definition of Variables

D.1 PUMA Characteristics

• **avwage** represents the average wage or per capita income in a region. There are several variations on this variable. **AVWAGE1**: Personal income per capita in 1984 (dollars); The Census defines total personal income as the current income received by residents of an area from all sources. It is measured before deductions of income and other personal taxes but after the deduction of personal contributions for social security, government retirement, and other social insurance programs. The per capita data are based upon estimated population for 1984. **AVWAGE2**: Total personal income in 1984 (millions of dollars). **AVWAGE3**: Money income per capita in 1985. **AVWAGE4**: Money income per capita in 1979 in current dollars. The Census defines money income as the sum of wage or salary income, net nonfarm self-employment income, net farm self-employment income, social security and railroad retirement income, public assistance income, and all other regularly received income such as interest, dividends, veterans' payments, pensions, unemployment compensation, and alimony. Capital gains and inheritances were not included as income. This represents the income before personal income tax deductions, social security, bond purchases, union dues, Medicare deductions, etc. For 1979, the per capita data comes from the 1980 census population data and for 1985, the per capita data comes from 1986 estimated population data. **AVWAGE5**: Money income per capita in 1979 in constant dollars. **AVWAGE6**: Median household income in 1979. The definition of this is the same as that for money income. **AVWAGE7**: Manufacturing wages per production worker in 1982. The Census defines this as all forms of compensation, such as salaries, wages, commissions, dismissal pay, all bonuses, vacation and sick leave pay, and compensation-in-kind, prior to such deductions as employees' social security contributions, withholding taxes, group insurance, union dues, and savings bonds.

• **realwag** is a variable that accounts for living costs. It is defined as **REALWAG** = AVWAGE/COSTS1.

• **tax** is a variable to indicate the local taxes in a region. **TAX1**: Local government taxes per capita. **TAX2**: Local government property taxes per capita. Both were obtained from the 1982 Census of Governments. Taxes consist of compulsory contributions exacted by governments for public purposes, except employee and employer assessments for retirement and social insurance purposes, which are classified as insurance trust revenue, and special assessments, which are classified as nontax general revenue. Property taxes are taxes conditioned on ownership of property and measured by its value. See the Census for more details.

• **futtax** is the future tax expected in a region. **FUTTAX1**: Local government debt per capita in 1981-1982. This includes all long-term debt obligations of the government and its agencies (exclusive of utility debt) and all interest-bearing short-term (i.e. payable within one year) debt obligations remaining unpaid at the close of the fiscal year. **FUTTAX2**: Local government debt outstanding in 1981-1982. See the Census for more details.

• **futgro** is a variable that attempts to get at the expected future growth of a region. **FUTGRO1**: Personal income, percentage change per capita from 1980-1984. Personal income has been defined above. The 1984 per capita data is obtained from estimates of the 1984 resident population.[1] **FUTGRO2**: New private housing units authorized by permit in 1986. **FUTGRO3**: Same from 1980-1986. **FUTGRO4**: Value of nonresidential building authorized by permit construction in 1986. **FUT-**

[1]The personal income is in nominal dollars, and thus is not inflation adjusted.

GRO5: The value of residential building authorized by permit in 1986. **FUTGRO6**: Retail sales, percentage change from 1977-1982; Sales represent sales of establishments in business at any time during the period. An establishment is a single physical location at which business is conducted. When two or more activities were carried on at a single location under a single ownership, all activities were grouped together as a single establishment, unless distinct and separate activities were being carried out (i.e. different industry classification codes were appropriate). Value represents the cost of construction as recorded on the building permit. The data relate to new housing units intended for occupancy on a housekeeping basis. They exclude mobile homes (trailers), hotels, motels, and group residential structures, such as nursing homes and college dormitories. They also exclude conversions of and alterations to existing buildings. **FUTGRO7**: Growth in money income per capita from 1980 to 1984. **POPCHG2**: Percent population change from 1980-1986. **NETMIG**: net migration from 1980-1986; These come from a combination of the Census of 1980 with another report entitled *1986 Population Estimates by County with Components of Change.* I actually use net migration divided by the population of the region.

• **unemp** is the unemployment of the location. **UNEMP1**: Unemployment rate in 1986. **UNEMP2**: Number of unemployed in 1986. **UNEMP3**: Change in the labor force from 1985-1986. **UNEMP4**: Civilian labor force in 1986. The unemployed are defined as all persons who did not work during the survey week, made specific efforts to find a job in the prior four weeks, and were available for work during the survey week (except for temporary illness). Persons waiting to be called back to a job from which they had just been laid off and those waiting to report to a new job within the next thirty days are also included in the unemployment figures. The source is the Bureau of Labor Statistics annual survey.

• **crime** It is interesting to consider whether it is per capita or total crime that people respond to. **CRIME1**: Serious crimes reported to police in 1985. **CRIME2**: Violent crimes reported to police in 1985. **CRIME3**: Serious crimes per 100,000 people reported in 1985. Violent crimes are defined as (a) murder or nonnegligent manslaughter, as defined in the UCR Program, is the willful killing of one human being by another. This excludes deaths caused by negligence, suicide, or accident; justifiable homicides; and attempts to murder or assaults to murder. (b) Forcible rape. Assaults or attempts to commit rape by force are also included. However, statutory rape (without force) and other sex offenses are excluded. (c) Robbery is the taking or attempting to take anything of value from the care, custody, or control of a person or persons by force or threat of force or violence and/or by putting the victim in fear. (d) Aggravated assault is an unlawful attack by one person upon another for the purpose of inflicting severe or aggravated bodily injury. This type of crime is usually accompanied by the use of a weapon. Attempts are also included. Serious crimes include violent crimes and the property crimes of burglary, larceny-theft, and motor vehicle theft. Arson is not included in this data. The source is the Federal Bureau of Investigation, unpublished data.

- **educ** This is a measure of both the education level of PUMA residents and an indication of the quality of schools in the region. **EDUC1**: Percentage of people with 12 or more years of school. **EDUC2**: Percentage of people with 16 or more years of school. One might also care about what type of schools are available for children. These come from the Census of 1980. **EDUC3**: Government expenditure for education per capita this is for 1982 expenditure on elementary and secondary schools. **EDUC4**: Public school enrollment in 1986-1987. **EDUC5**: Public school enrollment in 1980. **EDUC6**: Local government expenditures for education in 1982 (millions of dollars). **EDUC7**: Government expenditure for education, percentage of budget. This includes expenditure on higher education.
- **climate** This is a measure of the weather in an area. **ANN**: Mean average daily annual temperature from 1951-1980. **JAN**: Mean average daily January temperature from 1951-1980. **JULY**: Mean average daily July temperature from 1951-1980. **RAIN**: Annual average precipitation in inches from 1951-1980.[2]
- **race** This controls for the predominant race in an area. **RACEW**: Percentage white in 1984. **RACEB**: Percentage black in 1984. **RACEA**: Percentage asian in

[2]This data was obtained from the weather books of actual weather stations in various counties across the United States. See the bibliography for the source.

1984. **RACEH**: Percentage hispanic in 1984. **RACEI**: Percentage indian in 1984. Computed from population estimates again.

- **demographics** This controls for the median age of the population in that location. **MEDAGE**: The mean age of an area. **OLDPER**: The percentage of people over the age of 55.
- **soccond** This refers to social condition of the area. **SOCCOND1**: Percentage of households with a female head in 1980. **SOCCOND2**: Percentage of single headed households in 1980. The occupants of a household may be a single family, one person living alone, two or more families living together, or any other group of related or unrelated groups of people that share a household. The data does exclude people in group quarters. **SOCCOND3**: Births to mothers less than twenty years old in 1984. **SOCCOND4**: Percent of persons below poverty level in 1979. **SOCCOND5**: Percentage of families below poverty level in 1979. Families and persons were classified as below poverty level if their total family income or unrelated individual income was less than the poverty threshold specified for the applicable family size, age of householder, and number of children under 18 present. Inmates of institutions, persons in military group quarters and college dormitories, and unrelated individuals under 15 years old are excluded.
- **health** This represents the availability of good access to health care. **HEALTH1**: Physicians per 100,000 in 1985. The number of physicians represents the distribution of non-Federal physicians with known addresses who are professionally active in the United States. The source is the American Medical Association's Physician Master File which contains all physicians who are in the United States and meet the U.S. education standards for primary recognition as physicians. **HEALTH2**: Hospital beds per 100,000 in 1985.
- **diverse** This refers to the amount of variety there is in terms of entertainment and goods.[3] **DIVERSE1**: Retail sales per capita in 1982.

---

[3]This tries to capture the Dixit-Stiglitz sort of preferences that people move to where there is more variety. There is no obvious measure that is suitable. I use an indicator of metropolitan area to capture some of this feature. See the variable listed under central city indicator below.

• **density**[4] **DENSITY1**: Population per square mile in 1986. From population estimates for July 1, 1986. **DENSITY2**: Farm acreage in 1982, convert to per capita. The acreage designated as "land in farms" consists primarily of agricultural land used for crops, pasture, or grazing. It also includes woodland and wasteland not actually under cultivation or used for pasture or grazing, provided it was part of the farm operator's total operation. All land in Indian reservations used for growing crops or grazing livestock was included as land in farms. The census did its best to correct for this problem of some farms spilling over county boundaries. **DENSITY3**: Average size of farm in 1982 (acres).

• **central city indicator** A dummy variable that is one for central city and zero otherwise. **CENCTY1** is one for metropolitan areas and zero otherwise. **CENCTY2** is one for the central part of the city and zero otherwise.

