

ACADEMY of
ECONOMICS
and FINANCE
JOURNAL

Volume 1	Table of Contents	2010
Knowledge in Economic Science		1
<i>George Carter</i> <i>(Lecture by the 2010 Fellow of the Academy of Economics and Finance, Presented at the 2010 Meetings of the Academy of Economics and Finance)</i>		
Dollar Demise or Rise: The Long and the Short of It		11
<i>Shuming Bai and Oviedo-Toral Gerardo</i>		
U.S. Mortgage Foreclosures: An Analysis by State		17
<i>William Brent, Lynne Kelly, Debbie Lindsey, and Russell Price</i>		
A Panel Model of the Interest Sensitivity of Net Interest Margins: 1984—2009		31
<i>Albert E. DePrince, Jr. and Pamela D. Morris</i>		
Testing the Modigliani-Miller Risk-Class Property		39
<i>David P. Echevarria</i>		
A Nonparametric Kernel Method Test of the Symmetry of Security Returns		47
<i>Gary S. Moore, Qin Shao, and Haifeng Zhang</i>		
Socioeconomic Factors and Educational Attainment: Evidence from South Carolina		55
<i>Muhammad Mustafa and Rahul Singh</i>		
Student Performance in the Economics Classroom		61
<i>Richard Vogel and Wisdom Akpalu</i>		

Knowledge in Economic Science

George Carter, University of Southern Mississippi

Lecture by the 2010 Fellow of the Academy of Economics and Finance

Presented at the 2010 Meetings of the Academy of Economics and Finance

Abstract

Economic science is based on knowledge. Investigation of knowledge in its many aspects (philosophical, neurological, and empirical) reveals that knowledge is quite tentative. Thus, action based on economic science should be taken with great reservation and openness to alternatives.

Preamble

This paper is an unabashedly secular paper. Its purpose is to better understand knowledge as information that forms the basis of economic science. Theological answers to intractable questions eventually become “God’s will,” “Act of God,” or some other explanation that rests completely on faith. Faith has its place in human inquiry, including the question pursued here, but the discussion in this paper will be limited to secular explanation.

The parent discipline for all knowledge is philosophy. The role of philosophy has always been to address the intellectually unanswerable questions while the derivative disciplines, scientific and not, address questions for which philosophy has provided an intellectual roadmap to understanding. Thus, the sections of this paper devoted to philosophy inherently recognize lack of knowledge while the sections of this paper devoted to the derived disciplines recognize active development of conceptual and empirical knowledge.

Introduction

Income varies among the regions of the United States. One reason, of many, for that income variation, is the difference in the costs of living among the regions. How do we know that? Do we know these facts (regional income variation and a reason) to be absolutely true? Or, is there sufficient probability of truth in those statements that we can regard them as true for practical purposes? Or, does the truth of those statements border on guess, or even pure speculation? Humans live lives in which we continuously learn in order to better understand and decide. But, just how do we know what we know, and how does that knowledge provide a basis for science; in particular, economic science?

Knowing some fact is actually quite a bit more complicated than one would first imagine. By exploring what goes into “knowing” something, we will better understand those complications, and we will gain a better perspective on what we know. In examining that process, it becomes difficult to escape the conclusion that “knowledge” is extremely fragile.

The plan of this paper is to present traditional explanations of how we know something and how that knowledge provides a basis for scientific reasoning. The next section summarizes how philosophers have wrestled with this issue for at least twenty-seven centuries. An understanding of how the human brain acquires, processes and employs knowledge occupies a following section. After all, without human involvement, there is no knowledge. The following section addresses reality as understood in physics; reality being an essential ingredient for knowledge. Interestingly, studying the brain and physics leads us back into philosophy. The final two sections address issues in the scientific method and the philosophy of science.

Philosophy of Reality and Knowledge

The question, “How do we know?” has probably concerned us for as long as humans have been able to reason. Certainly, the question was posed in antiquity, and there are two ancient fields of philosophy that address the question: (1) ontology – theories on the nature of reality and (2) epistemology – theories of knowledge.

Ontology

What is “real?” This question started the field of philosophy, and must be addressed before we can “know” something. Thales of Miletus (a Greek colony in what is modern-day Turkey) earned the title, Father of Philosophy, by examining this question in the seventh century BCE. He concluded that reality is material substance that is made up of earth, wind, fire and

water. Of these four components, water is the most fundamental because water falls from the sky as rain, is absorbed into the earth and will condense with the temperature differential caused by fire.

Plato in the fourth century BCE, on the other hand, believed that reality is abstract, not material. To him, true reality is Forms that reside in your mind. Though you may believe that an object before you is real, it is not. What was real to Plato is the perfect conception in your mind that represents the object. He believed that abstract mathematics was more perfect than the mental image of a chair. Consequently, he developed an hierarchy from the inferior ordinary world of our senses to the exact world of intelligence like mathematics to the superior reality of the Forms that are perfect, unchanging and eternal. George Berkeley, in the eighteenth century CE, took Plato one step further by contending that the world about us has no reality apart from our perception of that world.

Philosophy, over its long history, provides the foundation that reality is material or ideal or somewhere in between. Materialism or idealism, this interpretation into extremes, with a range between, characterizes the kind of choices that we have to make in exploring this topic.

Epistemology

Having addressed “reality” as a range between two extremes of materialism and idealism, we now address how we “know.” In other words, how do we obtain “knowledge?” Again, the answer is a range between two extremes: rationalism and empiricism.

Plato extended his concept that reality is in the mind by explaining that reasoned thought (rationalism) led to knowledge. Rene Descartes, in the seventeenth century CE, expanded on Plato’s rationalism in leading the Continental Rationalists approach to knowledge. Through pure reason and strenuous doubt, Descartes determined that the only thing that he could know for sure is that he existed (*Cogito ergo sum*, I think, therefore I am.). Thus, only things that he perceived could be true, and the reasoning of his mind was the only dependable source of knowledge. The reaction to Descartes’ rationalism came from the British Empiricists.

John Locke, in the late seventeenth century, wrote that man is born with a blank slate for a mind. Experience from a material world writes knowledge onto that blank slate. David Hume, the ultimate skeptic in the rationalist-empiricist debate, rejected both. He argued that deductive reasoning, cause and effect, using the past to make claims about the future, is inherently flawed. The future is necessarily unknown, and any statements about the future can only be probabilistic. Inductive reasoning, on the other hand, empirically gains knowledge through the senses. But, any number of parlor games will convince a person that the senses can easily be fooled, and again knowledge is only probable. Apparently, he concluded, there is no true knowledge.

Immanuel Kant, in the eighteenth century, tried to reconcile the rationalist-empiricist debate by assuming that all minds have the same structure and interpret the real world identically. Because we are programmed to interpret the material world identically, we all accept “things as they appear to be” as true knowledge of “things as they are in themselves.” Then by accepting the legitimacy of causality and a causally-determined world, both rationalism and empiricism lead to knowledge.

The Human Brain

Philosophy sets an intellectual foundation for understanding how we “know” there is income variation among the regions, but it does not address in any convincing detail a fundamental component of our knowing anything: the human brain. We must understand the mechanics of how the human brain obtains, processes, decides and communicates knowledge.

Anatomy

The human brain is divided into two hemispheres, each with four lobes: the frontal (forward-most and evolutionary youngest), parietal (middle), occipital (hindmost) and temporal (on the side). The hemispheres have “gray matter” coats that are composed of more than 100 billion neurons. These neurons are connected at multiple places (synapses), and it is estimated that there are over 100 trillion synapses in the human brain. The interior of the brain is “white matter” that is composed of many long neural axons.

Learning

Learning is the process by which we acquire knowledge about the world, while memory is the process by which that knowledge is encoded, stored, and later retrieved. The physical representation of learning is a nucleus of neural networks. It is interesting to trace through how learning takes place from seeing something to writing about it (input-processing-output).

Input

A visual experience occurs when photons strike the rods and cones on the back and sides of the eye. There are 120 million rods and six million cones for collecting visual information. That information is transmitted to the visual cortex by only one million optic nerves and synaptically relayed through the thalamus before synaptic relay into cortical nuclei. Clearly, the amount and form of information reaching the visual cortex is substantially less and different than the information that entered the eye.

The visual cortex does not receive a photograph from the eye. Rather, the visual cortex creates a percept, or mental representation, of what was seen. On the basis of partial information, the brain constructs sight. In the best tradition of Plato's Forms, reality to humans is truly mental. And, this is equally true for other sensory inputs of hearing, touch, taste and smell. Basically, the primary visual cortex receives information on edges and contrast; then it distributes that information to other areas of the cortex for mental construction of what we see. Our mental image is not a faithful reproduction of the photon pattern that struck the eye. Rather, our mental image is an internal construction of what we perceive to be out there.

A long-lived but erroneous theory is the "inner theater." Common sense and superficial biological science reason that light enters the eyes and a mental picture is painted in the brain to be observed by homunculus (little man inside). Thus as we observe the world about us, we watch that reality in movie-like fashion. This has been a very appealing theory, and it is very common today. Empirical evidence, however, has shown that the brain constructs what we see. The brain's order of observations differs from reality with some observations occurring before their occurrence in reality. In addition, a theater screen in the brain is a specific location. Though not understood, it is clear that visual perception occurs simultaneously at wide-spread locations, not at a specific location.

Processing

There is general agreement that specific mental functions occur in specific brain areas. Visual information is initially received at the back of the occipital lobe (the primary visual cortex). That information is then distributed to the visual unimodal association cortex (front of the occipital lobe and rear of the temporal lobe) where a visual percept is constructed. Since action is expected, the visual percept must be integrated with memory and other sensory percepts, and a decision must be made. These higher-level activities take place in the multimodal association areas.

There are three multimodal association areas. The posterior multimodal association area is located where the occipital, parietal and temporal lobes come together. In this area, neural networks integrate the visual percept with other sensory percepts and language. The second multimodal association area is the limbic multimodal association area which is located at the front of the temporal lobe. Here, the integrated sensory percepts are further integrated with memory and emotion. In the third multimodal association area, the executive functions of planning, judgment and selection are conducted. This is the anterior multimodal association area that is located at the front of the frontal lobe.

Several comments are in order about this mental processing. First for pedagogical simplicity, the presentation above is presented as serial processing where one activity is completed before passing it on to another activity. The brain employs intuitive parallel processing. The primary, unimodal and multimodal areas work simultaneously in creating percepts and integrative results. As those percepts and results are constructed, the neural networks fill the gaps using memory and logic. Those filled gaps may be confirmed or denied as parallel-processed information is received. As noted above, much sensory information is not transmitted to the brain. So, many of the gaps remain solely neural constructs.

Second, emotion is a critical component of neural memory and decisions. There are common stereotypes of cold, calculating, emotionless decision-makers and emotional, flustered procrastinators. Somehow it is believed that emotion distorts judgment and memory. In fact, emotion is essential for the neural processes of memory and judgment, and any discussion of the limbic system is in the context of emotion.

Finally, consciousness will be addressed below. There essentially is no understanding of that mental state, but some believe that consciousness is based in the three multimodal association areas. Their complex integrative activities form the best biological basis for consciousness, as the brain is now understood.

Output

Taking action reverses the process described above. An anterior multimodal association area decision to pick up a pen to write is communicated to the premotor unimodal association cortex in the middle of the frontal lobe. There, this decision is condensed into required actions for the primary motor cortex at the back of the frontal lobe. The primary motor cortex then executes the required actions by firing neurons that travel through the spine to the fingers. As a result, muscles are electrically-stimulated, and the pen is picked up.

Carter: Knowledge in Economic Science

The whole process has a symmetry. Photons enter the eye. The primary visual cortex receives the information and distributes it to the visual unimodal association cortex for construction. The multimodal association areas combine that visual percept with other information and decide to pick up a pen. The premotor unimodal association cortex condenses that decision, and the primary motor cortex moves the fingers.

Memory

The physical representation of what we know is neural networks in specific functional areas of the brain. When we learn, we alter relevant neural networks.

Short-term learning occurs when a stimulus alters a neuron for an abbreviated period of time. There is an obvious resultant effect on other neurons and, therefore, on that particular neural network, a bit of knowledge. Learning has occurred because what we know (the specific neural network) has changed. The learning and memory are short-term because the initially-effected neuron, and the neural network, will revert to base rate of firing and neurotransmitter release after the stimulus, therefore losing the learning and memory.

Long-term learning occurs when a stimulus permanently alters a neural network. This requires repeated stimuli. DNA will then create a protein in the nucleus to permanently change the way that the cell works. The key to long-term learning, and therefore the permanent alteration of a neural network, is repetition. Spaced repetition, not mass repetition, is necessary to slowly build a changed neural network that represents new, long-term memory.

Consciousness

There is growing understanding of the biological basis for brain function. Though much is still unknown, neural scientists are confident that eventually the biological basis of the brain will be known, with one exception: consciousness. That mental activity escapes even biological definition. The self-awareness of humans does not seem to take place at a specific location or area as other mental activities do, and the wide-spread connections that characterize conscious mental activity have no known physical basis. There is so little known about conscious mental activity, and so few prospects for understanding it, that explaining conscious mental activity is known as the “hard problem” in philosophy of the mind. No explanation of “how we know” is complete without inclusion of conscious comprehension, yet that component of the human mind completely baffles neural science. Though neural scientists discuss consciousness, it essentially is a province of philosophy.

Physics

How we know something fundamentally rests in the biological world of the brain. The biological world, however, rests in the physical world so a quick orientation in physics helps since current philosophical discussions combine modern physical and biological knowledge.

Newtonian Physics

Newtonian physics is based on an assumption of extreme materialism. Reality is the physical world, and that world is governed by discoverable laws that describe cause and effect. In effect, the world is a clock. Once a cause is initiated, its effect and all the subsequent effects are determinable by the laws that govern their actions.

Quantum Physics

Early in the twentieth century, technology had advanced to the degree that subatomic observations could be made. As a result, the extreme materialism and determinism of Newtonian physics came under attack with observations such as one particle could be in two places at once, or a particle could be a material object at one and the some time that it was an immaterial energy wave, or a particle could directly and immediately influence another particle that was light years away. Culminating this attack was Heisenberg’s Uncertainty Principle. A material particle has position and velocity, but Heisenberg showed that at the subatomic level, if you know one of those two characteristics then the other is indeterminate.

The sum result is that we cannot know matter in its basic components with certainty. Rather, the reality of those components can be discussed only in probabilistic terms. That indeterminacy of physical reality and knowledge, combined with our incomplete knowledge of brain function and total lack of understanding of consciousness, brings us back into the philosophical realm.

Philosophy of the Mind

Scientific explanation of mental functioning stops short of explaining how brain activities lead to consciousness and self-awareness. There is belief that there is a biological basis for conscious thinking and reasoning, but no scientific, empirical work has found that basis. Thus, we have the mind-body problem that defines the philosophy of the mind.

Cartesian Dualism

Rene Descartes, in the seventeenth century, presented a theory that the universe is composed of two completely distinct things: a material physical realm composed of matter and an immaterial mental realm composed of thoughts. This Dualism, Cartesian Dualism after Descartes, has persisted and retains prominence today. The theory's weakness, then and now, is that the connection between mind and body is not given. Philosophers and scientists have struggled with dualism because conscious thought and physical action seem to be integrated activities rather than separate things that are somehow connected.

Dualism is so appealing that a number of hardly credulous attempts have been made to explain the illusive "connection." Thomas Huxley was a proponent of epiphenomenalism. That theory posits that physical events do not cause mental events. Likewise, mental events do not cause physical events. Rather, mental events just accompany physical events, but there is no connection. Nicolas de Malebranche supported occasionalism which held that there was no connection. God just made both happen simultaneously. Gottfried Leibniz preferred parallelism. As a coauthor of physics, with Newton, the clock-work universe explained physical reality. Naturally to him, a clock-work mental universe running in parallel with the physical universe made sense.

Monism

Monism holds that there is only one basic substance of the universe. There is no connection problem because there is only one thing, there is nothing to connect to. Idealism, the belief that mental is the only thing in the universe, has a long tradition, most notably Plato. Despite its persistence, a logical idealist must eventually reason that if the universe is all in the mind then it can only be in one mind, my mind. This is solipsism, and a dead end.

The other path of monism is materialism. The approach taken by neuroscientists is reductive materialism. Their belief is that the mental activities of thinking, dreaming, wishing, loving, etc. can be reduced to neural network and brain chemistry activities, in other words, brain states. While neural science has made tremendous strides and will continue to advance impressively, the idea that science will eventually be grossed up to include the mental realm seems logically flawed. Neural networks are unique. There is great variation in comparable networks of people. Many believe that it is not reasonable to assume that such varied physical structures can yield equivalent mental states. As an alternative to reductive materialism, Paul and Patricia Smith Churchland presented eliminative materialism. They contend that mental things do not exist, only physical things exist. Much as we understand that mythology of yore did not exist, so it is to them that future generations will understand that our mental states did not really exist.

Functionalism

Ludwig Wittgenstein changed the mind-body debate by pointing out that we cannot observe mental states. Rather, we observe behavior resulting from mental states. B.F. Skinner followed-up by applying behaviorism in psychology. Unfortunately, the connection between mental states and behaviors proved illusive, and behaviorism has fallen by the wayside. In its wake, functionalism has become the primary explanation in philosophy of the mind.