• **costs** This represents the cost of living in an area. **COSTS1**: Median value of occupied housing units in 1980 (dollars). **COSTS2**: Average value of land and buildings per farm in 1982 (thousands of dollars). **COSTS3**: Average value of land and buildings per acre in 1982. Respondents were asked to report their estimate of the current market value of land and buildings owned, rented, or leased from others, and rented or leased to others. Market value refers to the respondent's estimate of what the land and buildings would sell for under current market conditions. Where this figure was not reported, a value was placed there of a place with similar characteristics. The last two items come from the *1982 Census for Agriculture, State and County Data.*[5]

• **pollute** Matt Kahn of Columbia University has kindly given me pollution data on particulates for various counties in the United States. **POLLUTE**: Particulates level in various locations. This is divided by total land area to get a better idea of the density of pollution in an area. This data set does not contain carbon monoxide

[4]Density could be a proxy for land rents or just a dislike of crowding.

[5]A better source of costs would be a cost of living index which is provided by two sources. One is by McMahon article and the other is by the foundation ACCRA. However, the former merely has cost of living for states, and the latter for certain selected cities. Neither of these are geographically specialized sufficiently to be used in this chapter.

emissions. This and other measures of pollution can be obtained for metropolitan areas.[6]

• **green** The Parks and Recreation committee has data on various parks throughout the United States.

D.2 Individual Characteristics

• **family PERSON**: number of people following housing record.

• **racehead** This is the race of the migrant. **RACEH**: The race of the head of household.

• **agehead** This is the age of the migrant. **AGEH**: age of household head.

• **educhead** The education level of migrant. **ATTAINH**: The degree of education attained by the head of household.

• **foreign** This is variable that indicates whether the migrant is from a foreign country or not. **CITIZENH**: U.S. or non-U.S. citizen; **IMMIGH**: immigration year.

• **occuph** The occupation of household head. **INDUSTH**: industry; **OCCUPH**: occupation of head; **CLASSH**: class of worker

• **income** The household income of the migrants. **HHINC**: household income; **FAMINC**: family income.

• **married** A variable indicating the marriage status of the migrant.

• **child** A variable indicating the number of children in the household of the migrant.

• **sex pref** A variable indicating the sexual preference of the person. The possibilities are lesbian, gay, or heterosexual.

[6]The CFR Report from the National Affairs has this data. The Green Index is another good source for environmental statistics, although the data only exists starting in 1991.

D.3 Aggregating County Variables into PUMA Variables

All per capita variables were simply disentangled by population size of the individual counties and then averaged again using the county's population as weights. With variables that are nominal, the sum is taken. With climate, the average of the areas is taken. So that the reader can better understand the meaning of a PUMA location, I have included the actual locations for each PUMA of California.[7]

[7] For those who want a more detailed analysis of each state, programmed code may be requested from the author. One should also realize that my PUMAs differ from the ones defined by the census in certain ways. Specifically, the census is interested in geographical areas containing around 100,000 people, while I am concerned with county-type borders.

Table D.1: Aggregated PUMA Locations for California

PUMA	COUNTIES and/or CITIES
49	Del Norte, Lassen, Modoc, Siskiyou
50	Humboldt
51	Shasta
52	Lake, Medocino
53	Colusa, Glenn, Tehama, Trinity
54	Butte
55	Nevada, Plumas, Sierra
56	Sutter, Yuba
57	Napa
58	Yolo
59	Placer
60	El Dorado
61	Santa Rosa City
62	Sonoma
63	Marin
64	Solano
65	San Pablo City, Richmond City
66	Contra Costa
67	San Fransisco
68	Albany City, Berkeley City, Emeryville City
69	Alameda
70	Oakland City
71	Sam Mateo
72	San Joaquin
73	Modesto City
74	Stanilaus
75	Alpine, Mador, Calaveres, Inyo, Mariposa Mono, Tuolumne
76	Madera, San Benito
77	Sacramento City
78	Sacramento
79	Merced
80	Kings
81	Santa Barbara
82	San Diego
83	Santa Clara

Cities are indicated by city following proper name and counties have no such suffix.

Table D.2: Aggregated PUMA Locations for California: Continued

PUMA	COUNTIES and/or CITIES
84	Sunnyvale City
85	Tulare
86	Santa Cruz
87	Imperial
88	Salinas City
89	Monterey
90	Fresno City
91	Fresno
92	Santa Ana City
93	Orange
94	Fountain Valley City, Westminster City
95	Garden Grove
96	Bakersfield City
97	Kern
98	San Luis Obispo
99	Burbank City, San Fernando City
100	Glendale City
101	Monnterey Park City, Rosemead City, South San Gabriel City
102	Los Angeles
103	Lynwood City, South Gate City
104	El Monte City
105	Pomona City
106	Carson City
107	Inglewood City
108	Beverly Hills City, Culver City, Ladera Heights City Marina Del Rey City, West Hollywood City
109	Pasadena City
110	Los Angeles
111	Long Beach City
112	Ventura
113	Riverside
114	Rancho Cucamonga City
115	Ontario City
116	San Bernardino

Cities are indicated by city following proper name and counties have no such suffix.

Appendix E

Chapter 2: Summary Statistics

E.1 Statistics for the All Choice Homogeneous Group

Table E.1: Summary Statistics for the Individual Movers

Variable	Obs.	Mean	S.D.	Min.	Max.
Married(%)	28048	0.58	—	—	—
RaceHead(% W)	28048	0.90	—	—	—
EducHead (% HS)	24605	0.84	—	—	—
Foreign (%)	28048	0.06	—	—	—
Income (House)	28048	34557	33660	-2000	636456
Income (Person)	28048	15814	14975	-1685	245833
Age	28048	40	15	17	90
Children	28048	0.80	1.13	0	8

Table E.2: Summary Statistics for the Location Specific Variables

Variable	Obs.	Mean	S.D.	Min.	Max.
AVWAGE1	26640	12462	2893	6958	22582
COSTS1	28048	44690	16423	19453	124400
TAX1	28048	408	195	38	1962
TAX2	28048	329	159	32	1141
UNEMP1	28048	6.75	2.69	2.2	20.6
FUTTAX1	28048	1157	1006	34	10116
FUTGRO1	26640	42	11	1.43	122
NETMIG	26640	9132	45263	-197500	301400
CRIME3	28048	4326	2294	1022	15122
EDUC3	27656	1601	7253	51	64153
ANN	28032	55	8.8	39	76
RAIN	28032	39	13	4	65
SOCCOND4	28048	12	4	3	31
HEALTH1	26584	179	126	25	647
RACEB	28048	11	10	21	72

E.2 Individual Specific Statistics of Various Groups

Table E.3: Summary Statistics for the Elderly Movers

Variable	Obs.	Mean	S.D.	Min.	Max.
Married (%)	15303	0.527	—	—	—
RaceHead (% W)	15303	0.931	—	—	—
EducHead (% HS)	15303	0.652	—	—	—
Foreign (%)	15303	0.060	—	—	—
Income (Person)	15303	16991	20367	-5000	254119
Age	15303	67	8.26	55	90
Children	15303	0.05	0.31	0	5

Table E.4: Summary Statistics for the Young Movers

Variable	Obs.	Mean	S.D.	Min.	Max.
Married (%)	28048	0.581	—	—	—
RaceHead (% W)	28048	0.900	—	—	—
EducHead (% HS)	28048	0.838	—	—	—
Foreign (%)	28048	0.061	—	—	—
Income (Person)	28044	15813	14975	-1685	245833
Age	28024	40	15	17	90
Children	28048	0.801	0.111	0	8

Table E.5: Summary Statistics for the White Movers

Variable	Obs.	Mean	S.D.	Min.	Max.
Married (%)	7287	0.580	—	—	—
RaceHead (% W)	7287	—	—	—	—
EducHead (% HS)	7287	0.840	—	—	—
Foreign (%)	7287	0.033	—	—	—
Income (Person)	7287	16753	16027	0	209516
Age	7287	40	15	17	90
Children	7287	0.73	1.05	0	7

Table E.6: Summary Statistics for the Black Movers

Variable	Obs.	Mean	S.D.	Min.	Max.
Married (%)	7500	0.370	—	—	—
RaceHead (% W)	7500	—	—	—	—
EducHead (% HS)	7500	0.743	—	—	—
Foreign (%)	7500	0.064	—	—	—
Income (Person)	7500	11203	12258	-2078	195516
Age	7500	39	14	17	89
Children	7500	0.978	1.27	0	8

Table E.7: Summary Statistics for the Asian Movers

Variable	Obs.	Mean	S.D.	Min.	Max.
Married (%)	8433	0.760	—	—	—
RaceHead (% W)	8433	—	—	—	—
EducHead (% HS)	8433	0.910	—	—	—
Foreign (%)	8433	0.827	—	—	—
Income (Person)	8433	16804	15404	-3192	314629
Age	8433	38	11	16	90
Children	8433	1.17	1.21	0	7

Table E.8: Summary Statistics for the Hispanic Movers

Variable	Obs.	Mean	S.D.	Min.	Max.
Married (%)	7686	0.600	—	—	—
RaceHead (% W)	7686	—	—	—	—
EducHead (% HS)	7686	0.51	—	—	—
Foreign (%)	7686	0.533	—	—	—
Income (Person)	7686	9447	9386	0	85731
Age	7686	35	11	18	90
Children	7686	1.25	1.41	0	8

Table E.9: Summary Statistics for the High School Movers

Variable	Obs.	Mean	S.D.	Min.	Max.
Married (%)	11984	0.584	—	—	—
RaceHead (% W)	11984	0.880	—	—	—
EducHead (% HS)	11984	0.621	—	—	—
Foreign (%)	11984	0.069	—	—	—
Income (Person)	11980	10539	8855	0	135680
Age	11984	43	17	17	90
Children	11984	0.870	1.18	0	7

Table E.10: Summary Statistics for the Bachelors Degree Movers

Variable	Obs.	Mean	S.D.	Min.	Max.
Married (%)	14814	0.534	—	—	—
RaceHead (% W)	14814	0.919	—	—	—
EducHead (% HS)	14814	1	—	—	—
Foreign (%)	14814	0.058	—	—	—
Income (Person)	14814	24105	21218	-2100	253446
Age	14814	37	13	19	90
Children	14814	0.610	1.03	0	7

Table E.11: Summary Statistics for the Doctorate Movers

Variable	Obs.	Mean	S.D.	Min.	Max.
Married (%)	14865	0.628	—	—	—
RaceHead (% W)	14865	0.894	—	—	—
EducHead (% HS)	14865	1	—	—	—
Foreign (%)	14865	0.102	—	—	—
Income (Person)	14865	27402	22996	0	319031
Age	14769	41	12	21	90
Children	14865	0.720	1.05	0	7

Table E.12: Summary Statistics for the Gay Movers

Variable	Obs.	Mean	S.D.	Min.	Max.
Married (%)	3702	0.118	—	—	—
RaceHead (% W)	3702	0.912	—	—	—
EducHead (% HS)	3702	0.943	—	—	—
Foreign (%)	3702	0.060	—	—	—
Income (Person)	3702	29285	21880	-5999	224516
Age	3702	34	10	18	88
Children	3702	0.161	0.566	0	5

Table E.13: Summary Statistics for the Lesbian Movers

Variable	Obs.	Mean	S.D.	Min.	Max.
Married (%)	3312	0.025	—	—	—
RaceHead (% W)	3312	0.833	—	—	—
EducHead (% HS)	3312	0.938	—	—	—
Foreign (%)	3312	0.034	—	—	—
Income (Person)	3312	20706	12833	0	117905
Age	3312	38	11	18	90
Children	3312	0.170	0.584	0	6

Table E.14: Summary Statistics for the Foreign Born Movers

Variable	Obs.	Mean	S.D.	Min.	Max.
Married (%)	18690	0.681	—	—	—
RaceHead (% W)	18690	0.456	—	—	—
EducHead (% HS)	18690	0.709	—	—	—
Foreign (%)	18690	1	—	—	—
Income (Person)	18690	14979	15200	-5999	314629
Age	18690	42	14	18	90
Children	18690	1.05	1.26	0	9

Appendix F

Chapter 3: Mathematical Derivations

F.1 Derivation of Equation 3.7

$$\bar{\pi}_t(k) = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \gamma_{mt-k})R_{it} \quad \text{(F.1)}$$

$$= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \frac{N_p}{N}\gamma_{pt-k} - \frac{N_d}{N}\gamma_{dt-k})(R_{it} + R_t^p - R_t^p) \quad \text{(F.2)}$$

$$= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \frac{N_p}{N}\gamma_{pt-k})(R_{it} - R_t^p) \underbrace{- \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}\Omega_{d,pt-k}(R_{it} - R_t^p)}_{=0} +$$

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \gamma_{mt-k})R_t^p \quad \text{(F.3)}$$

$$= -\frac{1}{T}\sum_{t=1}^{T}N_p(\gamma_{t-k}^p - \bar{\gamma}^p)(R_t^p - \bar{R}^p) + \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \bar{\gamma}^p)(R_{it} - \bar{R}^p) +$$

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \gamma_{mt-k})R_t^p \quad \text{(F.4)}$$

$$= -\frac{1}{T}\sum_{t=1}^{T}N_p(\gamma_{t-k}^p - \bar{\gamma}^p)(R_t^p - \bar{R}^p) + \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \bar{\gamma}_i)(R_{it} - \bar{R}_i) +$$

$$\sum_{i=1}^{N_p}(\bar{\gamma}_i - \bar{\gamma}^p)(\bar{R}_i - \bar{R}^p) + \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \gamma_{mt-k})R_t^p, \quad \text{(F.5)}$$

where $\Omega_{d,pt-k} = \frac{N_d}{N} \frac{\sum_{i=1}^{N_d} V_{it-k}}{\sum_{i=1}^{N_p}(V_{it-k} - V_{mt-k})}$ is a measure of the weighted volumes of securities traded that are outside of the specific portfolio, yet part of the N total securities divided by the familiar volume weighting for that portfolio.

F.2 Derivation of Equation F.4

It will be sufficient to show that:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \gamma_{pt-k})(R_{it} - R_t^p) = -\frac{1}{T}\sum_{t=1}^{T} N_p(\gamma_{t-k}^p - \bar{\gamma}^p)(R_t^p - \bar{R}^p) + \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \bar{\gamma}^p)(R_{it} - \bar{R}^p). \quad \text{(F.6)}$$

$$\begin{aligned}
LHS: &= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \frac{N_p}{N}\gamma_{t-k}^p)(R_{it} - R_t^p) \\
&= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}[\gamma_{it-k}R_{it} - \gamma_{it-k}R_t^p - \frac{N_p}{N}\gamma_{t-k}^p R_{it} + \frac{N_p}{N}\gamma_{t-k}^p R_t^p] \\
&= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}\gamma_{it-k}R_{it} - \frac{1}{T}\sum_{t=1}^{T} N_p\gamma_{t-k}^p R_t^p. \\
RHS: &= -\frac{1}{T}\sum_{t=1}^{T} N_p(\gamma_{t-k}^p - \bar{\gamma}^p)(R_t^p - \bar{R}^p) + \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \bar{\gamma}^p)(R_{it} - \bar{R}^p) \\
&= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}\gamma_{it-k}R_{it} - \frac{1}{T}\sum_{t=1}^{T} N_p\gamma_{t-k}^p R_t^p + N_p\bar{\gamma}^p\bar{R}^p - N_p\bar{\gamma}^p\bar{R}^p \\
&= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}\gamma_{it-k}R_{it} - \frac{1}{T}\sum_{t=1}^{T} N_p\gamma_{t-k}^p R_t^p. \qquad \square
\end{aligned}$$

F.3 Derivation of Equation F.5

It will be sufficient to show that:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \bar{\gamma}^p)(R_{it} - \bar{R}^p) = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \bar{\gamma}_i)(R_{it} - \bar{R}_i) +$$

$$\sum_{i=1}^{N_p}(\bar{\gamma}_i - \bar{\gamma}^p)(\bar{R}_i - \bar{R}^p). \tag{F.7}$$

$$\begin{aligned}
LHS: &= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \bar{\gamma}^p)(R_{it} - \bar{R}^p) \\
&= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}[\gamma_{it-k}R_{it} - \gamma_{it-k}\bar{R}^p - \bar{\gamma}^p R_{it} + \bar{\gamma}^p\bar{R}^p] \\
&= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}\gamma_{it-k}R_{it} - N_p\bar{\gamma}^p\bar{R}^p. \\
RHS: &= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}(\gamma_{it-k} - \bar{\gamma}_i)(R_{it} - \bar{R}_i) + \sum_{i=1}^{N_p}(\bar{\gamma}_i - \bar{\gamma}^p)(\bar{R}_i - \bar{R}^p) \\
&= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}\gamma_{it-k}R_{it} - \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}\gamma_{it-k}\bar{R}_i - \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}\bar{\gamma}_i R_{it} + \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}\bar{\gamma}_i\bar{R}_i \\
&\quad + \sum_{i=1}^{N_p}[\bar{\gamma}_i\bar{R}_i - \bar{\gamma}_i\bar{R}^p - \bar{\gamma}^p\bar{R}_i + \bar{\gamma}^p\bar{R}^p] \\
&= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}\gamma_{it-k}R_{it} \underbrace{- \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}\gamma_{it-k}\bar{R}_i + \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}\bar{\gamma}_i\bar{R}_i}_{=0} \\
&\quad \underbrace{- \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}\bar{\gamma}_i R_{it} + \sum_{i=1}^{N_p}\bar{\gamma}_i\bar{R}_i}_{=0} - N_p\bar{\gamma}^p\bar{R}^p \\
&= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_p}\gamma_{it-k}R_{it} - N_p\bar{\gamma}^p\bar{R}^p. \qquad \square
\end{aligned}$$

The reader should be aware that the terms in underbraces are only equal to zero if the summations can be reversed, that is if N_p is independent across time. With our strategies, this is not so. Hence we cannot make the decomposition from equation F.4 to equation F.5. We left this step in for those who deal with a constant N over time.