Functionalism takes the position that mental states are functional states of an entire organism. In introducing functionalism, Hilary Putnam used hardware and software of computers as a model. Mental states would be analogous to software that performs a particular function. We recognize a particular software by its outputs and other software it initiates. So it is with mental states. We recognize a mental state such as love by its resulting behaviors (e.g., distraction, particular attention to the subject of love and gift-giving). In addition, that mental state initiates other mental states, with resulting behaviors, such as memories, hopes, and plans.

The mind-body problem today has evolved from dualism to functionalism which can be interpreted as either monistic or dualistic. The mental state where the mind is a separate mental state is the dualistic version. The monistic version returns to behaviorist roots. Functionalism, in the monistic version, changes mental states to complexly interacted activity of the mind with observable material behaviors. If pressed, functionalists would probably consider themselves materialists. That distinction, however, belies their belief in the integrated mind-body nature of mental state functioning.

Consciousness

To this point, this discussion has addressed the “Easy Problem of the Mind.” This problem is explaining how the mind operates, and functionalism is the primary explanation. The problem is considered “easy” because there is widespread belief that there is a biological basis for the mind. Now, we must address what David Chalmers has identified as the “Hard Problem of Consciousness.” The problem is “hard” because there is essentially no scientific explanation for consciousness. Philosophy must explain the phenomenon of consciousness while real-world research cannot provide a foundation from which to work. That is what philosophy does, and in this case, it is hard.

We noted above that functionalism can explain a mental state such as love. It does so by recognizing a particular neurological functional organization and resultant behaviors (e.g., distraction, particular attention to the subject of love and gift-giving). What functionalism does not do is explain the experience of love; the conscious, subjective feeling that either is part of or accompanies the mental state of love. Whether the subjective experience of consciousness is part of or accompanies a mental state, it clearly is special, immaterial, and different. It may even be unique to humans, or at least to higher-order animals. Its recognition would seem to reintroduce dualism and the intractable problem of connecting the dual states.

Since there is no scientific understanding of consciousness to give direction, philosophical explanations range between extremes. Functionalists maintain a monistic position by contending that consciousness is nothing more than high functionality. Anti-functionalists accept a dual argument. Both recognize that we just do not understand it yet.

The unity of the conscious experience has long perplexed thinkers. How is it that a sight, a smell, a sound, and a thought all merge into a conscious experience of incredible richness but never come together in a common neurological location? Science has identified specific brain locations for each component, but there appears to be no central consciousness brain location for their merger. This is the binding problem.

Francis Crick and Christof Koch contend that binding is a temporal rather than a spatial activity. Their belief is that neurons at spatially separate brain locations will adjust to a common rate of firing in the 40 Hertz range (40 firings per second). Though physically separate, the synchronous firing at a point in time produces a brain-wide mental state that we know as consciousness. Such synchronous activity has empirical support. Christiaan Huygens, in the seventeenth century, developed the concept of coupled oscillators by observing that grandfather clock pendulums would synchronize even when thrown out of synchrony. In our day, mirror neurons have been found that mimic neuronal experiences of others. Good scientific explanations for the coincidence of these physically-separated temporal phenomena have not been advanced. So it could be with consciousness.

As an alternative, Roger Penrose proposes a very interesting theory of consciousness. Most people believe that the brain is analogous to a computer. Following a set of preset logical rules (algorithms) such as addition and multiplication, a computer and therefore its analog the brain, will reach logical conclusions that are based on prior information. That certainly is true of computers, as now conceived. Penrose, however, shows that consciousness can transcend algorithmic activity to reach non-algorithmic conclusions. He not only shows that consciousness can occur; he shows that computers as now conceived cannot replace the brain because computers cannot achieve a state of non-algorithmic consciousness.

Penrose’s proof that consciousness is non-algorithmic is mathematically exact and based on two well-known mathematical theorems. The first is Turing’s halting problem solution. Turing showed that there are some problems that computer algorithms cannot solve. The second is Gödel’s incompleteness theorem. Gödel showed that the set of axioms on which algorithms are based limit algorithmic solution to less than everything that you want the algorithms to solve. Penrose then showed that our understanding those two theorems and their implications required non-algorithmic thought, and therefore, brains can do something that computers as now conceived cannot do. That something is conscious non-algorithmic thought.

The algorithmic approach is a very-Newtonian clockwork approach. Given initial conditions the algorithm leads to a defined result. Penrose’s non-algorithmic conscious thought leads to inspiration, creativity, free will and non-logical insights. These are conclusions that may not be based on stated premises and logical analysis. Penrose and Stuart Hammeroff appealed to quantum physics for reasoning. As described above in Physics, at the subatomic level probability rather than certainty rules. Non-algorithmic reasoning rules, and that is a hallmark of consciousness. Current neuroscience focuses on neurons, but that is too large a scale for quantum effects. Thus, Penrose and Hammeroff contend that consciousness research must look to subatomic brain structure and activity for understandable results. As an initial proposal, they suggest examining quantum effects in the microtubules of neuronal cytoskeletons.

Scientific Method

To this point, this presentation has addressed philosophical conceptual understanding of reality and knowledge and the biological process that creates knowledge in the brain. We now turn to the scientific method to understand how knowledge is developed through theory and experiment.

We establish what we know through the scientific method. This is a rationalist approach of developing a theory then finding if empirical evidence will support the theory. In the case at hand, the theory is that there is income variation among the United States regions that is partly caused by different costs of living among those regions.

Defining Terms

First, we must define terms. Exactly what do we mean by income, variation, regions and costs of living. From a scientific standpoint, those terms are not as evident as may have been first thought, and comparative and meta studies are heavily dependent on consistency of definitions.

Income can take many forms, most basically pecuniary and non-pecuniary. Even these categories are wide since pecuniary income can be gross, net of taxes, net of taxes and benefits, and net of a myriad other uses for pecuniary income. Income in-kind and other forms of non-pecuniary income likewise take many forms. These definitions are the normal lay ideas, but scientists also think in terms of resources. Thus, income can be determined in a variety of forms from the National Income and Product Accounts. Here, income represents the amount of resources that conceptual people command. In addition to income by person, income is collected and published where household or another multiple person unit is the reference. Further, income is often confused with wealth which is an entirely different concept.

Variation can lead to implications of inequality, and those measures have generated a wide literature. Whole books have been written on nothing but inequality measures alone. Among the statistical alternatives for measuring inequality are the standard deviation, variance, coefficient of variation, Theil Index, Herfindahl-Hirshman Index, Paasche Index, quintiles, regression coefficients and analysis of variance components. Each has its own strengths, and weaknesses.

United States regions also have multiple definitions. The U.S. Department of Labor publishes an eight-region division of the United States while the U.S. Census Bureau publishes a nine-region division. There are innumerable political and cultural divisions such as the 51 states (including District of Columbia), 3,300 counties, 12 Federal Reserve System districts, 435 Congressional districts, development districts, watershed districts, metropolitan/non-metropolitan areas, and urban/rural areas.

Finally, costs of living definitions present problems. Cost of living is defined in terms of a specific basket of goods. Clearly, different regions buy different representative baskets so comparison is apples and oranges. A common basket definition is necessary, but that basket is not necessarily representative of a particular region. Then, temporal changes in baskets are necessary which introduce often intractable temporal comparison problems. Temporal measurement also introduces index number definitions, of which there are many, such as the Laspeyres, Paasche, Fisher Exact and chained and unchained. Of real issue, is the fact that costs of living are not periodically collected in virtually all of the regions, however regions may be defined. Addressing the topic therefore requires proxies, their multiple definitions and conceptual inadequacies.

Just the basic definitions in so simple a topic create a host of concerns. Science advances by replication, yet the research papers on this topic use many combinations of these, and other, definitions, making replication questionable. Further, the definitions themselves are highly susceptible to change as physical reality or political convenience change.

Data Collection

Scientific studies rely on empirical support for theoretical conclusions. This is the information in “informed” decisions, and the information is always presented as numbers (data), even if it is qualitative information. Numbers are measurement, and measurement is the Achilles Heel of modern statistics and the practice of science.

As noted above, there are many definitions of income. One is Gross Domestic Income (GDI) as reported by the U.S. Department of Commerce, Bureau of Economic Analysis (BEA). The numerical value of GDI, part of the National Income and Product Accounts (NIPA), is based on a number of sources since there is no one survey for this purpose. BEA economists collect information from the U.S. Department of Labor, Bureau of Labor Statistics (BLS), Office of Management and Budget, Railroad Retirement Board, U.S. Department of Agriculture, Pension Benefit Guaranty Corporation, U.S. Treasury, U.S. Census Bureau, Social Security Administration, Internal Revenue Service, Federal Reserve Board, Federal Deposit Insurance Corporation, and Federal Financial Examination Council in order to compile this number and its components. As can be imagined, there is variation in timing and definition between these many agencies and their many

reports. Thus, many components of GDI must be imputed, and a very common phrase in the Summary of NIPA Methodologies that is published in BEA's Survey of Current Business is "judgmental trend." Even so important a number as GDI is very soft when its technical calculation is examined.

BEA and virtually all researchers depend on many surveys that collect primary data. A host of practical issues prevent comprehensive data collection so these surveys attempt to present representative sample information. There are clear questions about the accuracy of these many survey samples as description of one will illustrate.

BLS conducts the Current Employment Statistics (CES) survey each month. There are approximately nine million establishments in the United States that participate in state and federal unemployment insurance. BLS samples about 400,000 of these establishments for the CES. This is a very large sample so it is reasonable to assume that it is representative of the overall population of establishments. There is a chance, however, that the sample will not be representative, and we have no way of knowing since the population is unknown. If the sample is not representative, then the statistics is "garbage in – garbage out."

The vast majority of modern scientific work is based on samples that are extraordinarily small in relation to the population. In that work, the probability that the sample is not representative is relatively high which always puts a question mark on whether what we learn from those scientific experiments is actually true. An easy example can be found in the medical field. The many media outlets give great prominence to the latest medical determination. If you listen carefully, the news will mention the sample on which the remarkable finding is based. Very often, the sample will be less than one hundred, and the findings are extended to populations of millions. No matter how careful the researchers are, that small of a sample has a very high probability of being atypical, and the great findings being publicized have a high probability of being false.

There are a host of other problems with data. For example, a chronic problem with using government statistics is revision of the data. Research is often run, only to find out at a later date that the data has been revised. Thus, the research has been compromised. Also, comparability of the data, both temporal and geographic, is often suspect.

Statistical Inference

Modern science is based on statistical inference. It is rarely possible to know everything about a population (taking a census), so almost invariably, partial information is collected (taking a sample), and statistical methods are used to infer conclusions. This statistical inference is based on a host of assumptions.

First, the sample should be the unknown population in miniature. Randomness in collection and other sampling devices are used to assure this requirement as much as possible, but as noted above, one never knows for sure because the population is unknown. Second, the population's distribution is critical in statistical method selection. Again, the population is unknown so educated guesses have to be made. Third, a variety of statistical methods are available and within each method a variety of choices. Each method and each of the method choices are heavily assumption dependent. Checking assumptions is essential, but again, the unknown population prohibits complete confidence in selection. Finally, judgment plays a far more important role than scientists, and policy-makers if they know, would like. For example, Xavier Sala-i-Martin published a paper, "I Just Ran Two Million Regressions." He looked at sixty variables for explaining income inequality. He found that most could be found to significantly explain income inequality if the right specification was chosen. In other words, the researcher's judgment determined the scientific conclusions reached.

Philosophy of Science

The scientific method, despite the many problems identified above, defines science and the development of knowledge. But, just what is science, and therefore, non-science? And, is economics a science?

Physics is the discipline derived from philosophy by which the "scientific" nature of other derived disciplines is measured. This is because physics fundamentally led in developing empirically testable theories, and in empirically testing those theories. For Karl Popper, a science could be distinguished from a non-science by the presence of empirically falsifiable theories. Thomas Kuhn, on the other hand, distinguished science from non-science by the presence of agreement by a community of scholars on definitions of legitimate problems and methods. The difference is that Kuhn requires a science to have what is essentially political agreement while Popper demands empirical verification of falsifiable hypotheses.

Imre Lakatos reasoned with Popper that scientific knowledge has to result from empirical examination. He differed, however, because Popper discarded falsified knowledge while Lakatos included falsified knowledge as part of a larger program of knowledge. To him, though falsified, a theory should be retained because it could lead to insights of greater theoretical depth in an overall program that adhered to basic initial premises.

The theory of scientific evolution brings out the differences in these three concepts of “science.” Popper believed that science changed when empirically falsified knowledge was discarded. Lakatos believed that falsified information should be retained in a theory and used to constructively evolve change in a broader scientific base. Kuhn, on the other hand, required scientific authorities to agree on a change in science. Such change could be initiated from empirical information, but not necessarily. The change could come from political or social factors, as well.

Milton Friedman extended the concept of science by distinguishing between theoretical conclusions. Rationally developed and empirically supported theory provides conclusions that he identified as “positive” conclusions. Those “positive” conclusions, however, can be extended into policy prescriptions that he called “normative” conclusions since “should” and “ought” characterize policy proposals. The difference between positive and normative uses of science lies at the heart of the differences between Popper’s and Lakatos’ empirically tested science and Kuhn’s politically determined science.

Use of the scientific method in the social sciences, and therefore in economic science, has a special problem: the human element. While the physical sciences can control experiments, replicate experiments and work with elements that respond consistently; the social sciences also work with humans who have free will. Thus, experiments in the social sciences are difficult to control, and true replication is impossible. Humans respond inconsistently, and many controls are unacceptable on moral grounds. This raised the question that the social sciences, and economic science in particular, are not sciences since they fail to meet the positivist, Popper criterion of scientific respectability.

Alexander Rosenberg recognized the empirical theory predictive weakness in economics. He defended economics as a science because its development and practice paralleled physics, the reference science. Though predicting human behavior, aggregates react predictively and allow the Newtonian method of optimization and comparative equilibria to be used. This method leads to falsifiable hypotheses and science in both the Popper and Lakatos concepts.

Conclusion

This paper asserted a bit of knowledge: There is income variation among the United States regions, and that variation results, in part, from different costs of living among those regions. Then, an inquiry was made into, “How do we know that?” Do we know these facts (income variation and a reason) to be absolutely true? Or, is there sufficient probability of truth in those statements that we can regard them as true for practical purposes? Or, does the truth of those statements border on guess, or even pure speculation?

The inquiry found that the human brain internally constructs a significant proportion of knowledge, rather than faithfully internally reproducing the external world. In addition, that internal construction is conducted in a currently mysterious parallel processing manner. Furthermore though the external world appears exact, in its foundations (at the subatomic level), quantum physics has found that the external world can only be described in terms of chance. Thus, while philosophers speculate on possibilities, scientists have amazingly little understanding of how the brain and the world actually work. Compounding this lack of understanding is that most important part of our knowing something, our consciously knowing it. Scientists and the speculating philosophers are totally confounded by conscious knowledge, and therefore cannot actually describe how we consciously know something.

The conceptual difficulties described above are further compounded by the practical development of knowledge using the scientific method. Serious problems of definition, measurement, data collection and inference are addressed in every scientific investigation. Knowledge uncovered in these inquiries generate active, sometimes contentious, debate on what actually was determined, what should have been done that was not done, where the information actually is unreliable, and where the results do not really hold. At its heart, modern scientific investigation is based on inferences made from partial information about unknown populations. By definition, uncertainty characterizes such research.

Throughout this discussion, knowledge not belief has been the focus. Belief can be held in terms of certainty because belief is given by an external “giver of truth” whose omniscience is held on faith. As noted in the Preamble, what we know that is based on faith was not addressed in this discussion.

An inescapable conclusion of this inquiry is that knowledge is quite fragile. The inquiry brings wonderment at human advancements in philosophy and science, and even more wonderment at how the human mind works, as we understand it. Equally, the inquiry makes one wonder at how so many can have such certain opinions about what they “know.” But, they and we do!

References

- Craig, Edward. 2000. *Philosophy*. Oxford, U.K.: Oxford University Press.
- Flew, Anthony. 1991. *An Introduction to Western Philosophy*. New York, NY: Thames and Hudson.
- Friedman, Milton. 1996. "The Methodology of Positive Economics," in *Readings in the Philosophy of Social Science*, edited by Michael Martin and Lee C. McIntyre. Cambridge, MA: MIT Press.
- Hausman, Daniel M. 1984. *The Philosophy of Economics*. Cambridge, U.K.: Cambridge University Press.
- Kandel, Eric R., James H. Schwarz and Thomas M. Jessell. 2000. *Principles of Neural Science* (4th ed.). New York, NY: McGraw-Hill.
- Kuhn, Thomas S. 1970. *The Structure of Scientific Revolutions* (2nd ed.). Chicago, IL: Chicago University Press.
- Lakatos, Imre. 1976. "Proofs and Refutations," in *Philosophical Papers*, edited by J. Worrall and G. Currie. Cambridge, U.K.: Cambridge University Press.
- Larvor, Brendan. 1998. *Lakatos: An Introduction*. London, U.K. and New York, NY: Routledge Press.
- O'Shea, Michael. 2005. *The Brain*. Oxford, U.K.: Oxford University Press.
- Popper, Karl R. 1950. *The Poverty of Historicism*. Boston, MA: The Beacon Press.
- Popper, Karl R. 1965. *Conjectures and Refutation: The Growth of Scientific Knowledge*. New York, NY: Harper Torchbooks.
- Popper, Karl R. 1976. "The Logic of the Social Sciences," in *Positivist Dispute in German Sociology*, edited by Theodore W. Adorno. London, U.K.: Heinman Educational Books.
- Rosenberg, Alexander. 1996. "If Economics Isn't a Science, What is It?" in *Readings in the Philosophy of Social Science*, edited by Michael Martin and Lee C. McIntyre. Cambridge, MA: MIT Press.
- White, Thomas I. 2008. *Discovering Philosophy*. Upper Saddle River, NJ: Pearson/Prentice Hall.