Appendix G

Chapter 3: Tables

Table G.1: Mean Profits of All Four Strategies

Portfolio	$R>0$ $V>V_m$	$R>0$ $V<V_m$	$R<0$ $V>V_m$	$R<0$ $V<V_m$	$R>0$	$R<0$	$V>V_m$	$V<V_m$	Total Profit	Contrarian Profit
Panel A: Using the Volume Lead Strategy										
All Stocks	-0.285	-1.35	7.38	-1.92	-1.63	5.46	7.09	-3.2	3.82	10.38
	(-0.289)	(-2.42)	(7.667)	(-3.22)	(-2.296)	(7.89)	(8.12)	(-13.68)	(4.323)	(11.18)
Smallest	0.114	-3.7	12.75	-3.86	-3.58	8.8	12.8	-7.57	5.29	20.45
	(0.0741)	(-4.923)	(8.904)	(-4.44)	(-2.74)	(7.20)	(7.42)	(-13.82)	(2.94)	(11.15)
Medium	-0.124	-0.578	5.17	-1.23	-0.702	3.94	5.05	-1.8	3.23	6.88
	(-0.1216)	(-0.99)	(4.789)	(-2.207)	(-0.7752)	(4.32)	(4.03)	(-4.73)	(2.507)	(5.17)
Largest	-0.71	-0.621	4.365	-1.68	-1.34	2.67	3.6	-2.3	1.337	5.96
	(-0.716)	(-1.121)	(4.23)	(-2.964)	(-1.54)	(3.07)	(3.17)	(-6.70)	(1.12)	(4.94)
Panel B: Using Volume as an Indicator and Returns as Weights										
All Stocks	2.63	-3.77	10.66	-7.77	-1.14	2.89	13.29	-11.55	1.75	24.87
	(2.916)	(-5.531)	(11.264)	(-4.965)	(-1.847)	(1.890)	(17.133)	(-7.888)	(1.077)	(14.74)
Smallest	3.34	-8.09	15.7	-13.53	-4.7	2.19	19.07	-21.6	-2.55	40.73
	(2.266)	(-8.41)	(11.579)	(-5.077)	(-3.66)	(0.82)	(11.73)	(-8.26)	(-0.844)	(12.97)
Medium	1.102	-1.94	6.648	-3.866	-0.84	2.78	7.75	-5.8	1.93	13.57
	(1.18)	(-2.66)	(6.345)	(-5.483)	(-0.96)	(3.14)	(6.80)	(-8.79)	(1.504)	(10.08)
Largest	0.53	-0.73	4.58	-3.94	-0.207	0.639	5.115	-4.68	0.432	9.81
	(0.60)	(-1.269)	(4.268)	(-5.964)	(-0.28)	(0.70)	(4.64)	(-10.15)	(0.3699)	(8.04)
Panel C: Using the Volume Lead Strategy Combined with Return Information (VLR1)										
All Stocks	2.37	-4.14	12.4	-7.53	-1.77	4.89	14.8	-11.6	3.126	26.50
	(1.87)	(-0.027)	(9.73)	(-4.118)	(-1.63)	(2.43)	(9.73)	(-6.659)	(1.357)	(11.32)
Smallest	2.07	-8.57	17.1	-12.7	-6.5	4.3	19.2	-21.3	-2.12	40.62
	(1.07)	(-8.81)	(10.83)	(-4.459)	(-3.54)	(1.43)	(8.50)	(-7.51)	(-0.0579)	(11.24)
Medium	1.33	-1.68	6.56	-2.98	-0.34	3.58	7.89	-4.6	3.23	12.57
	(1.164)	(-2.202)	(5.239)	(-4.24)	(-0.30)	(3.06)	(5.10)	(-6.45)	(1.90)	(7.33)
Largest	-0.08	-0.88	5.244	-3.359	-0.96	1.88	5.16	-4.2	0.924	9.41
	(-0.073)	(-1.49)	(4.44)	(-5.1135)	(-0.94)	(1.81)	(3.86)	(-9.14)	(0.0639)	(6.64)
Panel D: Using the Volume Lead Strategy Combined with Return Information (VLR2)										
All Stocks	2.65	-3.78	10.67	-7.78	-1.14	2.90	13.32	-11.56	1.76	24.87
	(2.93)	(-5.54)	(11.26)	(-4.97)	(-1.85)	(1.89)	(17.15)	(-7.89)	(1.08)	(14.73)
Smallest	3.33	-8.10	15.75	-13.55	-4.77	2.20	19.08	-21.65	-2.57	40.72
	(2.25)	(-8.42)	(11.58)	(-5.08)	(-3.67)	(0.823)	(11.70)	(-8.26)	(-0.848)	(12.96)
Medium	1.10	-1.95	6.65	-3.87	-0.844	2.78	7.75	-5.82	1.94	13.57
	(1.18)	(-2.67)	(6.34)	(-5.48)	(-0.964)	(3.14)	(6.79)	(-8.79)	(1.50)	(10.07)
Largest	0.531	-0.739	4.59	-16.68	-0.208	-12.10	5.12	-17.42	-12.30	9.81
	(0.603)	(-1.27)	(4.27)	(-1.66)	(-0.281)	(-1.20)	(4.64)	(-1.74)	(-1.22)	(8.04)

These are all the returns and profits from various strategies using a one week lag, that is $k = 1$. The reader should realize that the various portfolios are shorthands for the strategies, thus $R > 0$ and $V > V_m$ indicates the portfolio of stocks for which the individual stock return was greater than zero and the individual volume was greater than market volume at $t - k$. The other portfolios are as discussed in the section 3.3 of chapter 3. The $R > 0$ portfolio indicates the combined profits from the portfolios with positive return in period $t - k$, yet still using the weights as defined by each strategy. Similarly, the $R < 0$ portfolio indicates the combined profits from the portfolios with negative return in period $t - k$. The $V > V_m$ portfolio is the combined profits from the high volume stocks in period $t - k$ using the weights described by the strategy. The $V < V_m$ portfolio is the combined profits from the low volume stocks in period $t - k$ using the weights described by the respective strategy. Total profit is the sum of returns of the four portfolios of each strategy. It is a costless portfolio, and it represents true profits. The contrarian profits indicate the profits from using the same strategies as outlined in section 3.3 however, we change them so as to sell winner stocks ($R > 0$) and buy loser stocks ($R < 0$); this is also a costless portfolio. The t statistics presented in the parentheses are Newey-West corrected using four lags. The sample consists of 1559 weeks. The returns and profits presented have been multiplied by 1000.

Table G.2: Mean Profits of All Four Strategies For Lags $k = 2, 4$, and 26: All Stocks

Lag	$R>0$ $V>V_m$	$R>0$ $V<V_m$	$R<0$ $V>V_m$	$R<0$ $V<V_m$	$R>0$	$R<0$	$V>V_m$	$V<V_m$	Total Profit	Contrarian Profit
				Panel A: Using the Volume Lead Strategy						
2	-0.2	-1.7	11.2	-4.41	-1.9	7.0	10.9	-5.9	5.0	16.86
	(-0.138)	(-1.496)	(5.358)	(-3.223)	(1.422)	(4.869)	(6.262)	(11.23)	(2.833)	(9.00)
4	-1.1	0.4	17.3	-7.6	-0.7	9.7	16.2	-7.2	9.0	23.46
	(-0.305)	(0.177)	(3.636)	(-2.944)	(0.292)	(2.803)	(4.125)	(6.887)	(2.263)	(5.64)
26	-2.9	12.7	31.2	-44.2	9.8	-13.0	28.2	-31.4	-3.2	59.74
	(-0.121)	(1.404)	(1.602)	(-3.371)	(0.528)	(1.058)	(1.595)	(4.432)	(-0.173)	(3.02)
				Panel B: Using Volume as an Indicator and Returns as Weights						
2	0.8	-4.2	12.8	-10.1	-3.4	2.7	13.6	-14.3	-0.7	28.03
	(0.260)	(-3.029)	(7.040)	(-5.111)	(-1.172)	(1.686)	(4.623)	(-9.006)	(-0.208)	(8.06)
4	-2.3	-3.1	21.0	-12.2	-5.4	8.7	18.6	-15.3	3.3	34.01
	(-0.666)	(-1.220)	(3.889)	(-4.059)	(-2.428)	(2.191)	(4.232)	(-9.070)	(0.767)	(6.68)
26	-19.4	29.78	48.4	-57.8	10.3	-9.3	29.0	-28.0	1.0	57.09
	(-0.984)	(1.793)	(2.117)	(-3.735)	(0.492)	(-0.655)	(1.349)	(-1.9228)	(0.0417)	(2.04)
				Panel C: Using the Volume Lead Strategy Combined with Return Information (VLR1)						
2	-1.1	-4.6	15.2	-9.7	-5.8	5.4	14.0	-14.4	-0.4	28.49
	(-0.255)	(-3.394)	(5.979)	(-4.421)	(1.261)	(2.159)	(2.920)	(7.654)	(-0.0776)	(5.34)
4	0.1	-3.6	32.2	-10.7	-3.4	21.4	32.3	-14.3	18.0	46.75
	(0.036)	(-1.474)	(2.902)	(-3.583)	(1.033)	(2.089)	(2.911)	(8.567)	(1.622)	(4.10)
26	-6.1	11.1	38.9	-56.9	4.9	-18.0	32.8	-45.0	-13.0	78.78
	(-0.222)	(1.180)	(1.571)	(-3.904)	(0.231)	(1.052)	(1.362)	(4.616)	(-0.595)	(2.65)
				Panel D: Using the Volume Lead Strategy Combined with Return Information (VLR2)						
2	0.804	-4.23	12.89	-10.12	-3.43	2.77	13.69	-14.35	-0.660	28.04
	(0.261)	(-3.03)	(7.04)	(-5.11)	(-1.17)	(1.69)	(4.62)	(-9.01)	(-0.204)	(8.05)
4	-2.33	-3.08	21.09	-12.26	-5.41	8.83	18.76	-15.35	3.42	34.11
	(-0.655)	(-1.22)	(3.88)	(-4.06)	(-2.41)	(2.19)	(4.23)	(-9.07)	(0.785)	(6.66)
26	-19.34	29.80	48.40	-57.83	10.46	-9.43	29.06	-28.03	1.03	57.09
	(-0.978)	(1.79)	(2.11)	(-3.74)	(0.498)	(-0.662)	(1.35)	(-1.93)	(0.043)	(2.04)

These are all the returns and profits from various strategies using two, four, and twenty six week lags, that is $k = 2, 4$, and 26. The reader should realize that the various portfolios are shorthands for the strategies, thus $R > 0$ and $V > V_m$ indicates the portfolio of stocks for which the individual stock return was greater than zero and the individual volume was greater than market volume at $t - k$. The other portfolios are as discussed in the section 3.3 of chapter 3. The $R > 0$ portfolio indicates the combined profits from the portfolios with positive return in period $t-k$, yet still using the weights as defined by each strategy. Similarly, the $R < 0$ portfolio indicates the combined profits from the portfolios with negative return in period $t - k$. The $V > V_m$ portfolio is the combined profits from the high volume stocks in period $t - k$ using the weights described by the strategy. The $V < V_m$ portfolio is the combined profits from the low volume stocks in period $t-k$ using the weights described by the respective strategy. Total profit is the sum of returns of the four portfolios of each strategy. It is a costless portfolio, and it represents true profits. The contrarian profits indicate the profits from using the same strategies as outlined in section 3.3 however, we change them so as to sell winner stocks ($R > 0$) and buy loser stocks ($R < 0$); this is also a costless portfolio. The t statistics presented in the parentheses are Newey-West corrected using four lags. The sample consists of 1559 weeks. The returns and profits presented have been multiplied by 1000.

Table G.3: Mean Profits of All Four Strategies For Lags k = 2, 4, and 26: Largest Quintile

Lag	$R>0$ $V>V_m$	$R>0$ $V<V_m$	$R<0$ $V>V_m$	$R<0$ $V<V_m$	$R>0$	$R<0$	$V>V_m$	$V<V_m$	Total Profit	Contrarian Profit
Panel A: Using the Volume Lead Strategy										
2	-2.90	-0.161	7.05	-2.66	-3.06	4.39	4.16	-2.82	1.33	6.98
	(-1.38)	(-0.161)	(3.32)	(-2.44)	(-1.65)	(2.40)	(1.81)	(-4.24)	(0.563)	(2.89)
4	-6.06	2.13	13.56	-4.75	-3.94	8.81	7.50	-2.62	4.87	10.12
	(-1.63)	(1.02)	(3.07)	(-2.27)	(-1.22)	(2.51)	(1.69)	(-1.83)	(1.08)	(2.09)
26	-6.30	5.36	71.63	-37.64	-0.934	1.00	65.35	-32.28	33.07	97.62
	(-0.318)	(0.623)	(2.86)	(-3.11)	(-0.048)	(1.44)	(2.63)	(-3.33)	(1.29)	(3.54)
Panel B: Using Volume as an Indicator and Returns as Weights										
2	-1.08	-0.517	7.91	-4.98	-1.60	2.92	6.82	-5.50	1.32	12.32
	(-0.570)	(-0.460)	(4.04)	(-3.95)	(-0.987)	(1.77)	(3.29)	(-6.12)	(0.610)	(5.25)
4	-5.57	2.60	13.46	-5.70	-2.98	7.76	7.89	-3.10	4.79	10.99
	(-1.65)	(1.17)	(3.41)	(-2.41)	(-1.05)	(2.50)	(2.07)	(-1.77)	(1.22)	(2.47)
26	-20.29	8.14	77.11	-46.20	-12.41	30.92	56.83	38.06	18.77	94.88
	(-0.979)	(0.758)	(3.11)	(-4.40)	(-0.653)	(1.47)	(2.11)	(-3.70)	(0.740)	(2.97)
Panel C: Using the Volume Lead Strategy Combined with Return Information (VLR1)										
2	-2.96	-0.911	8.60	-4.42	-3.87	4.18	5.64	-5.33	0.308	10.97
	(-1.34)	(-0.837)	(3.78)	(-3.58)	(-1.97)	(2.02)	(2.20)	(-5.85)	(0.116)	(3.93)
4	-6.70	1.44	15.00	-5.31	-5.26	9.70	8.31	-3.87	4.44	12.18
	(-1.62)	(0.641)	(3.15)	(-2.26)	(-1.45)	(2.39)	(1.65)	(-2.09)	(0.868)	(2.18)
26	-13.42	3.27	96.14	-44.77	-10.15	51.37	82.72	-41.50	41.22	24.23
	(-0.471)	(0.305)	(3.54)	(-4.12)	(-0.356)	(2.18)	(2.62)	(-4.05)	(1.36)	(3.45)
Panel D: Using the Volume Lead Strategy Combined with Return Information (VLR2)										
2	-1.09	-0.517	7.91	-4.98	-1.60	2.92	6.82	-5.50	1.32	12.32
	(-0.571)	(-0.460)	(4.04)	(-3.95)	(-0.987)	(1.77)	(3.29)	(-6.12)	(0.610)	(5.25)
4	-5.58	2.60	13.46	-5.70	-2.98	7.76	7.88	-3.10	4.78	10.99
	(0.067)	(1.16)	(3.40)	(-2.41)	(-1.05)	(2.50)	(2.07)	(-1.77)	(1.22)	(2.47)
26	-20.25	8.14	77.15	-46.20	-12.11	30.96	56.90	-38.05	18.85	94.96
	(-0.978)	(0.758)	(3.11)	(-4.40)	(-0.650)	(1.47)	(2.11)	(-3.70)	(0.743)	(2.97)

These are all the returns and profits from various strategies using two, four, and twenty six week lags, that is $k = 2, 4$, and 26. The reader should realize that the various portfolios are shorthands for the strategies, thus $R > 0$ and $V > V_m$ indicates the portfolio of stocks for which the individual stock return was greater than zero and the individual volume was greater than market volume at $t - k$. The other portfolios are as discussed in the section 3.3 of chapter 3. The $R > 0$ portfolio indicates the combined profits from the portfolios with positive return in period $t-k$, yet still using the weights as defined by each strategy. Similarly, the $R < 0$ portfolio indicates the combined profits from the portfolios with negative return in period $t - k$. The $V > V_m$ portfolio is the combined profits from the high volume stocks in period $t - k$ using the weights described by the strategy. The $V < V_m$ portfolio is the combined profits from the low volume stocks in period $t - k$ using the weights described by the respective strategy. Total profit is the sum of returns of the four portfolios of each strategy; it is a costless portfolio. It represents tr strategy. It is a costless portfolio, and it represents true profits. The contrarian profits indicate the profits from using the same strategies as outlined in section 3.3 however, we change them so as to sell winner stocks ($R > 0$) and buy loser stocks ($R < 0$); this is also a costless portfolio. The t statistics presented in the parentheses are Newey-West corrected using four lags. The sample consists of 1559 weeks. The returns and profits presented have been multiplied by 1000.

Table G.4: Mean Profits of All Four Strategies For Lags $k = 2, 4$, and 26: Medium Quintile

Lag	$R>0$ $V>V_m$	$R>0$ $V<V_m$	$R<0$ $V>V_m$	$R<0$ $V<V_m$	$R>0$	$R<0$	$V>V_m$	$V<V_m$	Total Profit	Contrarian Profit
Panel A: Using the Volume Lead Strategy										
2	-1.26	-1.30	8.74	-2.67	-2.56	6.07	7.48	-3.98	3.51	11.46
	(-0.632)	(-1.06)	(4.31)	(-2.25)	(-1.39)	(3.64)	(3.25)	(-4.53)	(1.42)	(4.66)
4	-2.62	-0.156	13.82	-6.45	-2.77	7.37	11.20	-6.61	4.60	17.81
	(-0.652)	(-0.073)	(3.15)	(-2.68)	(-0.800)	(1.98)	(2.27)	(-4.43)	(0.897)	(3.45)
26	-6.42	-0.992	15.48	-40.16	-7.41	-24.69	9.06	-41.16	-32.10	50.22
	(-0.253)	(-0.089)	(0.727)	(-2.91)	(-0.341)	(-1.46)	(0.313)	(-4.04)	(-1.19)	(1.48)
Panel B: Using Volume as an Indicator and Returns as Weights										
2	-1.67	-2.62	7.43	-6.03	-4.29	1.40	5.76	-8.65	-2.89	14.42
	(-0.922)	(-1.82)	(3.64)	(-4.26)	(-2.40)	(0.782)	(2.67)	(-7.15)	(-1.18)	(5.78)
4	-3.10	-2.43	10.13	-10.77	-5.53	-0.643	7.03	-13.20	-6.17	20.23
	(-0.862)	(-1.04)	(2.43)	(-3.25)	(-1.78)	(-0.169)	(1.48)	(-5.39)	(-1.23)	(3.57)
26	-13.79	4.06	35.86	-57.89	-9.73	-22.03	22.07	-53.83	-31.76	75.90
	(-0.683)	(0.332)	(1.46)	(-3.65)	(-0.552)	(-1.13)	(0.785)	(-3.70)	(-0.575)	(2.12)
Panel C: Using the Volume Lead Strategy Combined with Return Information (VLR1)										
2	-1.21	-2.59	8.69	-4.60	-3.80	4.09	7.48	-7.18	0.295	14.66
	(-0.527)	(-1.67)	(3.82)	(-3.32)	(-1.57)	(1.90)	(2.52)	(-5.12)	(0.053)	(4.48)
4	-1.31	-2.77	12.60	-8.17	-4.08	4.44	11.29	-10.94	0.355	22.23
	(-0.278)	(-1.23)	(2.52)	(-2.90)	(-0.928)	(0.968)	(1.78)	(-5.40)	(0.053)	(3.36)
26	-7.96	0.665	20.24	-51.55	-7.30	-31.31	12.28	-50.88	-3.86	63.17
	(-0.317)	(0.051)	(0.811)	(-3.23)	(-0.342)	(-1.56)	(0.377)	(-3.59)	(-1.25)	(1.60)
Panel D: Using the Volume Lead Strategy Combined with Return Information (VLR2)										
2	-1.66	-2.62	7.43	6.45	-4.28	13.88	5.77	3.83	9.60	14.43
	(-0.917)	(-1.82)	(3.64)	(0.521)	(-2.39)	(1.09)	(2.67)	(0.306)	(0.753)	(5.78)
4	-3.06	-2.43	10.14	-10.77	-5.49	-0.637	7.07	-13.20	-6.13	20.27
	(-0.851)	(-1.04)	(2.43)	(-3.25)	(-1.76)	(-0.168)	(1.48)	(-5.39)	(-1.22)	(3.57)
26	-13.81	4.07	35.83	-5.79	-9.74	-22.07	22.02	-53.83	-31.81	75.85
	(-0.683)	(0.332)	(1.46)	(-3.65)	(-0.552)	(-1.13)	(0.782)	(-3.70)	(-1.18)	(2.12)

These are all the returns and profits from various strategies using two, four, and twenty six week lags, that is $k = 2, 4$, and 26. The reader should realize that the various portfolios are shorthands for the strategies, thus $R > 0$ and $V > V_m$ indicates the portfolio of stocks for which the individual stock return was greater than zero and the individual volume was greater than market volume at $t - k$. The other portfolios are as discussed in the section 3.3 of chapter 3. The $R > 0$ portfolio indicates the combined profits from the portfolios with positive return in period $t-k$, yet still using the weights as defined by each strategy. Similarly, the $R < 0$ portfolio indicates the combined profits from the portfolios with negative return in period $t - k$. The $V > V_m$ portfolio is the combined profits from the high volume stocks in period $t - k$ using the weights described by the strategy. The $V < V_m$ portfolio is the combined profits from the low volume stocks in period $t - k$ using the weights described by the respective strategy. Total profit is the sum of returns of the four portfolios of each strategy. It is a costless portfolio, and it represents true profits. The contrarian profits indicate the profits from using the same strategies as outlined in section 3.3 however, we change them so as to sell winner stocks ($R > 0$) and buy loser stocks ($R < 0$); this is also a costless portfolio. The t statistics presented in the parentheses are Newey-West corrected using four lags. The sample consists of 1559 weeks. The returns and profits presented have been multiplied by 1000.