Dollar Demise or Rise: The Long and the Short of It

Shuming Bai, University of Texas Permian Basin

Oviedo-Toral Gerardo, University of Texas Brownsville

Abstract

The recent global financial crisis casts doubt on the two most prominent reserve currencies. In this paper, we examine and forecast the future value of this important currency pair by employing four fundamental models and the econometric model for the period of 1999:01 to 2009:12. First, results show that fundamental models are valid framework for analyzing the trend of the exchange rate movement. Second, multivariate cointegrating relations confirm this result and are enhanced by the short-run dynamics. Finally, this optimal model outperforms the 24-month-ahead out-of-sample forecasting than other alternatives. We project a gradual shift toward a tripodal reserve regime.

Introduction

The U.S. dollar has been the dominant reserve currency since the Bretton Woods in the 1940s; however, its monumental deficits and record spending have raised concerns as the global reserve role during the recent global financial crisis. Meanwhile in eurozone, the financial woes in Greece, Portugal, and Spain transcend the biggest fear over the survival of the single common currency “euro” since its debut in January 1999, sending euro to its weakest value since November 2008 (Financial Times, May 14, 2010). The two most competing reserve currencies possess similar size of population and economy.

In this paper, we investigate the fundamental value of the euro-dollar and forecast its long-term trend and short-run dynamics. The paper differs from the previous studies in three folds: First, instead of the commonly-applied monetary models of exchange rates, we employ several different versions of the fundamental model advocated in Madura (2010) to forecast the exchange rate movement based on international parity theories. Second, prior literature has focused on dollar and other major currency pairs other than the euro and excluded the recent financial crisis, which is critically testing the credibility of the euro as an economic union but not political union for the first time since its launch eleven years ago. Hence, it is crucial to examine where this most important pair is going to move as it greatly affects the international business and cash flows. Third, the unique model derived from the conventional monetary models via multivariate cointegration technique enhanced by the short-term error correction model outperforms all other alternatives including the random walk model for out-of-sample prediction based on the root mean squared errors. Fourth, the findings are unique in that two specifications outperform the random walk model for out-of-sample prediction in the short-run, a result that is void in prior literature which supports 3-4 years of outperformance. The result contradicts with the findings of Meese and Rogoff (1983 and 1988), but supports those by MacDonald and Taylor (1991 and 1994).

The remainder of the paper is organized as follows. Section two provides literature review, the theoretical framework, and the fundamental models. Section three describes the data and the econometric methodology, Section four discusses empirical results, and Section five concludes.

Literature Review and Fundamental Models

Literature Review and Theoretical Framework

The Forecasting of exchange rates has been attempted for the past few decade and numerous methods have been developed with each method aiming at being more consistent in predicting a currency’s future value. However, no particular forecasting model has been considered as the “gold standard” since each model has its strengths and limitations. Fundamental forecasting techniques use historical data to develop relationships between exchange rates and macroeconomic variables such as relative inflation, interest rates, and income levels (Madura, 2010). Among the fundamental approaches, the monetary model where the exchange rate is a function of money demand, national output, and interest rates, has been tested extensively (Isard 1995). Frankel (1979) was the first to use the model with expansion and concluded that the gap between the current exchange rate and its equilibrium level is proportionate to a real interest differential. Meese and Rogoff (1983 and 1988), based on their extensive set of carefully formulated tests, reported that existing empirical models failed to significantly outperform a random walk model in forecasting out-of-sample at horizons of up to 12 months. They concluded that monetary models performed poorly due to the real disturbances on exchange markets. They also found no cointegration between real

exchange rates and real interest differentials, even though it is supposed to be linked by international parity. However, Oberlechner (2001) reported that both financial journalists and foreign exchange traders found fundamental forecasting to be superior to technical forecasting when predicting currency values over longer time frames. However, Mehran and Shahrokhi (1997) supported the random walk to predict the future exchange rate of the U.S. dollar and the Mexican peso and found the ARIMA model to be most predictive of future spot rates based on the lowest number of errors, even though different models were found to outperform others at different time periods.

Fundamental Approach

In this paper, we adopt the fundamental approach to forecast exchange rates using various models as explained in Madura (2010, pp. 276-278) that some of the variables that affect the spot rate of the U.S. dollar against the euro are the changes in the differential values of inflation, interest rates, and income growth between the U.S. and euroarea (EA). The EA consists of sixteen European countries (Austria, Belgium, Cyprus, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Malta, the Netherlands, Portugal, Slovakia, Slovenia, and Spain) that have adopted the euro as their common currency over time. The four different structural models of the fundamental approach can be specified as follows:

$$\text{Model 1 (Complete):} \quad \Delta s_t = \beta_0 + \beta_1 INF_{t-1} + \beta_2 INT_{t-1} + \beta_3 INC_{t-1} + \mu_t \quad (1)$$

$$\text{Model 2 (based on international Fisher equation):} \quad \Delta s_t = \beta_0 + \beta_1 INF_{t-1} + \beta_2 INT_{t-1} + \mu_t \quad (2)$$

$$\text{Model 3 (based on uncovered interest parity):} \quad \Delta s_t = \beta_0 + \beta_1 INF_{t-1} + \beta_3 INC_{t-1} + \mu_t \quad (3)$$

$$\text{Model 4 (based on purchasing power parity):} \quad \Delta s_t = \beta_0 + \beta_2 INT_{t-1} + \beta_3 INC_{t-1} + \mu_t \quad (4)$$

where: Δs_t = the percentage change in the spot rate in natural logarithm
 INF_{t-1} = the previous monthly percentage change in the inflation differential (U.S. inflation – foreign inflation)
 INT_{t-1} = the previous monthly percentage change in the interest rate differential (U.S. interest rate – foreign rate)
 INC_{t-1} = the previous monthly percentage change in the income growth differential in natural logarithm
 μ_t = the random error term at t with mean zero
 β 's = model parameters.

Model 1 is the complete model based on the key macroeconomic variables affecting exchange rates (Madura 2010). Model 2 is based on the theory of international Fisher equation and we are interested in re-testing here with the euro and with recent data to see whether we confirm or rebuttal the findings in Meese and Rogoff (1988). Model 3 represents the uncovered interest parity assumption while adding an exogenous variable of income level, while Model 4 contains the purchasing power parity assumption.

Data and Methodology

Data

We obtain the monthly and quarterly data of exchange rate of euro in dollar, inflation rate, interest rate, and national income (industrial production) from Federal Reserve Bank of St. Louis and from the database of the Organization for Economic Co-operation and Development (OECD) Eurostat. The exchange rate and the industrial production are in natural logarithm. We chose the period of all data to start with the inception of the euro in January 1999 and ending in the last month of 2009, the longest available data span (132 observations) for the euro value.

Methodology

We use rolling regression methodology for the sample data from 1999:01 to 2008:12 as the in-sample estimation period and the data from 2009:01 to 2009:12 as the out-of-sample forecasting period. The various fundamental forecasting models which are then compared with the multivariate Johansen cointegration model (Johansen 1988) which simultaneously estimate the long-run trend as well as the short-run dynamics of the variables. Therefore, we can derive the optimal forecasting model to predict the changes in the value of the dollar in respect to the euro in the following 12 month out-of-sample time span.

The flexible-price monetary model (Isard 1995, pp. 134) reduced form equation can be written as follows:

$$s_t = \beta_1 m_t - \beta_2 m_t^* + \beta_3 r_t - \beta_4 r_t^* + \beta_5 y_t - \beta_6 y_t + \mu_t \tag{5}$$

Empirical evidence shows that most of the macroeconomic variables are empirically nonstationary. Furthermore, with major currencies like the U.S. dollar and the euro which are widely used in international trade and capital flows, the fundamental variables will be sufficiently interrelated such that the exchange rates themselves will likely share a common trend. Therefore, it is expected that there exist at least one linear combination of this most liquid bilateral exchange rates that is stationary. We estimate a stationary monetary cointegration vector using the multivariate Johansen estimation technique (Johansen 1988) by determining how many of the characteristic roots are less than unity as follows:

$$\text{Model 5 (Johansen long-run monetary model): } \Delta X_t = A_0 + \pi X_{t-1} + \sum_{i=1}^{k-1} \pi_i \Delta X_{t-i} + \varepsilon_t \tag{6}$$

where $X_t = (m_t, m_t^*, y_t, y_t^*, r_t, r_t^*)$ is a p-dimensional vector of variables, A_k are $(N \times N)$ coefficient matrices, A_0 is $(N \times 1)$ vector of constants, k is the appropriate lag length determined by the Akaike information criterion (AIC) and the Schwarz Bayesian criterion (SBC) to whiten the residuals, and ε_t is $(N \times 1)$ the white noise error vector with non-diagonal covariance matrix Ω . Where $\pi = \alpha\beta'$, where β represents parameters of r cointegrating vectors or equilibrium errors from the long-run relationship, and A_j is $(n \times n)$ coefficient matrices.

Empirical Results

Table 1 reports the regression results from the four models of the fundamental approach. The regression analysis determines the direction and magnitude to which euro is affected by each independent variable. The regression coefficient β_1 exhibits a positive sign and is statistically significant at the 1% level across Models 1-3 because when U.S. inflation increases relative to that of the euro area, there is an upward pressure on the value of the euro. $\beta_1 = 0.024$ indicates that for a one percent increase in the inflation differential, the euro is expected to appreciate by 2.4 percent. Similarly, the coefficient β_3 , which measures the impact of industrial output on euro is also expected to be positive because when U.S. income growth exceeds that of the British, an upward pressure will be exerted on the value of the euro. These relationships are consistent with the international finance theories. Finally, the negative sign of β_2 indicates an inverse relationship between interest rate and the euro-dollar exchange movements such that higher interest rate country implies higher expected inflation rate according to the international Fisher effect. $\beta_2 = -0.003$ indicates that for one percent decrease in the interest rate differential, the dollar is expected to appreciate against the euro.

Table 1: Results of the fundamental models

Variable	Model 1		Model 2		Model 3		Model 4	
	Coefficient	T stat	Coefficient	T stat	Coefficient	T stat	Coefficient	T stat
Constant	0.002	0.758	0.002	0.759	0.002	0.797	0.001	0.569
InfDif	0.024	3.925**	0.023	3.947**	0.024	3.935**		
IntDif	-0.003	-0.314	-0.002	-0.229			0.002	0.224
IpDif	0.123	0.608			0.132	0.581	-0.151	-0.639
F stat	5.298**		7.805**		7.960**		0.216	
R ²	0.121		0.119		0.121		0.004	

* and ** are significant at the 5% and 1% levels, respectively

Table 2 sets out the multivariate Johansen cointegration tests of the monetary variables based on the TRACE eigenvalue statistics. We reject the null hypothesis that there are no cointegrating vectors (CIVs) and we identify a long-run CIV among money supply, industrial production, and short term interest rates in the U.S. and the euro area at the 5 percent significance level, indicating that the monetary model seems to have long-run stability. This result contrasts sharply the findings of some researchers such as Meese and Rogoff (1983 and 1988), but concur with MacDonald and Taylor (1991 and 1994) and Nautz and Offermanns (2006). Since every cointegrating vector has a corresponding error correction term which adjusts the short-term deviation from the long-run equilibrium of the exchange rate, we incorporate the error correction term into the monetary model and employ it for both in-the-sample and out-of-sample forecasting. We are able to obtain the lowest root mean squared error for those variables in forecasting performance compared to all other alternatives.

Table 2: Results of multivariate Johansen long-run relationship for the monetary variables

Number of CIVs	λ_{trace}	5% CV	Probability	Deterministic Specification	Lag length (k)
$r = 0$	132.845**	125.615	0.017	Trend in CE and VAR	3
$r \leq 1$	95.557	95.754	0.052	Trend in CE and VAR	3
$r \leq 2$	66.543	69.819	0.089	Trend in CE and VAR	3
$r \leq 3$	43.429	47.856	0.123	Trend in CE and VAR	3
$r \leq 4$	24.240	29.797	0.191	Trend in CE and VAR	3
$r \leq 5$	10.406	15.495	0.251	Trend in CE and VAR	3
$r \leq 6$	0.439	3.841	0.507	Trend in CE and VAR	3

Note: The λ_{trace} statistics are computed against the critical values in MacKinnon-Haug-Michelis (1999). The VAR is estimated with lag length (k) selected by the sequential test-down until white residual using the likelihood ratio (LR) test. r is the number of cointegrating vectors (CIVs). The deterministic specification in VAR has trend in CE and VAR. * and ** denotes rejection of the hypothesis at the 5% and 1% levels, respectively.

Out-of-Sample Forecasts

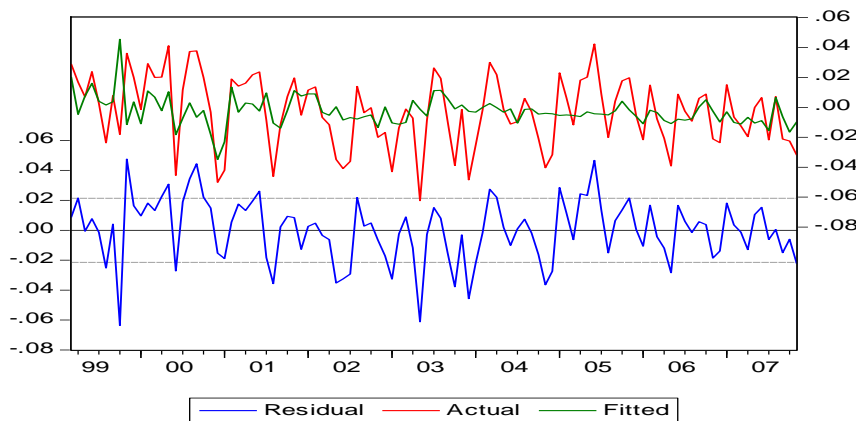
Based on the five models we estimated, we now use the models to predict the out-of-sample forecasting for the exchange rate changes for 12 month ahead from 2009:01 to 2009:12. The out-of-sample forecasts are conditional based on Model 1, Model 2 and Model 3 of the fundamental approach and the error correction model from the cointegrating estimation using the observed values for inflation, interest rate, and income level.

The evaluation of forecast accuracy is based on the root mean squared error (RMSE) evaluated by the following equation:

$$RMSE_h = \sqrt{\frac{1}{N} \sum_{t=1}^n (\hat{s}_t^h - s_t)^2} \tag{7}$$

where $t_0 = 2008:12$, $n = 2009:12$, $t = t_0 + h$, \hat{s}_t^h is the h-step forecast of s_t (or the spot rate value predicted by the model) for sample case t (out of n sample cases) and $N = 12-h+1$ is the number of predicted values; s_t is the target spot rate value for sample case t , For a perfect fit, $\hat{s}_t^h = s_t$ and $RMSE_h = 0$. So, the $RMSE_h$ index ranges from 0 to infinity, with 0 corresponding to the ideal.

Figure 1: Actual and fitted values of the exchange rate 1999:01-2008:12



The RMSE for the forecast equations and the random walk for $h = 1, \dots, 12$ are depicted in Figure 1. The figure shows that the error correction model of the monetary variables in which the predicted value tracks the actual exchange rate change and captured some important turning points. It exhibits lower root mean squared forecast errors for the euro-dollar exchange rate at every forecast horizon than the other three fundamental models which in turn outperform the random walk over forecast

horizons within a year. This is a relatively short time span compare with the prior literature that structural models of nominal exchange rates can outperform the random walk at long-horizons of 3-4 years (see Mark, 1995; Chinn and Meese, 1995). Our models (Models 1 and 2) provide evidence that the fundamental model of exchange rates can outperform the random walk at short horizons of one year ahead.

Table 3 reports the out-of-sample forecasting using the 3 significant fundamental models and the error correction models and compare with the random walk. The error correction model was estimated up to the end of 2008. This estimated equation was then used to forecast the exchange rate for five different horizons over the period 2009:01 to 2009:12. The mean squared errors are calculated over the forecasting horizon for the total 5 models employed in the paper. All models outperform the random walk models and the error correction model dominates all alternative models with the lowest mean squared errors.

Table 3: Comparison of out-of-sample forecast errors of different models based on RMSE (2009:01 to 2009:12)

Forecast horizon (months)	Model 1	Model 2	Model 3	Model 5 (ECM)	Random walk
1	0.0235	0.023	0.0234	0.0209	0.061
3	0.0238	0.024	0.0237	0.0211	0.037
6	0.0239	0.024	0.0239	0.0215	0.038
9	0.0236	0.0236	0.235	0.022	0.035
12	0.0238	0.026	0.0237	0.021	0.033

Note: Root mean squared errors (RMSE).

Conclusions and Implications

In this paper we re-examined the various models of the fundamental approach and the econometric models to assess their validity and usefulness in forecasting the future euro-dollar currency pair. Prior literature shows that exchange rate models estimated with different data usually lead to different results.

The fundamental models derived from the most important economic performance indicators of a country like national income, interest rates, and inflation have been widely used as the determinants of exchange rate and three of the four models exhibit some predicting power, contrasting what prior studies have found.