Table G.5: Mean Profits of All Four Strategies For Lags k = 2, 4, and 26: Smallest Quintile

Lag	$R>0$ $V>V_m$	$R>0$ $V<V_m$	$R<0$ $V>V_m$	$R<0$ $V<V_m$	$R>0$	$R<0$	$V>V_m$	$V<V_m$	Total Profit	Contrarian Profit
Panel A: Using the Volume Lead Strategy										
2	-0.944	-2.78	7.53	-7.37	-3.72	10.16	16.58	-10.14	6.44	26.73
	(-0.292)	(-1.66)	(4.79)	(-3.85)	(-1.28)	(3.21)	(4.08)	(-9.73)	(1.56)	(6.25)
4	1.49	1.44	29.14	-11.60	2.93	17.54	30.63	-10.16	20.48	40.79
	(0.286)	(0.434)	(2.65)	(-3.12)	(0.725)	(1.72)	(2.95)	(-5.35)	(1.92)	(3.90)
26	6.20	47.78	34.84	-58.48	53.98	-23.64	41.04	-10.70	30.34	51.74
	(0.191)	(2.43)	(1.02)	(-2.71)	(2.83)	(-1.02)	(1.33)	(-0.683)	(0.959)	(1.39)
Panel B: Using Volume as an Indicator and Returns as Weights										
2	-1.63	-6.88	17.53	-17.96	-8.51	-0.428	15.90	-24.84	-8.94	40.74
	(-0.186)	(-8.41)	(6.71)	(-4.90)	(-0.937)	(-0.123)	(1.82)	(-7.12)	(-0.943)	(4.35)
4	1.95	-3.66	34.15	-19.27	-1.71	14.88	36.10	-22.93	13.17	59.03
	(0.390)	(-0.916)	(3.47)	(-4.45)	(-0.433)	(1.71)	(3.94)	(-6.86)	(1.41)	(5.81)
26	10.15	85.65	55.40	-67.91	95.80	-12.51	65.55	17.74	83.29	47.80
	(0.318)	(2.15)	(1.41)	(-2.57)	(2.96)	(-0.502)	(1.78)	(0.457)	(2.26)	(0.72)
Panel C: Using the Volume Lead Strategy Combined with Return Information (VLR1)										
2	-3.56	-7.13	18.42	-17.37	-10.68	1.47	14.86	-24.50	-9.64	39.36
	(-0.427)	(-3.37)	(6.23)	(-4.40)	(-1.24)	(0.255)	(1.80)	(-6.50)	(-1.08)	(4.28)
4	4.27	-4.69	39.27	-16.48	-0.422	22.79	43.53	-21.17	22.36	64.70
	(0.760)	(-1.22)	(2.67)	(-3.74)	(-0.089)	(1.64)	(3.04)	(-6.54)	(1.55)	(4.31)
26	16.12	48.58	33.03	-67.86	64.70	-34.84	49.14	-19.28	29.86	68.42
	(0.474)	(2.13)	(0.857)	(-2.47)	(2.52)	(-1.13)	(1.23)	(-0.908)	(0.709)	(1.42)
Panel D: Using the Volume Lead Strategy Combined with Return Information (VLR2)										
2	-1.63	-6.88	17.55	-17.96	-8.51	-0.047	15.92	-24.84	-8.92	40.76
	(-0.187)	(-3.15)	(6.72)	(-4.90)	(-0.938)	(-0.117)	(1.82)	(-7.12)	(-0.942)	(4.35)
4	1.99	-3.66	34.27	53.44	-1.67	87.71	36.25	49.78	86.03	59.19
	(0.099)	(-0.915)	(3.47)	(0.744)	(-0.423)	(1.22)	(3.94)	(0.689)	(1.20)	(5.80)
26	10.22	85.66	55.31	-67.90	95.88	-12.59	65.54	17.76	83.29	47.78
	(0.320)	(2.15)	(1.40)	(-2.57)	(2.96)	(-0.506)	(1.78)	(0.458)	(2.26)	(0.72)

These are all the returns and profits from various strategies using two, four, and twenty six week lags, that is $k = 2, 4$, and 26. The reader should realize that the various portfolios are shorthands for the strategies, thus $R > 0$ and $V > V_m$ indicates the portfolio of stocks for which the individual stock return was greater than zero and the individual volume was greater than market volume at $t-k$. The other portfolios are as discussed in the section 3.3 of chapter 3. The $R > 0$ portfolio indicates the combined profits from the portfolios with positive return in period $t-k$, yet still using the weights as defined by each strategy. Similarly, the $R < 0$ portfolio indicates the combined profits from the portfolios with negative return in period $t-k$. The $V > V_m$ portfolio is the combined profits from the high volume stocks in period $t-k$ using the weights described by the strategy. The $V < V_m$ portfolio is the combined profits from the low volume stocks in period $t-k$ using the weights described by the respective strategy. Total profit is the sum of returns of the four portfolios of each strategy. It is a costless portfolio, and it represents true profits. The contrarian profits indicate the profits from using the same strategies as outlined in section 3.3 however, we change them so as to sell winner stocks ($R > 0$) and buy loser stocks ($R < 0$); this is also a costless portfolio. The t statistics presented in the parentheses are Newey-West corrected using four lags. The sample consists of 1559 weeks. The returns and profits presented have been multiplied by 1000.

Table G.6: Mean Absolute Weight Invested in Each Stock for All Four Strategies: All Stocks

Portfolio	$R > 0$ $V > V_m$	$R > 0$ $V < V_m$	$R < 0$ $V > V_m$	$R < 0$ $V < V_m$	$R > 0$	$R < 0$	$V > V_m$	$V < V_m$	Global
				Panel A: For All Strategies					
All Stocks	0.027	0.009	0.026	0.006	0.012	0.009	0.020	0.006	0.017
	(0.033)	(0.009)	(0.025)	(0.003)	(0.011)	(0.004)	(0.009)	(0.001)	(0.023)
Smallest	0.130	0.049	0.123	0.025	0.065	0.039	0.095	0.029	0.082
	(0.133)	(0.053)	(0.108)	(0.008)	(0.048)	(0.013)	(0.026)	(0.003)	(0.101)
Medium	0.144	0.045	0.157	0.031	0.060	0.046	0.110	0.029	0.094
	(0.155)	(0.055)	(0.178)	(0.023)	(0.052)	(0.030)	(0.070)	(0.004)	(0.134)
Largest	0.205	0.057	0.198	0.044	0.071	0.056	0.164	0.031	0.126
	(0.258)	(0.100)	(0.254)	(0.063)	(0.099)	(0.057)	(0.191)	(0.010)	(0.205)

These weights are for the calculations using a one week lag, that is $k = 1$. The weights are defined as the mean investment per stock in each portfolio as a fraction of the total dollar investment. The reader should realize that the various portfolios are shorthands for the strategies, thus $R > 0$ and $V > V_m$ indicates the portfolio of stocks for which the individual stock return was greater than zero and the individual volume was greater than market volume at $t - k$. The other portfolios are as discussed in section 3.3 of chapter 3. The $R > 0$ portfolio indicates the combined portfolio from the portfolios with positive return in period $t - k$. Similarly, the $R < 0$ portfolio indicates the combined portfolio from the portfolios with negative return in period $t - k$. The $V > V_m$ portfolio indicates the combined portfolio from the high volume stock portfolios in period $t - k$. Similarly, the $V < V_m$ portfolio indicates the combined portfolio from the low volume stock portfolios in period $t - k$. The standard errors are presented in parentheses. The sample consists of 1559 weeks.

Table G.7: Mean Absolute Weight Invested in Each Stock for All Four Strategies: Different Lags

Lag	Portfolio	$R>0$ $V>V_m$	$R>0$ $V<V_m$	$R<0$ $V>V_m$	$R<0$ $V<V_m$	$R>0$	$R<0$	$V>V_m$	$V<V_m$	Global
$k=2$	All Stocks	0.027	0.009	0.029	0.006	0.012	0.010	0.020	0.006	0.018
		(0.047)	(0.008)	(0.030)	(0.004)	(0.011)	(0.006)	(0.009)	(0.001)	(0.030)
	Smallest	0.125	0.049	0.129	0.027	0.066	0.043	0.094	0.029	0.085
		(0.127)	(0.053)	(0.116)	(0.013)	(0.061)	(0.021)	(0.026)	(0.003)	(0.049)
	Medium	0.139	0.046	0.171	0.034	0.061	0.050	0.109	0.029	0.098
		(0.150)	(0.065)	(0.203)	(0.026)	(0.061)	(0.033)	(0.069)	(0.004)	(0.046)
	Largest	0.196	0.055	0.194	0.047	0.067	0.060	0.163	0.031	0.124
		(0.243)	(0.089)	(0.256)	(0.069)	(0.086)	(0.077)	(0.188)	(0.010)	(0.200)
$k=4$	All Stocks	0.029	0.010	0.031	0.007	0.014	0.011	0.020	0.006	0.019
		(0.073)	(0.019)	(0.030)	(0.006)	(0.025)	(0.008)	(0.009)	(0.001)	(0.042)
	Smallest	0.131	0.053	0.148	0.030	0.068	0.049	0.094	0.028	0.090
		(0.146)	(0.076)	(0.149)	(0.020)	(0.075)	(0.037)	(0.027)	(0.003)	(0.122)
	Medium	0.131	0.049	0.191	0.042	0.062	0.057	0.109	0.029	0.103
		(0.136)	(0.092)	(0.234)	(0.068)	(0.077)	(0.052)	(0.073)	(0.004)	(0.159)
	Largest	0.186	0.057	0.212	0.052	0.065	0.067	0.168	0.031	0.127
		(0.237)	(0.105)	(0.277)	(0.084)	(0.083)	(0.096)	(0.203)	(0.010)	(0.207)
$k=26$	All Stocks	0.027	0.009	0.055	0.009	0.013	0.014	0.021	0.006	0.025
		(0.034)	(0.011)	(0.137)	(0.013)	(0.015)	(0.021)	(0.010)	(0.001)	(0.074)
	Smallest	0.135	0.073	0.149	0.051	0.087	0.067	0.088	0.028	0.102
		(0.184)	(0.145)	(0.185)	(0.078)	(0.144)	(0.076)	(0.029)	(0.005)	(0.160)
	Medium	0.150	0.046	0.208	0.061	0.065	0.080	0.109	0.028	0.116
		(0.193)	(0.059)	(0.248)	(0.134)	(0.083)	(0.141)	(0.064)	(0.005)	(0.185)
	Largest	0.199	0.039	0.205	0.058	0.051	0.072	0.192	0.029	0.125
		(0.268)	(0.044)	(0.234)	(0.079)	(0.043)	(0.081)	(0.236)	(0.009)	(0.199)

These weights are for the calculations using two, four, and twenty six week lags, that is $k = 2, 4$, and 26. The weights are defined as the mean investment per stock in each portfolio as a fraction of the total dollar investment. The reader should realize that the various portfolios are shorthands for the strategies, thus $R > 0$ and $V > V_m$ indicates the portfolio of stocks for which the individual stock return was greater than zero and the individual volume was greater than market volume at $t - k$. The other portfolios are as discussed in section 3.3 of chapter 3. The $R > 0$ portfolio indicates the combined portfolio from the portfolios with positive return in period $t - k$. Similarly, the $R < 0$ portfolio indicates the combined portfolio from the portfolios with negative return in period $t-k$. The $V > V_m$ portfolio indicates the combined portfolio from the high volume stock portfolios in period $t-k$. Similarly, the $V < V_m$ portfolio indicates the combined portfolio from the low volume stock portfolios in period $t - k$. The standard errors are presented in parentheses. The sample consists of 1559 weeks.

Table G.8: Mean Number of Stocks in Each Portfolio for All Four Strategies

Portfolio	$R > 0$ $V > V_m$	$R > 0$ $V < V_m$	$R < 0$ $V > V_m$	$R < 0$ $V < V_m$	$R > 0$	$R < 0$	$V > V_m$	$V < V_m$
Panel A: For All Strategies								
All Stocks	58	155	57	201	213	258	115	356
	(32)	(63)	(33)	(66)	(84)	(84)	(39)	(41)
Smallest	11	27	11	44	38	55	22	70
	(6)	(11)	(6)	(11)	(14)	(14)	(6)	(6)
Medium	11	32	11	39	43	51	23	71
	(7)	(14)	(7)	(15)	(18)	(18)	(9)	(10)
Largest	12	32	12	36	45	48	24	68
	(12)	(18)	(11)	(18)	(21)	(21)	(17)	(17)

These mean numbers are for the calculations using a one week lag, that is k=1. The numbers are similar for other lags, given that we use non-overlapping lags. The reader should realize that the various portfolios are shorthands for the strategies, thus $R > 0$ and $V > V_m$ indicates the portfolio of stocks for which the individual stock return was greater than zero and the individual volume was greater than market volume at $t - k$. The other portfolios are as discussed in section 3.3 of chapter 3. The $R > 0$ portfolio indicates the combined portfolio from the portfolios with positive return in period $t - k$. Similarly, the $R < 0$ portfolio indicates the combined portfolio from the portfolios with negative return in period $t - k$. The $V > V_m$ portfolio indicates the combined portfolio from the high volume stock portfolios in period $t - k$. Similarly, the $V < V_m$ portfolio indicates the combined portfolio from the low volume stock portfolios in period $t - k$. The standard errors are presented in parentheses. The sample consists of 1559 weeks. The mean numbers of stocks in each portfolio is independent of the weighting strategy given that positive or negative return and the volume lead are the ones that separate the portfolios. The mean number of stocks and standard deviations were rounded to the nearest integer.

Table G.9: The Profits from Various Studies of Contrarian Strategies Using Different Weighting Schemes

Portfolio	Lehmann (1990) $u_{it}(k) = \frac{R_{it-k} - \bar{R}_{t-k}}{\sum(R_{it-k} - \bar{R}_{t-k})}$	Lo and MacKinley (1990) $u_{it}(k) = -\frac{1}{N}(R_{it-k} - R_{mt-k})$	Conrad *et al.* (1994) $u_{it}(k) = \frac{R_{it-k}}{\sum R_{it-k}}$	Chincarini/Llorente (1995) $u_{it}(k) = \frac{R_{it-k}}{\sum R_{it-k}}$
All Stocks	17.9	0.169	1.16	24.87
	(41.07)	(20.81)	(1.60)	(14.74)
Smallest	10.6	0.453	n.a.	40.73
	n.a.	(18.81)		(12.97)
Medium	2.1	0.106	n.a.	13.57
	n.a.	(13.84)		(10.08)
Largest	0.4	0.062	n.a.	9.81
	n.a.	(11.22)		(8.04)

These are all the profits from various strategies using a one week lag, that is $k = 1$. These portolios are all costless portfolios. The profits are multiplied by 1000. n.a. indicates that the data was not available.

Bibliography

[1] *County and City Databook 1988.* Data User Services Division, Bureau of the Census; Washington D.C., 1989.

[2] "Air Pollution: Wheeze". *The Economist,* pages 52–64, January 1991.

[3] *Climate Normals for the U.S. (base, 1951-1980).* Detroit, Michigan : Gale Research Co., 1991.

[4] *Census of Population and Housing: Public Use Microdata Samples U.S.* Bureau of the Census; Washington D.C., 1992.

[5] *Business Control Atlas.* American Map Corporation: Langensheidt Publishing, 1993.

[6] "Drugs and Crime: The Wage of Crack". *The Economist,* page 29, September 1994.

[7] "London. Yet It Moves". *The Economist,* September 1994.

[8] "Clean Air: To Some Dirty Words". *The USA Today,* page 8A, March 1995.

[9] ABBOTT, LAURIE. "Sun, Surf, and Smog: California Fights Back". *American City and County,* pages 57–58, March 1990.

[10] AMEMIYA, TAKESHI. *Advanced Econometrics.* Harvard University Press, 1985.

[11] BEN-AKIVA, MOSHE AND STEVE R. LERMAN. *Discrete Choice Analysis: Theory and Applications to Travel Demand.* The MIT Press, 1985.

[12] BERGER, MARK C. AND GLENN C. BLOMQUIST. "Mobility and Destination in Migration Decisions: The Roles of Earnings, Quality of Life, and Housing Prices". *Journal of Housing Economics,* pages 37–59, September 1992.

[13] BIGGAR, C., CHARLES F. LONGINO, JR., AND DALE E. YEATTS. "Distance Versus Destination: Stream Selectivity of Elderly Interstate Migrants". *Journal of Gerontology,* pages 288–294, March 1987.

[14] BLACK, FISCHER. "Noise". *The Journal of Finance,* pages 529–543, July 1986.

[15] BLUME, L., D. EASLEY, AND M. O'HARA. "Market Statistics and Technical Analysis: The Role of Volume". *Journal of Finance*, pages 153–181, March 1994.

[16] BOEHM, THOMAS P., HENRY W. HERZOG, JR., AND ALAN M. SCHLOTTMAN. "Intra-Urban Mobility, Migration, and Tenure Choice". *The Review of Economics and Statistics*, September 1991.

[17] CAMPBELL, J., S. GROSSMAN AND J. WANG. "Trading Volume and Serial Correlation in Stock Returns". *Quarterly Journal of Economics*, pages 905–940, November 1993.

[18] CEBULA, RICHARD J. "The Impact of Living Costs on Geographic Mobility". *The Quarterly Review of Economics and Finance*, pages 101–105, September 1993.

[19] CHINCARINI, LUDWIG B. *"The Dynamics of City Formation"*. Ph.D. dissertation: Chapter 1, Massachusetts Institute of Technology, Department of Economics, June 1995.

[20] CHOPRA, NAVIN, JOSEF LAKONISHOK, AND JAY R. RITTER. "Measuring Abnormal Performance: *Do Stocks Overreact?*". *Journal of Financial Economics*, pages 235–268, September 1992.

[21] CHRISTALLER, W. *Die Zentralen Orte in Suddeutschland.* Jena: Fischer, 1933.

[22] CHRISTALLER, W. *Central Places in Southern Germany.* Englewood Cliffs, NJ: Prentice-Hall. Translation by C.W. Baskin, 1966.

[23] CONRAD, J., A. HAMEED, AND C. NIDEN. "Volume and Autocovariances in Short-Horizon Individual Security Returns". *Journal of Finance*, pages 1305–1329, September 1994.