More interestingly, the Multivariate Johansen cointegration tests of the monetary variables demonstrate at least a long-run stationary relationship in forecasting the trend of the exchange rate. This finding is important because it lends future evidence to the many widely-held parities in international finance. The cointegrating relations contrast sharply with prior studies, too as pointed out in Meese (1983). The out-of-sample forecasting using the error correction models shows the lowest forecasting errors, outperforming other fundamentals and the random walk. The findings show that the dollar is to continue with the falling trend in the short- and medial run, since the U.S. inflation and national income are higher than their European counterparts. The current worsening situation in Greece, Spain and Portugal may temporary help the dollar to rally. However, dollar won't be strengthening without solving its fundamental debt problems just as the PIGS has been involved. The current eurozone sovereign debt issues among several members poses immense threat to the community and the world because of the close tie among its EA members. The sovereign debt, along with the intended enlargement and disparity of political and economic development may lead to further challenge to test the real viability of this common currency. Meanwhile in Asia, concerns over the American sovereign debt and high unemployment have motivated the Chinese to expand the role of its currency yuan in international trade and capital flows. The world currency arena may be played off by the tripolarity of the reserve currency regimes in the future.

The findings of this paper are show that there is no strong tendency for the dollar to appreciate in the short and medium. However the tendency shown in our models for a weak dollar in the future seems consistent with the current low global economic activity and the recent large spending from the U.S. government. This gives insight that a large dollar appreciation in the future is unlikely.

Future research in forecasting can account for government intervention to help to understand the exchange rate behavior better especially in the current euro zones besides the structural economic models. Government intervention plays an important role in currency values and consequently in exchange rate determination.

References

- Enders, Walter. 1995. *Applied Econometric Time Series*. New York: John Wiley & Sons, Inc.
- Frankel, Jeffrey A. 1979. "On the Mark: A Theory of Floating Exchange Rates Based on Real Interest Differentials." *American Economic Review*. 64: 610-622.
- Garnham, Peter, David Oakley, and Henny Sender. 2010. "Fears Grow over Weaker Euro." *Financial Times*. May 14.
- Isard, Peter. 1995. *Exchange Rate Economics*. Cambridge: Cambridge University Press.
- Johansen, Soren. 1988. "Statistical Analysis of Cointegration Vectors." *Journal of Economic Dynamics and Control*. June/September, 12: 231-254.
- Johansen, Soren, and Katarina Juselius. 1990. "Maximum Likelihood Estimation and Inference on Cointegration – With Applications to the Demand for Money." *Oxford Bulletin of Economics and Statistics* 52: 169-210.
- Liu, W. 2008. "Forecasting Exchange Rate Change Between USD and JPY by Using Dynamic Adaptive Neuron-Fuzzy Logic System." *Asia Pacific Journal of Finance and Banking Research*, 2(2): 2-12.
- MacDonald, Ronald and Mark P. Taylor. 1991. "A Stable U.S. Money Demand Function, 1874-1975." *Economics Letters* 39: 191-198.
- MacDonald, Ronald and Mark P. Taylor. 1994. "The Monetary Model of the Exchange Rate: Long-run Relationships, Short-run Dynamics and How to Beat a Random Walk." *Journal of International Money and Finance* 13: 276-290.
- Madura, Jeff. 2010. *International Financial Management. Forecasting Exchange Rates* (pp. 252-261). Mason, Ohio: Thomson South-Western.
- Mark, N. C. 1995. Exchange Rates and Fundamentals: Evidence on Long-horizon Predictability. *American Economic Review* 85: 201-218.
- Mehran, J. and M. Shahrokhi. 1997. "An Application of Four Foreign Currency Forecasting Models to the U.S. Dollar and Mexican Peso." *Global Finance Journal*, 8(2): 211-220.
- Meese, Richard M. and Kenneth Rogoff. 1983. "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?" *Journal of International Economics* 14: 3-74.
- Meese, Richard M. and Kenneth Rogoff. 1988. "Was It Real? The Exchange Rate-Interest Differential Relation over the Modern Floating-Rate Period." *The Journal of Finance* 43(4): 933-948.
- Nautz, Dieter and Christian j. Offermanns. 2006. "Does the Euro Follow the German Mark? Evidence from the Monetary Model of the Exchange Rate." *European Economic Review* 50: 1279-1295.
- Oberlechner, T.. 2001. "Importance of Technical and Fundamental Analysis in the European Foreign Exchange Market." *International Journal of Finance and Economics*, 6: 81-93.
- Rajan, R. and J. Kiran. 2006. "The US Dollar and the Euro as International Currencies." *Intereconomics*, 41(3): 124-141. doi:10.1007/s10272-006-0183-6.

U.S. Mortgage Foreclosures: An Analysis by State

William Brent, Lynne Kelly, Debbie Lindsey, and Russell Price, Howard University

Abstract

This paper examines the pattern of mortgage foreclosures across the US in order to gain insight into the key factors that explain geographic variation in the incidence of foreclosures. A two-way fixed effects model is used to estimate foreclosure rates on loans, prime loans, and subprime loans for 2004 - 2007. Overall, the factors that significantly influence foreclosure rates are: lagged personal income; credit score (prime borrowers); percent NIV loans; growth in the housing price index; construction starts; unemployment rate; loans purchased by non-GSEs; and denial rates. The findings reveal evidence to support initiatives designed to help mitigate foreclosure.

Introduction

In 2009, foreclosure actions were initiated on 2,824,674 US properties, one in every forty-five households, according to RealtyTrac. This represents a 21 percent increase in total properties from 2008 and a 120 percent increase from 2007. In July 2009, foreclosure filings peaked at 361,000 households. This was followed by four straight monthly declines in the number of foreclosure filings driven primarily by legislative and industry-related delays in processing delinquent loans. However, in December 2009, foreclosure filings increased 14 percent from November 2009 to 349,519 households.

Nevada, Arizona, Florida, and California accounted for roughly 50 percent of the foreclosure filings in the country. In 2009, foreclosure actions were initiated on 10.2 percent of the mortgages in Nevada, 6.1 percent of the mortgages in Arizona, 5.9 percent of the mortgages in Florida, and 4.8 percent of those in California. Other states with 2009 foreclosure rates ranking among the nation's ten highest were Utah with 2.9 percent, Idaho and Georgia with 2.7 percent, Michigan with 2.6 percent, Illinois with 2.5 percent, and Colorado with 2.4 percent.

The number and geographic dispersion of foreclosures in recent years has led researchers to document the problem and attempt to inform intervention strategies. This paper examines the pattern of mortgage foreclosures across the US in order to gain insight into the key factors that explain geographic dispersion in foreclosure rates. Using cross-sectional time series regressions, the importance of borrower characteristics, loan characteristics, economic conditions, and lending conditions in explaining foreclosure rates by state is examined. The findings reveal evidence to support, at the state and local levels, initiatives, policies, and programs designed to help mitigate foreclosure. The paper proceeds as follows. In the second section, we discuss determinants of mortgage default and foreclosure. In the third section, we formulate an econometric model of US mortgage foreclosure rates. In the fourth section, we present data on mortgage foreclosure rates across the US. In the fifth section, we present our empirical analysis. Our conclusions are presented in the sixth section.

Determinants of Mortgage Default and Foreclosure

Foreclosure occurs in two steps. In the first step, the borrower is unable to make a scheduled payment and becomes delinquent. Typically, lenders assume the initial delinquency is temporary and that the borrower will resume payments in the future. However, once the borrower misses three or more payments, he is considered to have defaulted. Lenders may then proceed with foreclosure filings. A significant amount of research examines the factors that lead to default and subsequent foreclosure. Following is an overview of the literature.

One explanation for default is a pure wealth maximization motive. Individuals can increase their wealth by defaulting on a mortgage when the market value of the mortgage equals or exceeds the value of the house. Such defaults are termed "ruthless defaults" in the sense that the borrower has the ability to pay but chooses not to do so. Ambrose and Capone (1998) find that ruthless defaulters have high loan-to-value (LTV) ratios and experience local price declines prior to default. Foster and Van Order (1984), Schwartz and Torous (1993) and Case and Shiller (1995) find a strong relationship between property price declines and default incidence.

Although most defaults tend to be consistent with a pure wealth maximization motive, researchers find that the option to default is under exercised. One explanation for this behavior is that there are significant direct costs such as legal fees and indirect costs such as a reduction in creditworthiness associated with foreclosure. Considering these other costs, the borrower may not default even if the option to do so is in the money.

Another departure from a pure wealth maximization motive is that some borrowers default with positive equity. In this case the amount of equity does not exceed the cost of selling and the household does not have the liquid assets to make up the

difference upon sale. These liquidity-constrained borrowers often have difficulties that arise from “trigger events” such as jumps in mortgage payments because of higher interest rates on an adjustable-rate mortgage (ARM). Zorn and Lea (1989) examine ARMs in Canada and find that they have a higher probability of default than fixed-rate mortgages in the US.

Job loss is an important trigger event that negatively impacts the borrower’s ability to handle mortgage payments. The potential for job loss is expected to be higher during periods of high unemployment. Several studies have examined the relationship between unemployment and mortgage default. The results of these studies are mixed. Foster and Van Order (1984) Vandell and Thibodeau (1985), and Claurette (1987) find unemployment to be insignificant for explaining default. Cunningham and Capone (1990) find unemployment to be significant but, negatively related to default. Campbell and Dietrich (1983), Lea and Zorn (1986) and Deng, Quigley, and Van Order (2000) find unemployment to be highly significant and positively related to default.

The borrower’s financial capabilities and general creditworthiness are key factors in assessing default risk. The affect of borrower income on default probability has been studied extensively. Case and Shiller (1995) find that the change in log per capita personal income over the last four quarters is negatively significant. Capone (2002) finds that moderate income homeowners have a lower propensity to default than both very low and high income borrowers, conditional on having an incentive to default.

Loan ratio is the ratio of loan amount to income of the borrower. The higher the loan ratio, the greater the proportion of the borrower’s income that goes toward making mortgage payments and the higher the probability of facing financial difficulties in the case of a trigger event. Findings for the affect of loan ratio on default risk are mixed. Herzog and Earley (1970) find the influence to be insignificant or negative. Teo (2004) finds the influence to be negative.

Loan documentation status is related to loan ratio because uncertainty about the borrower’s true income translates into uncertainty about the stated loan ratio and, therefore, the borrower’s ability to handle mortgage payments. Traditional mortgage underwriting requires verification of the borrower’s income. In recent years, lenders developed no income verification (NIV) loans to accommodate creditworthy borrowers with irregular incomes. However, the practice has been abused in recent years in order to qualify borrowers who probably would not qualify under traditional underwriting criteria. Hayre, Saraf, Young, and Chen (2008) demonstrate through mortgage default models higher default rates for NIV loans.

A borrower’s credit score is often used as an indicator of creditworthiness. Many large lenders have developed proprietary credit scoring methods to screen borrowers. Findings for the affect of credit score on default risk are mixed. Pennington-Cross (2003) examines the performance of thirty-year fixed rate, owner-occupied, home purchase mortgages for the period 1995 through 1999 and finds that higher credit scores are associated with lower levels of default for both prime and nonprime mortgages. Merry and Wilson (2006) examine mortgage default rates in each state and the District of Columbia from 2003 through 2005. They find that subprime loan default rates decrease as the share of borrowers with credit scores below 700 increases. Hayre, Saraf, Young, and Chen (2008) argue that credit scores may become less predictive of credit performance because borrowers used previous home equity gains to smooth consumption and improve their credit records.

Spread at origination is the difference between the coupon rate on a loan and a measure of average prevailing mortgage rates. A high spread indicates the lender perceived that the borrower had above-average risk. Therefore, spread is a measure of the credit impairment of the borrower. Pennington-Cross (2003) uses spreads to categorize mortgages as prime and nonprime and finds that thirty-year fixed rate mortgages with high spreads default at elevated rates.

Owner-occupancy status is considered to be important in explaining the propensity to default on a mortgage. If the mortgaged property is the full-time residence of the borrower, then it is owner-occupied. Otherwise, it is an investor property. Mortgages on owner-occupied properties are considered to be less risky because borrowers who live in a house are more motivated to keep it. However, Teo (2004) finds the influence of owner-occupancy status to be insignificant in explaining default incidence.

The availability of credit has also been examined in studies of mortgage default. Hendershott and Kane (1992) argue that excess mortgage supply can lead to disproportionate space supply, low property returns, and high default rates. Stevenson (1995) concludes that default and foreclosure follow rapid build-ups in lending in the commercial market. Mian and Sufi (2009) find that subprime ZIP codes experience mortgage credit growth that is more than twice as high as prime ZIP codes during the period 2002 to 2005.

In general, much of the research on the factors that explain mortgage default and foreclosure has been conducted on large databases of loans that are held in portfolios. These studies use data from individual loans in logit or probit regressions where the incidence of default is modeled as an indicator variable. Other related studies estimate timing to default. Drawing from this literature, we model mortgage default rates in each state and Washington, DC using variables that can be classified into the following four categories: borrower characteristics, loan characteristics, economic conditions, and lending conditions. The affect these variables have on foreclosure rates is estimated in cross-sectional time-series regressions.

Model of US Mortgage Foreclosure Rates

We use a two-way fixed effects model to examine cross-sectional time series data on mortgage foreclosure rates in the US. Our data consists of foreclosure rates for all fifty states and Washington, DC for the period 2004 through 2007. Therefore, there are fifty-one cross-sectional observations and four time series observations. Greene (2003) expresses the two-way fixed effects model as follows:

$$y_{it} = \sum_{k=1}^K x_{itk} \beta_k + u_{it} \quad (1)$$

$$u_{it} = \gamma_i + \alpha_t + e_{it} \quad (2)$$

where y_{it} is the dependent variable, β_k represents the slope coefficients that are common to all cross-sections, x_{it} are the explanatory variables, γ_i s are nonrandom parameters capturing the unobserved cross-sectional effects, α_t s are nonrandom parameters capturing the unobserved time effects, and e_{it} are the residuals which are independent and identically distributed random variables with $E[e_{it}] = 0$ and $E[e_{it}^2] = \sigma_e^2$.

Since the dataset is balanced, we can write the following:

$$\tilde{y}_{it} = y_{it} - \bar{y}_i - \bar{y}_t + \bar{y} \quad (3)$$

$$\tilde{x}_{it} = x_{it} - \bar{x}_i - \bar{x}_t + \bar{x} \quad (4)$$

where the symbols:

y_{it} and x_{it} are the dependent variable (a scalar) and the explanatory variables (a vector whose columns are the independent variables not including a constant), respectively, \bar{y}_i and \bar{x}_i are cross section means, \bar{y}_t and \bar{x}_t are time means, and \bar{y} and \bar{x} are the overall means.

The two-way fixed effects model is simply a regression of \tilde{y}_{it} on \tilde{x}_{it} . Therefore, the two-way β is given by:

$$\tilde{\beta}_S = (\tilde{X} \tilde{X}')^{-1} \tilde{X}' \tilde{y} \quad (5)$$

The calculations of cross section dummy variables, time dummy variables, and intercepts follows. Denote the cross-sectional effects by γ and the time effects by α . These effects are calculated from the following relations:

$$\hat{\gamma}_i = (\bar{y}_i - \bar{y}) - \tilde{\beta}_S (\bar{x}_i - \bar{x}) \quad (6)$$

$$\hat{\alpha}_t = (\bar{y}_t - \bar{y}) - \tilde{\beta}_S (\bar{x}_t - \bar{x}) \quad (7)$$

Denote the cross-sectional dummy variables and time dummy variables with the superscript C and T. When no intercept is specified, the dummy equations can be written as follows:

$$D_i^C = \hat{\gamma}_i + \hat{\alpha}_T \quad (8)$$

$$D_t^T = \hat{\alpha}_t - \hat{\alpha}_T \quad (9)$$

The sum of squared errors is:

$$SSE = \sum_{i=1}^N \sum_{t=1}^{T_i} (y_{it} - \gamma_i - \alpha_t + X_{it} \tilde{\beta}_S)^2 \quad (10)$$

The estimated error variance can be written:

$$\hat{\sigma}_e^2 = SSE / (M - N - T - (K - 1)) \quad (11)$$

The variance covariance matrix of $\tilde{\beta}_s$ is given by

$$\text{Var}[\tilde{\beta}_s] = \hat{\sigma}_e^2 (\tilde{X}_s' \tilde{X}_s)^{-1} \quad (12)$$

The variance covariance matrix of the dummy variables are specified as follows:

$$\text{Var}[D_i^C] = \hat{\sigma}_e^2 \left(\frac{1}{T} + \frac{1}{N} - \frac{1}{NT} \right) + (\bar{x}_i + \bar{x}_t - \bar{x})' \text{Var}[\tilde{\beta}_s] (\bar{x}_i + \bar{x}_t - \bar{x}) \quad (13)$$

$$\text{Var}[D_i^T] = \frac{2\hat{\sigma}_e^2}{N} + (\bar{x}_i - \bar{x}_T)' \text{Var}[\tilde{\beta}_s] (\bar{x}_i - \bar{x}_T) \quad (14)$$

$$\text{Cov}[D_i^C, D_j^C] = \hat{\sigma}_e^2 \left(\frac{1}{N} - \frac{1}{NT} \right) + (\bar{x}_m i i + \bar{x}_m i \cdot t - \bar{x})' \text{Var}[\tilde{\beta}_s] (\bar{x}_m i j + \bar{x}_m i \cdot t - \bar{x}) \quad (15)$$

$$\text{Cov}[D_i^T, D_u^T] = \frac{\hat{\sigma}_e^2}{N} + (\bar{x}_i - \bar{x}_T)' \text{Var}[\tilde{\beta}_s] (\bar{x}_u - \bar{x}_T) \quad (16)$$

$$(15) \text{Cov}(D_i^C, D_i^T) = -\frac{\hat{\sigma}_e^2}{N} + (\bar{x}_i + \bar{x}_t - \bar{x})' \text{Var}[\tilde{\beta}_s] (\bar{x}_i - \bar{x}_T) \quad (17)$$

$$\text{Cov}[D_i^C, \beta] = -(\bar{x}_i + \bar{x}_t - \bar{x})' \text{Var}[\tilde{\beta}_s] \quad (18)$$

$$\text{Cov}(D_i^T, \beta) = -(\bar{x}_i - \bar{x}_T)' \text{Var}[\tilde{\beta}_s] \quad (19)$$

Data

The data on foreclosure start rates are from Mortgage Bankers Association National Delinquency Survey. The data are for each state and Washington, DC for the years 2004 through 2007.

During the sample period, states with foreclosure rates ranking among the nation's five highest are Indiana with 1.04 percent, Ohio with 0.99 percent, Michigan with 0.93 percent, Mississippi with 0.78 percent, and Georgia with 0.74 percent. As measured by standard deviation, the most volatile states for foreclosure activity are Nevada with 0.62 percent, Florida with 0.58 percent, California with 0.45 percent, Rhode Island with 0.41 percent, and Mississippi with 0.37 percent. The most volatile state, Nevada, is thirty-one times more volatile than the least volatile state, Montana. Comparing foreclosure rates and volatility is also interesting. Both increase over the sample period. In 2004, foreclosure rates and volatility are 0.43 percent and 0.19 percent, respectively. In 2007, foreclosure rates and volatility are 0.72 percent and 0.30 percent, respectively. [Data table available upon request.]

Foreclosure rates of prime loans are now considered. Our data shows that these rates are much lower than those for all loans, 0.25 percent on average over the sample period versus 0.51 percent. The five states with the highest foreclosure rates among prime borrowers are Indiana with 0.51 percent, Ohio with 0.47 percent, Michigan with 0.41 percent, South Carolina with 0.39 percent, and Mississippi with 0.38 percent. The five most volatile states for foreclosure activity among prime borrowers are Florida with 0.34 percent, Nevada with 0.30 percent, California with 0.22 percent, Mississippi with 0.21 percent, and Michigan with 0.18 percent. Foreclosure rates and volatility of prime loans increased over the sample period. [Data table available upon request.]

Our data also shows that foreclosure rates of subprime loans are much higher than those for all loans and prime loans. On average, subprime foreclosure rates are 2.1 percent over the sample period versus 0.51 percent for all loans and 0.25 percent for prime loans. Subprime borrower foreclosure rates are 8.4 times those for prime borrowers. The five states with the highest foreclosure rates among subprime borrowers are Michigan with 3.48 percent, Ohio with 3.22 percent, Indiana with 3.18 percent, Iowa with 2.79 percent, and Minnesota with 2.72 percent. The five most volatile states for foreclosure activity among subprime borrowers are Nevada with 2.26 percent, Rhode Island with 2.11 percent, California with 2.07 percent, Florida with 1.94 percent, and Massachusetts with 1.63 percent. Foreclosure rates and volatility of subprime loans increased over the sample period. [Data table available upon request.]

The explanatory variables used to estimate foreclosure rates for each state are listed by broad categories and discussed and illustrated in Tables 1- 2. A trend analysis for each independent variable over the sample period is then provided.

Table 1: Dependent Variable Foreclosure Starts

Independent Variables		Theoretical Underwriting or Economic Relationship	Direction	
			Direct (+)	Inverse (-)
Borrower Characteristics				
credscr	credit score	Higher credit scores signify lower credit risk, and in turn, lower risk of defaults and foreclosures.		X
laginc	lagged personal income (current dollars)	Higher incomes reduce credit risk, and in turn, lower risk of defaults and foreclosures.		X
gwthinc	growth in personal income (current dollars)	Income growth lowers credit risk, and in turn, lowers risk of defaults and foreclosures.		X
Loan Characteristics				
loanrat	loan ratio loan amount/income	Higher loan ratios mean that the proportion of borrower income that goes toward making mortgage payments is higher. This increases credit risk and, in turn, increases foreclosures.	X	
loanratl	lagged loan ratio	Higher loan to income ratios from the previous year should result in higher foreclosures within 12 months.	X	
ltv	loan to value ratio	Higher LTVs signify lower equity investment by borrowers, which increases credit risk and, in turn, increases risk of defaults and foreclosures.	X	
pctadj	adjustable loans /total loans (%)	ARMs are subject to payment shocks. This increases credit risk, and, in turn, increases risk of defaults and foreclosures.	X	
nivpct	no income verification loans/ total loans (%)	NIV loans contain the risk that the borrower has insufficient income to support the loan thereby increasing credit risk, defaults and foreclosures.	X	
nonocc	non-owner occupied loans/ total loans (%)	Investor loans are considered to be more risky than owner-occupied loans because investors are not as motivated to keep properties that are not their primary residence. Investor loans have greater risk of default and foreclosure.	X	
spread	loans with a reportable spread/ total loans (%)	High cost loans indicate that the lender perceives the borrower has above-average risk. Also, loans with a spread have higher payments, and increased credit and default risk.	X	
Economic Conditions				
delhpidx	Change (growth in) Housing Price Index	Increases in home prices increase borrower equity position resulting in an increased desire to remain in the home or the ability to sell or refinance if a trigger event occurs.		X
unempl	unemployment rate (%)	Higher unemployment increases the probability of job loss which increases the potential for default and foreclosure.	X	
constart	construction starts	Construction starts provide a leading indicator of strong economic conditions, particularly in housing. Strong economic conditions should result in fewer foreclosures.		X
Lending Conditions				
dloanorg	Rate of change in the percent of loans for home purchase	Higher rates of change in mortgage loans for home purchase signify a loosening of mortgage supply to borrowers who may not have previously been offered mortgages. This is expected to increase the pool of loans that fall into foreclosure.	X	
nongse	loans purchased by non-GSEs (commercial banks, institutions, affiliates)/ total loans (%)	Higher percent of Non-GSE loans imply more non-conforming loans are sold in the secondary market. This serves as an indicator that less toxic loans are originated; thus risk of defaults and foreclosures is less likely.		X
denpct	denial rate	Higher denial rates suggest higher credit standards, lower credit risk, and lower defaults. There are however divergent patterns of denial rates for prime and subprime lenders. Subprime lenders attract many applications from less creditworthy individuals resulting in high denial rates. To the extent that this variable includes subprime activity it may not have the inverse relationship as expected.		X
loanpur	loans purchased/ total loans (%)	High loan purchases are a barometer of more correspondent lending behavior or wholesale activity in the mortgage market. Correspondent lending could increase the likelihood of fraud and inflated values and in turn increase foreclosures.	X	

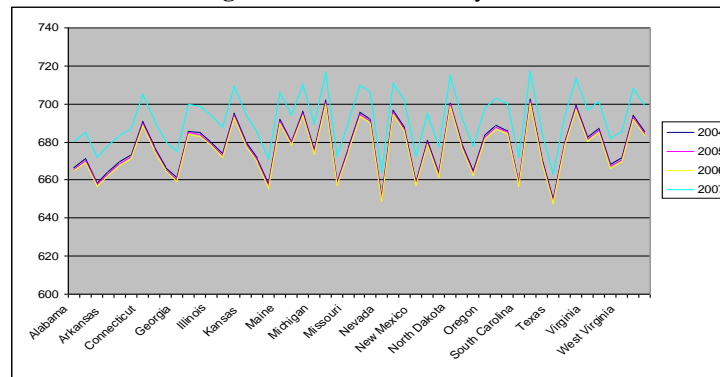
Table 2: Dependent Variable Foreclosure Starts

Independent Variables	2004	2005	2006	2007
credscr	678	677	676	692
laginc	\$31364	\$32336	\$33863	\$35768
gwthinc	3.05%	4.52%	5.49%	5.07%
loanrat	5.01	5.33	5.43	5.43
loanratl	4.47	5.01	5.33	5.43
ltv	77.29	75.65	77.83	79.43
pctadj	14.68%	17.11%	17.27%	16.65%
nivpct	1.959%	1.958%	1.960%	1.959%
nonocc	11.788%	12.924%	13.178%	12.979%
spread	15.862%	26.134%	28.665%	18.095%
delhpidx	235.703	265.010	281.752	285.494
unempl	4.869	4.514	4.100	4.461
constart	31974.350	33153.410	27232.250	19339.780
dloanorg	67.08	7.63	3.06	-5.37
nongse	0.350%	0.399%	0.395%	0.261%
denpct	19.940%	20.632%	21.954%	24.683%
loanpur	1.930%	1.942%	1.944%	1.935%

Borrower Characteristics

Credscr is credit score and is computed by a numerical expression based on a statistical analysis of a person's credit files. A credit score represents the creditworthiness of that person. Higher credit scores signify lower credit risk, and in turn, lower defaults and foreclosures. Figure 1 graphs credit score by state from 2004 to 2007. On average, the credit score for the United States was 678 in 2004, 677 in 2005, 676 in 2006, and 692 in 2007. This upward trend in credit scores continues in 2008 (683) and 2009 (686). The credit score data suggests that potential borrowers' creditworthiness has improved over time.

Figure 1: Credit Scores by State



Laginc is personal income lagged one year. Income is measured in current dollars and includes earnings received by all persons from all sources, including the sum of net earnings by place of residence, rental income of persons, personal dividend income, personal interest income, and personal current transfer receipts. Lagged income has increased progressively during the study period, i.e. 3.1 percent in 2004, 4.7 percent in 2006, and 5.6 percent in 2007.

Gwthinc is the growth rate in personal income. The higher the income growth rate, then the lower the credit risk, and in turn, lower defaults and foreclosures. The percent change in the growth rate of income is reflexive of the economic trends during the study period. The growth rate in incomes increases at an increasing rate between 2004 and 2006. But between

2006 and 2007, the growth rate for income declined. This latter trend is consistent with the start of the late 2000s Great Recession, which began in December 2007.

Loan Characteristics

Loanrat is the loan to income ratio and is computed as average loan amount divided by personal income. Loan to income ratio can be used as a crucial personal financial health indicator. This ratio is used here similar to the housing to income ratio to capture whether a person is qualified for a mortgage. Higher loan to income ratios mean that the proportion of the borrower's income that goes toward making mortgage payments is higher. This increases credit risk and the likelihood of foreclosure. The loan to income ratio increased in 2005 and 2006. Loan to income ratios for those years are 6.3 percent and 2.0 percent, respectively. From 2006 to 2007, the loan ratio is constant.

Loanratl is loan to income ratio lagged one year. Assuming that a sizeable number of foreclosures occur within the first year of mortgage origination, the higher the lagged loan to income ratio, then the greater the number of foreclosures within the first twelve months. The lagged loan ratio increases sharply between 2004 and 2005 (11.9 percent), then increases more gradually in 2006 (6.4 percent) and 2007 (2 percent).

LTV is the loan to value ratio, which expresses the amount of a first mortgage lien as a percentage of the total appraised value of real property. Higher LTV leads to lower equity investment by borrowers, which in turn increases credit risk, defaults, and foreclosures. LTV decreased by 2.1 percent in 2005 and increased in 2006 and 2007 by 2.9 percent and 2.1 percent, respectively.

Pctadj is the percentage of adjustable rate mortgages (ARMs). Adjustable mortgages have interest rate changes based on a standard rate index. Most ARMs have a cap on how much the interest rate may increase. The higher the percentage of ARMs, the greater the number of borrowers that are exposed to payment shocks. This increases credit risk and leads to higher defaults and foreclosures. The percentage of adjustable mortgages increased by 16.6 percent from 2004 to 2005. In 2006, the percent of ARMs peak at 17.27 percent and in 2007 decline slightly by 3.6 percent.

Nivpct is the percent of loans in which income is not reported. While HMDA rules make no specific reference to NIV applications per se, the NIV (No-Income Verification) characteristic can be implicitly derived from the absence of a reported value in the income field. It is also assumed that NIV loans do not use income in the underwriting decision. The higher the percentage of NIV loans, then the higher the percentage of borrowers that may have insufficient income to support loan repayment thereby increasing credit risk, defaults, and foreclosures. After close review, The percentage of NIVs basically declined in 2005 and increased in 2006. There is a similar .1 percent noted in the decline in 2007.

Nonocc is the percent of loans where the borrower does not occupy the home. The HMDA rules require lenders to report whether or not the borrower is occupying the home. Owner occupants are considered less risky and command the best terms, other things being equal. Non-owner occupants are considered to be real estate investors. The higher the percentage of non-owner occupancy, then the higher the chance of investor loan speculation, and rental vacancy risk thereby increasing credit risk, defaults, and foreclosures. The percentage of non-owner occupant loans peak in 2006 and decline by 1.5 percent in 2007.

Spread is the percent of loans with a spread at origination. The spread is reported in HMDA if the difference between the annual percentage rate (APR) and the applicable Treasury yield is equal to or greater than 3 percentage points for first-lien loans or 5 percentage points for subordinate-lien loans. A loan with a high spread indicates that the lender perceives the borrower has above-average risk. Figure (10) shows that the percentage of loans with a spread increase by 64.8 percent between 2004 and 2005. In 2006, loans with spreads peak and then decline by 36.9 percent in 2007.

Economic Conditions

Delhpidx is the growth in the Freddie Mac Conventional Mortgage Home Price Index (CMHPI). This index provides a measure of typical price inflation for houses within the US. Increases in HPI should increase the value of property and home equity, which in turn reduces risk of foreclosure and puts homeowners in a better position to refinance. The growth rate in the conventional mortgage home price index has been steadily declining i.e., down 12.4 percent in 2005, 6.3 percent in 2006, and 1.3 percent in 2007. The increase in home prices indicates a gain of home equity between 2004 and 2007.

Unempl is the unemployment rate representing the number unemployed as a percent of the labor force. The unemployment rate declined by 7.3 percent between 2004 and 2005. It declined again from 2005 to 2006 by 9.2 percent. As the economy started to slow down, the unemployment rate increased between 2006 and 2007 by 8.8 percent.

Constart is the number of residential building construction projects begun during a given year. Construction starts is a closely-watched indicator of the health of the home construction industry. Homebuilding is a major component of the overall economy due to its size and its linkages with finance, basic materials, home furnishings and appliances, employment, and overall growth. As construction investments increase, home prices, home values, and equity increase which helps decrease

the risk of defaults and foreclosures. Construction starts peak in 2005 and steadily decline in 2006 and 2007, by 17.6 percent and 29 percent respectively. This downward trend in construction serves as a barometer that the economy moved toward a recession.

Lending Conditions

Dloanorg is the rate of change in the percentage of conventional loans for home purchase. A conventional mortgage is a lender agreement to provide funds for property that is not guaranteed or insured by the federal government under the Veterans Administration, the Federal Housing Administration, or the Rural Housing Service of the U.S. Department of Agriculture. The higher the rate of change in mortgage loans for home purchase, then the greater the pool of loans that may face foreclosure. The rate of change in the percentage of conventional loans for home purchase peaks in 2004. Then, the rate of change declines dramatically from 2005 to 2007, i.e., by 88.6 percent in 2005, 59.9 percent in 2006, and 275 percent in 2007.

Nongse is the percent of loans sold to Non-Government Sponsored Enterprises (Non-GSE) during a given year. Loans sold is an indicator of secondary market activity in the mortgage market. HMDA requires disclosure of loans sold by each institution and the purchaser of such loans by class, e.g., Fannie Mae, Ginnie Mae, Freddie Mac, Farmer Mac Code, Private Securitization, Commercial Bank, Savings Bank or Savings Association, Life Insurance Company, Credit Union, Mortgage Bank, or Finance Company, Affiliate Institution, and other type of purchase. Higher percentages of loans sold only to Non-Government Sponsored Enterprises can serve as an indicator that lenders are originating less risky loans. Therefore, in general there are less toxic loans in circulation as a result of the higher percent of loans sold to Non-GSEs and the lower the risk of foreclosure. It is assumed that rates on GSE loans are subsidized and terms must conform to standards while higher non-conforming might mean less consumer oriented terms, i.e., subprime. The percentage of loans sold after 2005 decline. From 2005 to 2007, the percent of loans sold decrease, i.e., by no change in 2006 and 4.2 percent in 2007. These declines signal that secondary mortgage market activity slowed.

Denpct is the percent of applications that were not approved. Higher denial rates suggest higher credit standards, lower credit risk, and lower defaults. There are, however, divergent patterns of denial rates for prime and subprime lenders. Subprime lenders attract many applications from less creditworthy individuals resulting in high denial rates. To the extent that this variable includes subprime activity it may not have the inverse relationship as expected. The percentage of applications denied increase at an increasing rate over the time period studied. From 2004 to 2005 percent of applications denied increase by 3.5 percent and from 2005 to 2006 by 6.4 percent. From 2006 to 2007, the percent of applications denied swell to 24.68 percent which is a 12.45 percent increase from 2006.