[24] CREADY, W.M., AND P.G. MYNATT. "The Information Content of Annual Reports: A Price and Trading Response Analysis". *The Accounting Review*, pages 291–312, September 1991.

[25] DEBONDT, WERNER F.M., AND RICHARD THALER. "Does the Stock Market Overreact?". *The Journal of Finance*, pages 793–807, January 1986.

[26] FOSTER, WILLIAM S. "The Misconceptions of Urban Concerns". *American City and County*, pages 44–54, September 1984.

[27] FOX, WILLIAM F., HENRY W. HERZOG, JR., AND ALAN M. SCHLOTTMAN. "Metropolitan Fiscal Structure and Migration". *Journal of Regional Science*, pages 523–536, November 1989.

[28] FREES, EDWARD W. "Forecasting State-to-State Migration Rates". *Journal of Business and Economic Statistics*, pages 153–167, April 1992.

[29] FRIEDRICH, K. AND R. KOCH. "Federal Republic of Germany". In A. Rogers and W. Serow, editor, *Elderly Migration: An International Comparative Study.*

[30] GALLANT, A.R., P.E. ROSSI AND G. TAUCHEN. "Stock Prices and Volume". *Review of Financial Studies*, pages 199–242, September 1992.

[31] GARREAU, JOEL. *Edge City.* Doubleday, 1988.

[32] GLAESER, EDWARD, HEDI D. KALLAL, JOSE A. SCHEINKMANN, AND ANDREI SCHLEIFER. "Growth in Cities". *Journal of Political Economy*, pages 1126–1152, September 1992.

[33] GOLANT, STEPHEN M. "Post-1980 Regional Migration Patterns of the U.S. Elderly Population". *Journal of Gerontology*, pages S135–S140, September 1990.

[34] GRADSHTEIN, I.S., AND I.M. RHZHIK. *Tables of Integrals, Series, and Products.* Academic Press, New York, New York, 1980.

[35] GREENWOOD, MICHEAL J. "Research on Internal Migration in the United States: A Survey". *Journal of Economic Literature*, pages 397–433, September 1975.

[36] GROSSMAN, S. AND J.E. STIGLITZ. "Information and Competitive Price Systems". *American Economic Review*, pages 246–253, September 1976.

[37] GROSSMAN, S. AND J.E. STIGLITZ. "On the Impossibility of Informationally Efficient Markets". *American Economic Review*, pages 393–408, September 1980.

[38] HARRIS, M. AND A. RAVIV. "Differences of Opinion Make a Horse Race". *Review of Financial Studies*, pages 473–506, September 1993.

[39] HE, H. AND J. WANG. "Differential Information and Dynamic Behavior of Stock Trading Volume". *Working Paper*, November 1994.

[40] HELLER, DANIEL. "A Spatial Model with Local Goods". *Stanford Working paper*, September 1993.

[41] HSIAO C. AND K. SMALL. "Multinomial Logit Specification Tests". *Working Paper, Department of Economics, Princeton University*, September 1982.

[42] JACOBS, JANE. *The Economy of Cities.* Random House Publishers, New York, New York, 1969.

[43] KAHLEY, WILLIAM J. "Population Migration in the United States: A Survey of Research". *Economic Review of the Kansas City Federal Reserve Bank*, September 1991.

[44] KANDEL, E. AND N. D. PEARSON. "Differential Interpretation of Public Signals and Trade in Speculative Markets". *Working paper, University of Rochester*, September 1992.

[45] KARPOFF, J. M. "The Relation Between Price Changes and Trading Volume: A Survey". *Journal of Financial and Quantitative Analysis*, pages 109–126, September 1987.

[46] KNOX, PAUL AND AGNEW, JOHN. *The Geography of the World Economy.* Edward Arnold, New York, New York, 1994.

[47] KRUGMAN, PAUL. "A Dynamic Spatial Model". *NBER Working Paper*, September 1991.

[48] KRUGMAN, PAUL. "Fluctuations, Instability, and Agglomeration". *NBER Working Paper*, January 1994.

[49] KYLE, A. "Continuous Auctions and Insider Trading". *Econometrica*, pages 1315–1335, September 1985.

[50] LEBARON, B. "Persistence of the Dow Jones Index on Rising Volume". *Working paper, University of Wisconsin*, September 1991.

[51] LEHMANN, B. "Fads, Martingales, and Market Efficiency". *Quarterly Journal of Economics*, pages 1–28, September 1990.

[52] LIN-YUAN, YIHUA AND LESZEK A. KOSINSKI. "The Model of Place Utility Revisited". *International Migration*, September 1994.

[53] LITWAK E. AND C. LONGINO. "Migration Patterns Among the Elderly: A Developmental Perspective". *Gerontologist*, pages 266–272, September 1987.

[54] LLORENTE, J. GUILLERMO. "Autocorrelation in Returns and Trading Volume on Individual Stocks". *MIT, mimeo*, March 1995.

[55] LO, A.W. AND A.C. MACKINLAY. "When are Contrarian Profits Due to Stock Market Overreaction?". *Review of Financial Studies*, pages 175–206, September 1990.

[56] LONGINO, CHARLES F. AND KENNETH J. SMITH. "Black Retirement Migration in the United States". *Journal of Gerontology*, pages S125–S132, September 1991.

[57] LÖSCH, AUGUST. *The Economics of Location.* New Haven: Yale University Press. Translation by W.H. Woglom., 1954.

[58] MADDALA, G.S. *Limited Dependent and Qualitative Variables in Econometrics.* Cambridge University Press, 1983.

[59] MCMAHON, W.W. "Geographic Cost of Living Differences: An Update". *American Real Estate and Urban Economics Association Journal*, pages 426–450, February 1991.

[60] MEYER, JUDITH W. "A Regional Scale Temporal Analysis of the Net Migration Patterns of Elderly Persons Over Time". *Journal of Gerontology*, pages 366–375, September 1987.

[61] MONTGOMERY, EDWARD. "Evidence on Metropolitan Wage Differences across Industries and over Time". *Journal of Urban Economics*, pages 69–83, January 1992.

[62] MORSE, D. "Asymmetrical Information in Securities Markets and Trading Volume". *Journal of Financial and Quantitative Analysis*, pages 1129–1146, September 1980.

[63] PADOA-SCHIOPPA, F. *Mismatch and Labor Mobility.* 1990.

[64] PFLEIDERER, P. "The Volume of Trade and The Variability of Prices: A Framework for Analysis in Noisy Rational Expectations Equilibria". *Working Paper, G.S.B. Stanford University*, September 1984.

[65] PIEROTTI, PIERO. *Urbanistica: Storia e Prassi.* Firenze: Marchi & Bertolli, 1972.

[66] RING, M.L. "Elderly Migration: An International Comparative Analysis". In A. Rogers and W. Serow, editor, *Elderly Migration: An International Comparative Study.*

[67] SEROW, W. "Why the Elderly Move: International Comparisons". *Research on Aging*, pages 582–597, September 1987.

[68] SEROW, WILLIAM J. "Determinants of Interstate Migration: Differences Between Elderly and Nonelderly Movers". *Journal of Gerontology*, pages 95–100, September 1987.

[69] SEROW, WILLIAM J. AND DAVID F. SLY. "Geographic Mobility of the Elderly in Industrialized Societies". In Wolfgang Lutz, editor, *Future Demographic Trends in Europe and North America*, chapter 20, pages 399–419. Academic Press, San Diego, first edition, September 1991.

[70] SMITH, KELLY, MELANIE MAVRIDES, JANET REYES, AMI WALSH, AND LESLIE WHITAKER. "The Best Places to Live in America". *Money*, September 1994.

[71] SPILIMBERGO, ANTONIO AND LUIS UBEDA. "Why Do Europeans Move So Little?". *Mimeo, M.I.T.*, 1993.

[72] STEIN, JEREMY. "Overreactions in the Options Market". *The Journal of Finance*, pages 1011–1022, September 1989.

[73] STICKEL, S.E. AND R.E. VERRECCHIA . "Price Behavior when Expectations are Conditioned over Trading Volume". *Working Paper, Wharton School and La Salle University*, September 1992.

[74] STRUMPF, KOLMAN. *"Credibility and the Structure of Taxes: The Local Earned Income Tax in Pennsylvania"*. Ph.D. dissertation: Chapter 1, Massachusetts Institute of Technology, Department of Economics, June 1995.

[75] SZYMANSKI, R. AND AGNEW, J.A. *Order and Skepticism: Human Geography and the Dialetic of Science.* Washington D.C: Association of American Geographers, 1981.

[76] TAGTOW, RICK. "The Need for Urban Forests". *American City and County*, pages 74–75, March 1990.

[77] TRAIN, KENNETH. *Qualitative Choice Analysis.* The MIT Press, 1986.

[78] VARIAN, HAL R. "Divergence of Opinion in Complete Markets". *Journal of Finance*, pages 309–317, September 1985.

[79] VARIAN, HAL R. "Differences in Opinion in Financial Markets". In C.C. Stone, editor, *Financial Risk: Theory, Evidence and Implications: Proceedings of the* 11^{th} *Annual Economic Policy Conference of the Federal Reserve Bank of St. Louis*, chapter 33, pages 399–419. Academic Press, first edition, September 1989.

[80] VON THÜNEN, J.K. *Der Isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie.* Hamburg, 1826.

[81] WANG, J. "A Model of Competitive Stock Trading Volume". *Working paper, Sloan School of Management, M.I.T*, September 1992.

[82] WARD, JANET AND LINDA PARHAM. "Cleaning the Air". *American City and County*, March 1990.

[83] ZAX, JEFFREY S. "The Substitution Between Move and Quits". *The Economic Journal*, pages 1510–1521, November 1991.

[84] ZIPF, GEORGE. *Human Behavior and the Principle of Least Effort.* Addison-Wesley Press, Inc., Cambridge, MA., 1949.