Loanpur is the percent of loans purchased by lending institutions during a given year. Loan purchases are a barometer of correspondent lending behavior or wholesale activity in the mortgage market. A correspondent lender delivers loans to a wholesale lender against prior price commitments the wholesaler has made to the correspondent. The commitment protects the correspondent against pipeline risk. Correspondent lending could increase the likelihood of fraud and inflated values and in turn increase foreclosures. There is basically a constant movement in the percentage of loans purchased between 2004 and 2007. The percent change from year to year is less than one percent whether that change is an increase or decrease.

Empirical Analysis

Our review of the results of the model estimations focuses on identifying factors that explain foreclosure rates for all loans, prime loans, and subprime loans. For each type of loan, six models are estimated using different combinations of the independent variables that represent borrower characteristics, loan characteristics, economic conditions, and lending conditions.

Table 3 provides the regression results for all loans. For all models, R^2 exceeds 0.97. The variables that significantly explain foreclosure rates and exhibit the expected direction of influence across models are lagged personal income (borrower characteristic); growth in the housing price index, construction starts, and unemployment rate (economic conditions); and loans purchased by non-GSEs (lending condition). None of the variables that represent loan characteristics significantly influenced foreclosure rates in the way that the authors expected. However, percent of NIV loans and loan ratio (loan characteristics); rate of change in the percent of loans for home purchase and denial rates (lending conditions); and growth in personal income (borrower characteristic) significantly explain foreclosure rates. The regression results suggest the following:

- Foreclosure rates decrease as the percent of NIV loans increases. During the sample period, NIV loans were only approved for borrowers with very high credit scores and sizeable down payments. These loans were considered to be very high quality with low levels of default risk. Therefore, the regression results seem reasonable.
- Foreclosure rates decrease as loan ratio increases. Loan ratio is loan amount/income. Because other measures of income are included in the models, this result may be spurious.

- Foreclosure rates decrease as the rate of change in the percent of loans for home purchase increases. This relationship is marginally significant (10 percent level) in only one of the models (Model 1).
- Foreclosure rates increase as denial rates increase. There are divergent denial rates for prime and subprime lenders. Subprime lenders attract many applications from less creditworthy borrowers. Thus, denials are higher for subprime loans. Increased denial rates capture increased subprime mortgage activity.
- Foreclosure rates increase as growth in personal income increases. This relationship may signify that high income borrowers acquired very high levels of mortgage debt during the sample period, a period of rising house prices.

Table 3: All Loans: Foreclosure Starts (Model Coefficients and T-stats)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
credscr	-0.0660	-0.0748	-0.0656	-0.0689	-0.0898	-0.0911
	-1.21	-1.21	-1.06	-1.14	-1.50	-1.57
laginc	-0.0400	-0.0300		-0.0300		-0.0300
	-2.76 ***	-1.69 *		-1.77 *		-1.91 *
gwthinc					1.3768	
					3.34 ***	
loanrat	-0.2101					
	-5.70 ***					
loanratl		-0.0175	-0.0191			
		-0.72	-0.79			
ltv	-0.0047	-0.0057	-0.0056	-0.0053	-0.0041	
	-1.17	-1.26	-1.22	-1.21	-0.95	
pctadj	0.0101	-0.0054	-0.0062	-0.0062	-0.0049	0.0004
	1.35	-0.70	-0.80	-0.81	-0.63	0.05
nivpct	-0.0276	-0.0333	-0.0349	-0.0330	-0.0305	-0.0198
	-2.30 **	-2.50 **	-2.61 ***	-2.49 **	-2.27 **	-1.55
nonocc	0.0007	0.0000	0.0004	0.0006	0.0015	
	0.24	0.00	0.12	0.19	0.46	
spread					-0.0018	-0.0087
					-0.28	-1.61
delhpidx	-0.0151	-0.0133	-0.0116	-0.0135	-0.0112	-0.0075
		**		**	**	
	-5.97 ***	-4.75 *	-4.40 ***	-4.85 *	-4.09 *	-2.34 **
unempl	0.0203	0.0279	0.0361	0.0253	0.0492	
	1.04	1.27	1.68 *	1.17	2.33 **	
constart	-0.0027	-0.0030	-0.0033	-0.0031	-0.0032	-0.0032
		**		**	**	
	-2.91 ***	-2.96 *	-3.26 ***	-3.04 *	-3.19 *	-3.38 ***
dloanorg	-0.0019	-0.0002	-0.0001	-0.0002		
	-1.68 *	-0.15	-0.11	-0.14		
nongse	-1.3789	-1.1457	-1.0685	-1.1722	-1.0400	
		**		**		
	-4.22 ***	-3.15 *	-2.94 ***	-3.26 *	-2.53 **	
denpct						0.0349
						4.34 ***
denpct_1		0.0074	0.0069			
		0.93	0.86			
loanpur					-0.0276	
					-0.94	
R ²	0.981	0.977	0.977	0.977	0.978	0.978

*** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level

Table 4 provides the regression results for prime loans. For all models, R^2 exceeds 0.97. The variables that significantly explain foreclosure rates and exhibit the expected direction of influence across models are credit score and lagged personal income (borrower characteristics); growth in the housing price index, construction starts, and unemployment (economic conditions); and loans purchased by non-GSEs (lending condition). Also, percent of NIV loans and loan ratio (loan characteristics); rate of change in the percent of loans for home purchase and denial rates (lending conditions); and growth in personal income (borrower characteristic) significantly explain foreclosure rates. The direction of influence for these variables was discussed in the previous section. Overall, the results are similar to those for all loans. The exception is that credit score significantly influences prime loan foreclosure rates but, does not influence foreclosure rates of all loans. This result seems reasonable because credit histories are likely impaired for subprime borrowers. Therefore, credit score information is not very useful for explaining foreclosure rates that include subprime borrowers.

Table 4: Prime Loans: Foreclosure Starts (Model Coefficients and T-stats)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
credscr	-0.0613 -2.11 **	-0.0650 -1.97 **	-0.0607 -1.84 *	-0.0629 -1.94 *	-0.0755 -2.34 **	-0.0703 -2.24 **
laginc	-0.0200 -2.54 **	-0.0100		-0.0100 -1.54		-0.0100 -1.79 ***
gwthinc					0.6835 3.08 ***	
loanrat	-0.1164 -5.92 ***					
loanratl		-0.0064 -0.49	-0.0071 -0.55			
ltv	-0.0016 -0.75	-0.00212 -0.87	-0.0021 -0.85	-0.0020 -0.83	-0.0013 -0.57	
pctadj	0.0052 1.3	-0.0034 -0.81	-0.0037 -0.89	-0.0039 -0.94	-0.0030 -0.71	-0.0018 -0.45
nivpct	-0.0128 -1.99 **	-0.0160 -2.24 **	-0.0168 -2.34 **	-0.0158 -2.22 **	-0.0138 -1.91 ***	-0.0100 -1.45
nonocc	0.0007 0.45	0.0002 0.12	0.0004 0.23	0.0006 0.37	0.0009 0.55	
spread					-0.0006 -0.17	-0.0027 -0.92
delhpidx	-0.0066 -4.92 ***	-0.0056 -3.74 ***	-0.0048 -3.41 ***	-0.0057 -3.84 ***	-0.0045 -3.05 ***	-0.0036 -2.05 **
unempl	0.0107 1.02	0.0147 1.25	0.0186 1.62	0.0134 1.16	0.0252 2.22 **	
constart	-0.0021 -4.29 ***	-0.0023 -4.22 ***	-0.0025 -4.51 ***	-0.0023 -4.28 ***	-0.0023 -4.31 ***	-0.0024 -4.57 ***
dloanorg	-0.0010 -1.57	0.0000 0.02	0.0000 0.06	0.0000 0.02		
nongse	-0.4832 -2.77 ***	-0.3523 -1.81 *	-0.3160 -1.63	-0.3691 -1.91 *	-0.3141 -1.42	
denpct						0.0134 3.09 ***
denpct_l		0.0058 1.35	0.0056 1.2900			
loanpur					-0.0190 -1.20	
R^2	0.978	0.973	0.972	0.972	0.974	0.973

*** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level

Table 5 provides the regression results for subprime loans. The variables that significantly explain foreclosure rates and exhibit the expected direction of influence across models are lagged personal income, growth in the housing price index, and loans purchased by non-GSEs. Also, percent of NIV loans, loan ratio, loan to value ratio, and loans with a reportable spread, denial rates and growth in personal income significantly explain foreclosure rates. The regression results suggest:

- Foreclosure rates decrease as the loan to value ratio increases. During the sample period, house valuations were inflated. Therefore, loan to value ratios may have been artificially deflated during periods of increasing foreclosures.
- Foreclosure rates decrease as the percent of loans with a reportable spread increase. Loans with a reportable spread indicate that the lender perceives that the borrower has above-average risk. The spread compensates the lender for the increased default risk. This finding may signify that lenders properly identify such borrowers and provide proper instruments when the lender is compensated with a spread.

Table 5: Subprime Loans: Foreclosure Starts (Model Coefficients and T-stats)

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
credscr	0.0943	0.0856	0.1087	0.0835	0.0393	0.0323
	0.44	0.36	0.45	0.35	0.17	0.14
laginc	-0.1200	-0.0700		-0.0700		-0.0600
	-2.08 **	-1.09		-1.23		-1.10
gwthinc					3.2540	
					1.99 **	
loanrat	-0.7675					
	-5.25 ***					
loanratl		-0.0272	-0.0314			
		-0.29	-0.33			
ltv	-0.0378	-0.0389	-0.0385	-0.0401	-0.0377	
	-2.39 **	-2.20 **	-2.18 **	-2.33 **	-2.21 **	
pctadj	0.0417	-0.0174	-0.0195	-0.0181	-0.0058	0.0195
	1.40	-0.57	-0.64	-0.60	-0.19	0.67
nivpct	-0.1162	-0.1369	-0.1409	-0.1362	-0.1188	-0.0827
	-2.44 **	-2.64 ***	-2.72 ***	-2.63 ***	-2.23 **	-1.65 *
nonocc	-0.0089	-0.0111	-0.0102	-0.0093	-0.0068	
	-0.77	-0.87	-0.80	-0.74	-0.54	
spread					-0.0235	-0.0578
					-0.92	-2.71 ***
delhpidx	-0.0605	-0.0532	-0.0489	-0.0546	-0.0489	-0.0302
	-6.05 ***	-4.88 ***	-4.81 ***	-5.04 ***	-4.52 ***	-2.39 **
unempl	0.0842	0.1088	0.1296	0.1024	0.1594	
	1.09	1.28	1.56	1.21	1.91	
constart	-0.0033	-0.0048	-0.0055	-0.0048	-0.0045	-0.0054
	-0.90	-1.21	-1.41	-1.20	-1.14	-1.43
dloanorg	-0.0057	0.0013	0.0014	0.0006		
	-1.27	0.26	0.29	0.13		
nongse	-6.7128	-5.9643	-5.7733	-5.9585	-4.9172	
	-5.19 ***	-4.21 ***	-4.11 ***	-4.24 ***	-3.02 ***	
denpct						0.1562
						4.93 ***
denpct_l		0.0335	0.0323			
		1.08	1.04			
loanpur					-0.1378	
					-1.18	
R ²	0.982	0.979	0.978	0.978	0.979	0.979

*** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level

Overall, the factors that significantly influence foreclosure rates are:

- Borrower characteristics - lagged personal income and credit score (prime borrowers)
- Loan characteristics – percent NIV loans
- Economic conditions - growth in the housing price index, construction starts, and unemployment rate
- Lending conditions – loans purchased by non-GSEs and denial rates

These results suggest that foreclosures can be mitigated with appropriate underwriting standards. Borrower income and credit history as well as the general health of the economy remain important in determining default risk. In addition, the affect that NIV loans have on foreclosure rates during the sample period suggests that significant down payment requirements can be used to effectively reduce default risk. Finally, the variables related to lending conditions suggest that rapid growth in the supply of subprime mortgages increase foreclosure rates. During periods of rapid subprime mortgage growth, lenders made it easier for borrowers with low credit scores to obtain loans with little or no down payment, terms beyond thirty years, or low initial teaser rates. Lenders also allowed borrowers to provide little or no proof of income or assets. While this relaxed underwriting occurred partially to respond to the challenge of maintaining the affordability of homeownership during periods of rapid house price appreciation, these actions contributed to the foreclosure crisis we currently face.

Conclusions

Problems in the subprime market and the resulting foreclosure crisis have led to complex recommendations for legislative and regulatory reform. These responses should not overly restrict consumer access to mortgage credit on fair and appropriate terms. The Obama Administration's efforts to curbe the devastating effects of forced foreclosure have been introduced in the President's Home Affordable Modification "Relief" Plan of March 4, 2009. The Plan addresses many issues presented in this paper. For mortgage loans that are in imminent default, loan modifications will be used as an effective tool to minimize losses to investors and help borrowers avoid foreclosure. The Plan reduces individual mortgage interest rates by as much as 200 basis points, extends the term of many of the mortgages up to as many as 40 years and, where possible, allows for forbearance or forgiveness of portions of mortgage principal balances to allow some mortgage payments to equal to thirty-one percent of the monthly household income. The incentive of the plan is to bring "responsible" homeowners to a current status on their mortgage loans and allow them to remain in their homes, to reduce the large number of foreclosures, and to assist homeowners in maintaining some remnants of their prior credit standing.

Although the Plan was implemented to assist in Fannie Mae and Freddie Mac mortgages, many non-GSE lenders are voluntarily participating in the program. Generally, GSE borrowers would be unable to refinance their loans because their homes have lost significant value, which essentially would cause their current loan-to-value ratios to exceed 80 percent. Since many of the GSE lenders and servicers currently have much of the borrower's information already, then documentation requirements will be greatly reduced. This should allow borrowers to refinance their loans more quickly while simultaneously reducing costs faced by both borrowers and lenders. This Plan should assist several million homeowners who have solid mortgage payment histories on their existing homes to remain in them and avoid foreclosure.

References

- Ambrose, Brent W. and Charles A. Capone, "Modeling the Conditional Probability of Foreclosure in the Context of Single-Family Mortgage Default Resolutions", *Real Estate Economics*, Fall 1998
- Campbell, Tim S. and J. Kimball Dietrich, "The Determinants of Default on Insured Conventional Residential Mortgage Loans", *Journal of Finance*, December 1983
- Capone, C.A., *Research Into Mortgage Default and Affordable Housing: A Primer*, Local Initiatives Support Corporation, 2002
- Case, K.E. and R. J. Shiller, "Mortgage Default Risk and Real Estate Prices: The Use of Index-Based Futures and Options in Real Estate", *Journal of Housing Research*, 1995
- Clauret, T. M., "The Impact of Interstate Foreclosure Cost Differences and the Value of Mortgages on Default Rates", *Journal of the American Real Estate and Urban Economics Association*, 1987
- Cunningham, D.F. and C.A. Capone, "The Relative Termination Experience of Adjustable to Fixed-Rate Mortgages", *The Journal of Finance*, 1990
- Deng, Yongheng, Quigley, John M. and Robert Van Order, "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options", *Econometrics*, Mar. 2000
- Foster, C. and R. Van Order, "An Option-Based Model of Mortgage Default", *Housing Finance Review*, 1984
- Foster, Chester and Robert Van Order, "FHA Terminations: A Prelude to Rational Mortgage Pricing", *AREUEA Journal*, 1985
- Greene, W.H., *Econometric Analysis*, 5th ed. Upper Saddle River, NJ: Prentice Hall, 2003
- Hayre, Lakhbir S., Saraf, Manish, Young, Robert, and Jiakai (David) Chen, "Modeling of Mortgage Defaults", *Journal of Fixed Income*, Spring 2008
- Hendershott, P.H. and E.J. Kane, "Causes and Consequences of the 1980s Commercial Construction Boom", *Journal of Applied Corporate Finance*, 1992
- Herzog, J.P. and J.S. Earley, *Home Mortgage Delinquency and Foreclosure*, Cambridge, MA: National Bureau of Economic Research, 1970
- Lea, M.J. and P.M. Zorn, "Adjustable-Rate Mortgages, Economic Fluctuations, and Lender Portfolio Change", *Journal of the American Real Estate and Urban Economics Association*, 1986
- Merry, Ellen A. and Michael D. Wilson, *The Geography of Mortgage Delinquency*, Federal Reserve Board of Governors, September 12, 2006
- Mian, Atif and Amir Sufi, "The Consequences of Mortgage Credit Expansion: Evidence from the US Mortgage Default Crisis", *The Quarterly Journal of Economics*, November 2009
- Pennington-Cross, Anthony, "Credit History and the Performance of Prime and Nonprime Mortgages", *Journal of Real Estate Finance and Economics*, Nov. 2003
- Schwartz, Eduardo S. and Walter N. Torous, "Mortgage Prepayment and Default Decisions: A Poisson Regression Approach", *Journal of the American Real Estate and Urban Economics Association*, 1993
- Stevenson, B.G., "Cyclical Capital Flows and the Cycles of Losses in Commercial Real Estate", *Real Estate Review*, 1995
- Teo, Alan, "Delinquency Risk In Residential ARMs: A Hazard Function Approach", *Journal of Real Estate Portfolio Management*, Sep-Dec 2004
- Vandell, K.D. and T. Thibodeau, "Estimation of Mortgage Defaults Using Disaggregate Loan History Data", *Journal of the American Real Estate and Urban Economics Association*, 1985
- VanderHoff, James, "Adjustable and Fixed Rate Mortgage Termination, Option Values and Local Market Conditions: An Empirical Analysis", *Real Estate Economics*, Fall 1996
- Zorn, P.M. and M.J. Lea, "Mortgage Borrower Repayment Behaviour: A Microeconomic Analysis with Canadian Adjustable Rate Mortgage Data", *Journal of the American Real Estate and Urban Economics Association*, 1989.

A Panel Model of the Interest Sensitivity of Net Interest Margins: 1984—2009

Albert E. DePrince, Jr., Middle Tennessee State University

Pamela D. Morris, Kaplan University

Abstract

Textbooks typically argue a positive relationship between the slope of the yield curve and net interest margin (NIM). However, some have found that banks have successfully neutralized interest rate risk. At the same time, others have shown that NIM is affected by the slope of the yield curve. These studies use panel model methodology but undertake various levels of aggregation in their analysis. Left unanswered is the link on a bank-by-bank basis between the slope of the yield curve and NIM for each. This study fills that void. It develops an unbalanced panel model using all banks with reported data for the relevant variables during the 1984-2009 sample period. This screen produced a sample of 19,641 banks in the final version. Results show that the NIM is positively related to the slope of the yield curve, and economic growth is inversely related to credit risk, concentration, and asset size. Additionally, the results show that responses of the bank NIMs are not instantaneous; rather, 85 to 98 percent of the responses to interest rate changes are completed in about one year. Finally, characteristics unique to each bank, or fixed effects, are an important determinant of each institution's NIM.

Introduction

As short-term interest rates rose during 2004 and 2006, concern over adverse effects of a flattened yield curve on net interest margins (NIMs) mounted. Indeed, an earlier longitudinal study (DePrince and Morris, 2007) showed that fears of reduced earnings stemming from a flattened yield curve were well founded. The slope and rotation of the yield curve were found to affect NIM.

This study extends that earlier study. DePrince and Morris (2007) stratified banks into one of five asset sizes and treated each asset size as a panel. Industry-wide average NIM was calculated for each of the five panels. The stratification was designed to capture effects of scale on NIM. In contrast, this study uses all banks in existence between 1984 and 2009 as panels, and scale is taken into account directly by each institution's asset size.

This study turns first to a background discussion on net interest margins, explaining the basic assumptions of forces affecting NIM. This is followed by a summary of recent research on NIM. The study's model is then developed, and the model is compared with earlier studies. Next, data and data methodologies are addressed, followed by the estimation and evaluation of the model. Key findings are discussed, and the paper ends with possible extensions of the study.

This study finds that NIM is affected by the level of interest rates, the slope of the yield curve, credit quality, asset concentration, and scale. Additionally, the results show that the responses of bank NIMs are not instantaneous; rather, completion of 85 to 98 percent of the adjustment takes about a year, depending on the model. Finally, characteristics unique to each bank, or fixed effects, are an important determinant of each institution's NIM.

Background on Net Interest Margins

Customarily, net interest margins are presumed to be inversely related to the level of interest rates. This relationship has four aspects. The first is repricing risk, which stems from a mismatch in the maturity structure of assets and liabilities. The presumption here is that banks are liability sensitive, meaning that liabilities reprice faster than assets as interest rates change. This is also reflected in the positive duration gap typical among banks. The sensitivity is presumed to be amplified by the rotation in the yield curve as interest rates change. In a rising rate environment, the curve flattens and may even invert. In a falling rate environment, the curve, if inverted, becomes positively sloped. Once positively sloped, it typically steepens as rates fall. This rotation is the second source of interest rate risk and is commonly known as yield curve risk.

The link between the other two aspects of interest rate risk and the inverse relationship between interest rate movements and NIM is less clear-cut. Basis risk occurs when there is imperfect correlation in the timing of interest rate changes for assets and liabilities. Banks may use the timing difference to amplify or moderate effects of the first two factors on NIM. The last is optionality risk. Typically, this stems from options embedded in mortgage-related instruments, but options may also be embedded in liabilities. Large institutions may also incur optionality risk as writers of over-the-counter interest rate option contracts and through exchange-traded options held in their trading account (Basel Committee on Banking Supervision).

This inverse relationship may be moderated or amplified by the phase of the business cycle. Other things equal, one would conceptually expect the NIM to be procyclical, rising in an economic boom and falling in recessions. Expanding loan demand in an upswing lessens competitive pressures, enabling banks to widen spreads. However, some argue as noted below that the state of the economy has a countercyclical effect on net interest margins.

Interest rate risk is under close regulatory scrutiny. Banks are required to measure interest rate risk, set limits and monitor such risk, and undertake stress tests. Bank size may play a role. Regulatory scrutiny is closer for the big banks (Basel Committee on Bank Supervision, 2003), reflected in the changes to the 1988 Basel capital standard applicable to the world's largest banks. While smaller banks are under less intense scrutiny, they cannot ignore interest rate risk.

Research on Net Interest Margins

Banks do not face interest rate changes passively but manage assets and liabilities to maximize net interest income, consistent with the interest rate, credit, and liquidity risks deemed acceptable. Due to the active management of asset and liability portfolios, the presumed inverse relationship between interest rates and net interest margin may not materialize. In fact, this point has been noted elsewhere. English (2002) examined 10 countries and found that banks were remarkably successful in isolating net interest margins from interest rate movements.

Recent research on net interest income has taken a number of different routes. Some studies simply present a clinical description of developments. This is seen in the quarterly Federal Deposit Insurance Corporation and Federal Reserve reports (FDIC, 2003; Carlson and Perli, 2004). Others examine net interest margins to assess their vulnerability to interest rate movements (English, 2002) or to changes in capital requirements (Zarruk and Madura, 1992). Beyond clinical descriptions, investigations into net interest margins have been sparse.

Two broad models of net interest margins have emerged: the dealer model and the cost of goods model (Hanweck and Ryu, 2005). The earliest cross-sectional study used the dealer model (Ho and Saunders, 1981). This study utilized a four-part explanation of a bank's interest margin, setting it equal to a pure spread, a markup for interest expense, the opportunity cost of reserves, and a default premium. It used a sample of 100 banks in which each bank was treated as a dealer demanding one type of deposit and supplying one type of loan. Emphasis was on estimating the pure spread. Subsequent studies broadened this approach to heterogeneous loans (Allen, 1988) as well as other countries (Saunders and Schumacher, 2000). Allen (1988) made a noteworthy observation that the introduction of heterogeneous loans provided a portfolio benefit to risk management.

Over time, the cross-sectional characteristics of the investigations expanded as the use of call report data became more common. Studies broadened the investigation by separating banks by asset size (Angbazo, 1997) and allowed for the isolation of the effects of interest rate risk and credit risk on margins by asset size.

The Hanweck and Ryu (2005) study is the most ambitious. The authors separate banks by business-line specialization into six groups. Within each, individual banks serve as the cross-sectional variable. The model for each group is estimated separately, and Hanweck and Ryu assess variations in net interest margin through variables accounting for interest rate risk, term structure risk, and credit risk. They use other institutional variables and seasonal variables to account for business risk. The same model is used for all six business-line specializations. The business-line specialization indirectly takes account of size. For example, the international banks would all likely fall among the nation's largest banks, while agriculture banks would be among the smallest.

The cost-of-goods model focused on the supply of deposits and addresses the cost of goods by introducing uncertainty into the supply function (Zarruk, 1989). Later, uncertainty was introduced through loan losses, deposit insurance, and capital regulation (Zarruk and Madura, 1992).

More recently, Aliaga-Diaz and Olivera (2005) examined the cyclical aspect of NIM and built on the earlier work of Deuker and Thorton (1997), which concluded that loan pricing, as reflected in the prime rate, exhibited a countercyclical pattern. This was due largely to "switching costs." Aliaga-Diaz and Olivera went beyond the prime rate and examined the cyclical behavior of NIM. They developed a time series on NIM from individual call reports and found that the business cycle exerts a countercyclical effect on NIM.

The current study takes a different approach from both the Angbazo and the Hanweck and Ryu studies. As with the Aliaga-Diaz and Olivero study, it examines the effects of interest rates and general business conditions on net interest margins of all banks. It finds that all institutions react in a similar manner to the variables within the model. Unique difference among banks is captured through fixed effects. This stands in contrast to Hanwick and Ryu, in which various types of banks were the cross-sectional variables; Aliaga-Diaz and Olivero, in which various income and expense items were aggregated across banks as industry-wide measures per period; and DePrince and Morris (2007), in which banks were aggregated into asset classes.

The Data

The study includes (1) commercial bank, national (federal) charter and Fed member, supervised by the Office of the Comptroller of the Currency (OCC); (2) commercial bank, state charter and Fed member, supervised by the Federal Reserve (FRB); (3) commercial bank, state charter and Fed nonmember, supervised by the Federal Deposit Insurance Corporation (FDIC); (4) savings banks, state charter, supervised by the FDIC; and (5) savings associations, state or federal charter, supervised by the Office of Thrift Supervision (OTS). The study excludes insured U.S. branches of foreign-chartered institutions since they are branches and not banks.

The data were obtained from the quarterly call reports for each institution for each year from 1Q1984 through 1Q2009. Banks were included if they were present at one time or another during the sample period. Since banks appeared and disappeared during the sample period, blanks were inserted if there were no data for a given bank in a given quarter. NIM for each bank was calculated as the ratio of net interest income to average assets for each quarter. Prior to the calculation of NIMs, data were purged of extreme outliers involving interest income, interest expense, and assets.

GDP and employment data were obtained from online databases of the Bureau of Economic Analysis and the Bureau of Labor Statistics, respectively. The various interest rates were obtained from the Federal Reserve's online database of historic data from the weekly H15 statistical releases. Quarterly averages of the interest rates and employment data were calculated and compound annual growth rates of the economic variables computed.

Before proceeding with a discussion of the basic model, it is useful to review the general characteristics of the NIM data. This task is complicated, since the study deals with a total of nearly 20,000 banks that were in existence at one time or another during the sample period. To simplify matter, only the first and last quarter of the sample period and a quarter roughly mid-way through the sample are used. To illustrate the relationship between size and NIM, the sample is split into two broad groups: small and large banks. Small banks are defined as any bank with assets of one billion dollars or less in the first quarter of 2009. Of the 7,528 banks with usable observations, 6,951 (92.34% of the total) met this criteria. The remaining 577 banks (7.66%) represent large banks. Moving back in time to the other two quarters, large banks are the largest 7.66% of the banks with usable observations in both quarters, and small banks are the remaining banks. Results are reported in Table 1. Admittedly this approach is a gross simplification, but it does provide a useful overview.

Table 1: Summary Data for Net Interest Margin by Asset Class

Period	Class	Max Asset for Small Banks*	Average Class NIM	Average Class Assets*	Number of Banks
1984Q1	Total				15,352
	Large		3.68%	\$ 1,754,081	1,176
	Small	\$ 235,551	4.21%	\$ 45,702	14,176
1997Q1	Total				10,071
	Large		4.13%	\$ 5,128,511	771
	Small	\$ 499,700	4.25%	\$ 93,453	9,300
2009Q1	Total				7,528
	Large		3.07%	\$ 19,256,335	577
	Small	\$ 1,000,000	3.40%	\$ 196,622	6,951

*Data expressed in thousands of dollars.

In reviewing Table 1, the reader quickly sees the dramatic consolidation in banking, as the number of banks was roughly cut in half over the sample period. Also, it is quite evident that size of big banks exploded over the sample period, as the average size of the largest banks rose from roughly \$1.8 billion in 1984 to \$19.3 billion in 2009. In contrast, the increase in the size of small banks was less dramatic, rising from roughly \$45 million to \$197 million over the 25-year period. Finally, as hypothesized in the basic model, there is a visible inverse relationship between size and NIM, and shifts in the NIM between periods would be explained by the various factors in the basic model.

The Basic Model

Each bank is treated as a panel in this study. The general hypothesis is that net interest margin (*NIM*) of each bank or panel is affected by economic factors (*econ*), financial factors (*fin*), and industry structure (*is*). A general model can be specified as

$$NIM_{i,t} = \beta_{0,i} + \beta_{j,i}econ_{j,i,t} + \beta_{k,i}fin_{k,i,t} + \beta_{l,i}is_{l,i,t} + \mu_{i,t} \quad (1)$$

where NIM = net interest margin,
 i = 1, ..., n and represents the banks in the sample,
 j = 1, ..., m and represents the various economic variables,
 k = 1, ..., q and represents the various financial variables, and
 l = 1, ..., s and represents the industry structure variable.

This is a panel model in which the individual institutions serve as the panels as well as the cross-section identifiers. Thus, *NIM* of each institution is assumed to depend on factors unique to each institution. These may be unidentified or unobserved factors and may be either time-invariant (fixed effects) or time-variant (random effects). The $\beta_{0,i}$ coefficient now represents unidentified factors unique to each bank (or panel) that influence the dependent variable ($NIM_{i,t}$). To account for these unidentified factors, longitudinal models employ concepts known as fixed effects and random effects. Fixed effects are portrayed as different constant terms for each state and are presumed to capture the effects of variables that were omitted from the estimated function, either because data were not available or because the investigators failed to identify such variables. These omitted variables are assumed to be time-invariant across the sample period. Random-effects coefficients also resemble separate constant terms, but these represent effects of time-varying variables.

If the economic, financial, or structure variables have a common effect across all banks, the coefficient would be the same for all n banks or panels. For example, if the first economic factor has only a common and not a cross-sectional effect, $\beta_{1,1} = \beta_{1,2} = \dots = \beta_{1,n}$ and becomes β_1 .

Equation 1 above represents a reduced-form equation that begins with the simple identity that, for any bank,

$$NIM_{j,t} = \frac{Interest\ Income_{j,t}}{Average\ Assets_{j,t}} - \frac{Interest\ Expense_{j,t}}{Average\ Assets_{j,t}}, \quad (2)$$

$$\frac{Interest\ Income_{j,t}}{Average\ Assets_{j,t}} = F(Lending\ Rates, Risk\ Premiums, Market\ Structure), \quad (2a)$$

and

$$\frac{Interest\ Expense_{j,t}}{Average\ Assets_{j,t}} = F(Lending\ Rates, Risk\ Premiums, Market\ Structure). \quad (2b)$$

In other words, translating Equation 2 into a behavioral regime, interest income (Equation 2a) depends on lending rates, risk premiums, market structure, and loan volume. Market structure depends, in turn, upon concentration and scale. Loan volume depends, in turn, on the state of the economy. Similarly, interest expense depends on borrowing rates, risk premiums, market conditions, and loan volume. Market structure depends, in turn, upon concentration and scale, while loan volume depends, in turn, on the state of the economy. Scale has been shown elsewhere to have an inverse effect on *NIM* (DePrince and Morris, 2007; FDIC, 2009).

In collapsing Equations 2a and 2b into Equation 2, *GDP* growth and employment growth (*EMP*) capture the effects of the economy on loan volume; the slope of the yield curve (*YC*) is used to approximate the difference between lending and borrowing rates. The general hypothesis would see a positive coefficient on the yield curve slope, reflecting the view that net interest income should fall as the yield spread becomes less positive or increasingly negative. The change in the slope directly captures the effects of the rotations of the yield curve. If the change is positive, the yield curve is becoming more positively or less negatively sloped, and this should have a positive effect on *NIM*. The reverse would hold if the change is negative. The spread between AAA and BAA bonds approximates risk premiums (*RISK*). Two variables capture industry structure. The Herfindahl-Hirschman Index (*HHI*) approximates concentration, and “average assets of each bank” captures scale.

Finally, the model also assumes that industry structure can be represented by a concentration measure and a scale measure. Thus, equation 2 becomes

$$NIM_{i,t} = \beta_{0,i} + \beta_{1,i}GDP_{i,t} + \beta_{2,i}EMP_{i,t} + \beta_{3,i}YC_{i,t} + \beta_{4,i}\Delta YC_{i,t} + \beta_{5,i}RISK_{i,t} + \beta_{6,i}HHI_{i,t} + \beta_{7,i}SCALE_{i,t} + \mu_{i,t} \quad (3)$$

Empirical Results

Preliminary screening equations indicated that common, rather than cross cross-sectional, coefficients were appropriate for the financial, economic, and structure data. Thus, the estimation results contain common coefficients for these variables in all three variations of Equation 3, reported in Table 1. Model 1 reports results without fixed effects. Model 2 reports results with fixed effects, and Model 3 reports results for random effects, or fixed effects and a partial adjustment process. Of the 22,000 or so banks that existed over the 25-year sample period, 19,641 had sufficient data to be recognized by EVIEWS 7, the statistical software used in the estimation phase. The final data set is an unbalanced panel with 1,056,524 observations.

Turning first to Model 1, one quickly sees that the adjusted R^2 is 0.39, and all coefficients differ from zero with virtual certainty (p-value < 0.001). More important, the coefficient on the proxy variable for the yield curve effects (the long-term less the short-term rate) is the appropriate sign, pointing to an increase in the NIM as the curve steepens and vice versa. The risk premium (the BAA less the AAA rate) has a negative sign, suggesting that as credit quality deteriorates (i.e., the spread widens), NIM is adversely affected. NIMs are inversely related to bank size, consistent with general findings regarding size and NIM, and NIMs are positively related to economic momentum. Faster growth is associated with higher NIM and vice versa.

Contrary to expectations, the concentration ratio has a negative effect on the NIMs. In other words, higher nationwide concentration reduces pricing power rather than increasing it. While surprising, this could reflect the link between size and concentration in a national market. As the big become bigger, concentration also rises. Thus, the HHI could reflect the same information as bank size in the model. At a local market level, however, research shows a positive link between concentration and performance, which can be attributed to stronger pricing power in concentrated markets (DePrince and Morris, 2010). The findings in the present study point to a break between local and national markets. In a national market, this ratio may not reflect pricing power (an issue best left to local markets) but rather other forces such as possible diseconomies associated with the increasing number of very large banks. Obviously, this difference could be the subject of a follow-up study.

Finally, the lagged dependent variable captures the delayed adjustment of NIMs to changes in the economic and financial variables. Due to the mix between interest-sensitive and non-sensitive assets and liabilities, adjustments to changes in the slope of the yield curve are not instantaneous, and the adjustment of loan volumes or loan pricing to growth should not be, either. Taking account of the 0.619 coefficients in Models 1 and 3, this version suggests that 85 percent of the adjustment is completed within four quarters.

Next, Model 2 contains the same economic and financial variable as in Model 1 along with fixed effects. The fixed effects added considerable explanatory power to Model 2. However, before preceding any further, the choice of fixed over random effects needs to be discussed. Here, the Hausman (1978) test was conducted to determine whether fixed or random effects were appropriate. The null hypothesis (no statistical difference between the fixed and random effects parameter estimates) was rejected with a p-value of <0.001. From this, it appears that the estimates from the random effects model are not efficient compared with the fixed effects model, and the fixed effects model is preferred. Additionally, a test for the redundancy of the fixed effects terms was also conducted. For both the F-test and the Chi square test for redundancy, the null was rejected with a p-value of <0.001, implying that the fixed effects contain information not captured elsewhere in the model. In other words, there remains unidentified information unique to each institution that is time invariant.

Beyond the fixed effects, all coefficients are significant at least at the $p < .001$ level. All coefficients have the same signs as in Model 1, with only minor differences in magnitude in most cases. NIMs are marginally less responsive to the slope of the yield curve and more responsive to credit quality in Model 2 than in Model 1. The negative effect of concentration is a bit stronger in Model 2 as is the negative effect of growth. Again, the similar qualitative movement in both these variables between Models 1 and 2 reinforces the notion that the concentration ratio reflects the same basic information as is reflected in the asset size variable. Effects of GDP growth are similar in both, while the effects of employment growth are stronger in Model 2. The most important difference is probably the speed of adjustment coefficient. It is roughly half the size in Model 2, and now 98 percent of the adjustment is completed within one year (four quarters).

Model 3 incorporates random effects. The earlier tests show that fixed effects are superior to random effects. This finding is borne out in a comparison of the coefficients in Models 3 and 1. Results for Model 3 are virtually indistinguishable from Model 1.

DePrince and Morris: Interest Sensitivity

Table 2: Estimation Results for Net Interest Margin (NIM)

	Model 1	Model 2	Model 3
Variable	Coefficient (P Value)	Coefficient (P Value)	Coefficient (P Value)
Constant	0.477 (<0.001)	0.871 (<0.001)	0.477 (<0.001)
10-Year Less 1-Year Treasury Yield	0.003 (<0.001)	0.002 (<0.001)	0.003 (<0.001)
AAA Less BAA Corporate Bond Yield	-0.020 (<0.001)	-0.024 (<0.001)	-0.02 (<0.001)
HHI*1000	-0.092 (<0.001)	-0.105 (<0.001)	-0.092 (<0.001)
LOG (Average Assets)	-0.006 (<0.001)	-0.017 (<0.001)	-0.006 (<0.001)
GDP Growth	0.002 (<0.001)	0.003 (<0.001)	0.002 (<0.001)
Employment Growth	0.003 (<0.001)	0.004 (<0.001)	0.003 (<0.001)
Lagged Dependent variable	0.619 (<0.001)	0.364 (<0.001)	0.619 (<0.001)
Effects	NA	Fixed	Random
R-squared	0.391	0.486	0.391
Adjusted R-squared	0.391	0.476	NA
Method: Pooled Least Squares			
Sample (Adjusted)	84:Q2-09:Q1	84:Q2-09:Q1	84:Q2-09:Q1
Periods Included after adjustments	100	100	100
Cross-Sections Included:	19641	19641	19641
Total Panel (Unbalanced) Observations	1056524	1056524	1056524

Summary

This study takes a different approach from earlier studies. As with Aliaga-Diaz and Olivero, it examines the effects of interest rates and general business conditions on net interest margins for all banks, but the data construction methodology differs considerably. Aliaga-Diaz and Olivero aggregate various income and expense items into industry-wide measures for each period. In Hanwick and Ryu, banks were aggregated into various classes based on key business lines. These business classes were the cross-sectional variables. In DePrince and Morris (2007), banks were aggregated into asset classes.

The important feature of this study is the fact that no aggregation of any sort was needed to produce a strong relationship between bank-level NIMs and a variety of financial and economic variables. While developing a data set of more than one million observations on nearly 20,000 banks over a 25-year sample period is itself a complicated task, it enabled the development of a study that used each bank as a panel and successfully accounted for differences in NIM by bank. NIM rises as the yield curve steepens and vice versa, increasing levels of risk premiums point to lower NIM, economic activity positively affects the NIM while positive effects of pricing power associated with concentration are evident, and the usual

inverse relationship between size and NIM is confirmed. The collective positive effect of the real variables is as expected, and the speed of adjustment coefficient confirms that effects are not immediate.

Finally, fixed effects are found to be an important determinant of bank-level NIMs, confirming the view that all banks do not respond in an identical manner to changes in economic conditions. At the same time, however, Model 1 stands on its own. Even without taking into account fixed effects, coefficients are significant with plausible signs. As important, fixed effects (Model 2) had minor effects on the coefficients in Model 1 except for the speed of adjustment coefficient. Thus, it appears that their effect on the speed of adjustment is the major contribution of the fixed-effects estimation.

In sum, the main purpose of the study was accomplished. A panel model of all banks can successfully explain bank-level NIMs through a variety of financial and economic variables and industry structures.

Acknowledgement

Thanks to Robert Kao (discussant) of Park University and the session attendees at the 37th Annual Meeting of the AEF in Houston, Texas, for their valuable comments. Their suggestions have been incorporated into this version of the study.

References

- Allen, Linda. 1988. The Determinants of Bank Interest Margins: A Note. *Journal of Financial and Quantitative Analysis* 23, no. 2:231-235.
- Aliaga-Diaz, Roger, and Maria Pia Olivero. 2005. "The Cyclical Behavior of Net Interest Margins: Evidence from the United States by Banking Sectors." <http://www4.ncsu.edu/~raaliaga/papers/Empirical%20paper%2058p%20x%2090.pdf> .
- Angbazo, Lazarus. 1997. Commercial Bank Net Interest Margins, Default Risk, Interest-Rate Risk, and Off-Balance Sheet Banking. *Journal of Banking and Finance* 21, no. 1:55-87.
- Basel Committee on Banking Supervision. 2003. *Principles for Management and Supervision of Interest Rate Risk*. Bank for International Settlements. Consultative Document. September.
- Carlson, Mark, and Roberto Perli. 2004. Profits and Balance Sheet Developments at U.S. Commercial Banks in 2003. *Federal Reserve Bulletin* 90, no. 2:162-191.
- Critchfield, Tim, Tyler Davis, Lee Favison, Heather Gratton, George Hanc, and Katherine Samolyk. 2004. Community Banks: Their Recent Past, Current Performance, and Future Prospects. *FDIC Banking Review* 16, no. 3:1-56.
- DePrince, Albert E., Jr., and Pamela D. Morris. 2007. A Longitudinal Model of Net Interest Margin by Bank Asset Size: 1992-2005. *Journal of Economics and Finance* 31, no. 1:20-32.
- DePrince, Albert E., Jr., and Pamela D. Morris. 2010. Causes of Interstate Differences in Community Bank Performance. *Journal of Economics and Finance*, forthcoming.
- Dueker, Michael J., and Daniel Thornton. 1997. Do Bank Loan Rates Exhibit a Countercyclical Mark-Up? Federal Reserve Bank of St. Louis. Working Paper 1997-004A.
- English, William B. 2002. Interest Rate Risk and Bank Net Interest Margins. *BIS Quarterly Review* 67-82. Federal Deposit Insurance Corporation (FDIC). 2003. *Quarterly Banking Profile*. Fourth Quarter.
- Federal Deposit Insurance Corporation (FDIC). 2009. *Quarterly Banking Profile*. Third Quarter <http://www2.fdic.gov/qbp/2009sep/qbp.pdf> .
- Gunther, Jeffrey W., and Robert R. Moore. 2004. Small Banks' Competitors Loom Large. *Southwest Economy*. Federal Reserve Bank of Dallas. January/February:1, 9-13.
- Hausman, J. A. 1978. Specification Tests in Econometrics. *Econometrica* 46:1251-1272.
- Hanweck, Gerald, and Lisa Ryu. 2005. The Sensitivity of Bank Net Interest Margins and Profitability to Credit, Interest Rate, and Term Structure Shocks across Bank Product Specializations. Federal Deposit Insurance Corporation. Working Paper 2005-02. January.
- Ho, Thomas S., and Anthony Saunders. 1981. The Determinants of Bank Interest Margins: Theory and Empirical Evidence. *Journal of Financial and Quantitative Analysis* 16, no. 4:581-600.
- Kahn, George A., Linda Schroeder, and Stuart Weiner. 2003. The Role of Community Banks in the U.S. Economy. *Economic Review*. Federal Reserve Bank of Kansas City 88, no. 2:5-43.
- Saunders, Anthony, and Liliana Schumacher. 2000. The Determinants of Bank Interest-Rate Margins: An International Study. *Journal of International Money and Finance* 19, no. 6:813-832.
- Zarruk, Emilio. 1989. Bank Spread with Uncertain Deposit Level and Risk Aversion. *Journal of Banking and Finance* 13:797-810.
- Zarruk, Emilio, and Jeff Madura. 1992. Optimal Bank Interest Margin under Capital Regulation and Deposit Insurance. *Journal of Financial and Quantitative Analysis* 27, no. 1:143-149.

Testing the Modigliani-Miller Risk-Class Property

David P. Echevarria, University of North Carolina Wilmington

Abstract

The Modigliani-Miller (1958, 1963) firm valuation irrelevancy propositions rest on the notion that firms are classified into risk classes. Assignment to a risk class drives the value-establishing discount rate decision. Most academic discussions of the irrelevancy propositions are focused on the impact of debt and taxes. Indeed, those observations motivated the 1963 revisions to the irrelevancy argument. Largely unexamined is the nature of the risk-class property. There are two potential factors that act on the risk-class property. The first is the volatility of market valuations. The second is the volatility of earnings. A number of studies suggest that prices, and by implication market valuations, are more volatile than earnings (and/or dividends). The efficient markets notion suggests that these two factors should be highly correlated. Statistical testing of an index composed of these two factors suggests that risk classes are highly volatile. Firms with low index values (market values more volatile than earnings) are very likely to score higher index values and vice-versa over time. Data analysis also suggests that while many market valuations are consistently more volatile than earnings, a non-trivial proportion of securities exhibit market value volatilities significantly less volatile than earnings. These results suggest that the risk-class property is more attractive as a theoretical construct than as a practical valuation concept.

Introduction

The Modigliani-Miller (1958) Capital Structure Irrelevance theorems form the basis for determining the value of the firm. The building blocks of the MM theorem are two propositions. Proposition I is well known: "...the market value of any firm is independent of its capital structure and is given by capitalizing its expected return at the rate ρ_k appropriate to its risk class." ¹ The capitalization rate for each risk class k is defined as the average cost of capital and is "...equal to the capitalization rate of a pure equity stream in its class". ² In effect, the earnings stream divided by the market value of its securities assigns the firm to a particular [yield-driven] risk class.

The Modigliani and Miller (MM) cost of capital and firm valuation propositions continue to be the object of discussion and conjecture. Most of the discussion centers on the impact of debt (capital structure) and taxes on the cost of equity and the cost of capital in general. The cost of capital determines the value of the firm. A measure of the discussion activity can be made using a simple query and the JSTOR collections (*Economics*). Entering "Modigliani and Miller" produces more than 140 references exclusive of the original 1958 paper and the 1963 "correction". The last major spate of papers is those written to celebrate the 30th anniversary of the original MM paper. These papers appeared in the *Journal of Economic Perspectives* (Volume 2, Number 4, Fall 1988). Four papers are briefly reviewed.

Literature Review

Modigliani (1988) focuses on the impact of taxes and tax rates on the value added by debt (and the tax-deductible interest thereon) and the impact of tax policy on dividends. Modigliani takes a much broader view of the MM propositions and the manner in which they have been incorporated in finance and economics.

Miller (1988) begins by observing that many of the controversies have been settled, especially "*Our Proposition I, holding the value of a firm to be independent of its capital structure (that is, its debt/equity ratio) is accepted as an implication of equilibrium in perfect capital markets.*" However, Miller notes that empirical attempts to settle issues regarding the value-invariance propositions have not been achieved by them or anybody else. The *invariance* issue is complicated by a number of factors. Chief among these are dividends, dividend policy, investment policy, issues regarding limited liability, risky debt and tax policy (rates).

Ross (1988) reviews the MM propositions in three critical areas: (1) The *Arbitrage* argument, (2) The Risk-Class concept, and (3) Taxation. Ross reflects on Miller's suggestion that the risk-class notion so necessary to the MM arbitrage argument has passed from favor. However, Ross believes that further research into the risk-class concept may yield implications for empirical work on MM theory. Ross's suggestion is the motivation for this study.

Bhattacharya (1988) notes work by DeBondt and Thaler (1987). The latter observe that stock returns display significant negative serial correlation over relatively long periods of time. They note that over time, stars become dogs and vice versa. Moreover, these reversals are accompanied by significant changes in their rank orders. DeBondt and Thaler conclude that there is over-reaction to current earnings. Bhattacharya suggests that DeBondt and Thaler do not recognize the mean reversion tendency of prices over time.

The volatility of market prices in relation to dividends has been extensively analyzed. Shiller (1981) explores the relationship between earnings, dividends, and [excessive] price volatility. Shiller observes that the volatility of prices cannot be explained by changes in dividends. Kleidon (1986) takes issue with the excessive volatility argument. Kleidon suggests that major changes are necessary in the behavioral assumptions underlying economic models. These changes are driven by apparent anomalies in financial economics. Kleidon observes that “arguments for such changes based on claims of ‘excess volatility’ in stock prices appear flawed for two main reasons: there are serious questions whether the phenomenon exists in the first place and, even if it did exist, whether radical change in behavioral assumptions is the best avenue for current research.” (p. S469)

One important expression of these beliefs is the notion that stock prices are set rationally as the present value of expected future cash flows. Kleidon notes that some researchers have challenged this proposition: “Studies by Grossman and Shiller (1981), LeRoy and Porter (1981), Shiller (1981), and others, use what has become known as *variance bounds* or volatility tests to examine the validity of the valuation model.” (p. S474) The primary claim of this literature is that the model is grossly violated in empirical tests - so grossly violated, in fact, that “for the aggregate stock market, there is no evidence at all that stock price movements have been followed by corresponding dividend movements” (Shiller 1984a, p. 476). Moreover, Kleidon notes that the literature claims that stock prices show excess volatility and that “the variability of stock prices cannot be accounted for by information regarding future dividends” (quoted from Grossman and Shiller 1981, p. 222). (p. S474).

The conclusions drawn from the articles above suggest that valuation models (and theories) are victimized by a significant number of economic, financial, and statistical factors. Moreover, a significant background factor is the notion of how economic and financial variables should interact in an efficient market. The simplest form of efficient market theory suggests that prices reflect all known information. Empirical studies indicate that there is a certain degree of chaos or over-reaction to information events. However, the central tendency of economic and financial series over time should permit a reasonable estimate for the mean of expectation and its variance. The proverbial “fly-in-the-ointment” is whether or not we have a normally distributed variable of interest. When the [test] sample is large, we tend to assume normality. However, we know that stock returns are not normally distributed and that “fat tails” present a significant problem with significance testing.

Measuring the Association between Business Risk and Market Risk

In an efficient market, the generally accepted notion is that earnings expectations drive prices. Moreover, earnings and value are generally positively correlated. Earnings are subject to a variety of macro-economic forces. These forces may be collectively described as *business risk*. A scale-free proxy measure for business risk is computed simply as the coefficient of variation (CV) for operating income (earnings before interest, taxes, amortization and depreciation expenses or EBITDA). Stock prices (and the market value of the firm’s equity or MVE) are also subject to a complex variety of forces. These forces are collectively termed as *market risk*. In the same manner as business risk, the scale-free measure of market risk is the CV for the market value of equity.

The research hypothesis examined herein is that in an efficient market, business risk and market risk ought to move in tandem. Increases in business risk ought to be reflected with increases in market risk. The Modigliani & Miller (MM) risk class property is [of necessity] the result of the interaction of business and market risk. It is reasonable to suggest that an index can be formed from the CV for EBITDA and the CV for MVE as shown in equation (1):

$$MM\ Index = CV_{EBITDA} / CV_{MVE}. \quad (1)$$

If markets are efficient, then the MM Index (MMI) ought to hover around unity. The expected value of MMI should be equal to 1. However, studies by Shiller (1981) and others suggest that prices are much more volatile than changes in dividends. Dividend payouts are driven by earnings which are subject to business risk. Prices (and the market value of equity) are subject to market risk. Accordingly, we ought to expect that the empirical MMI will be less than 1; prices are more volatile than earnings.

The efficient markets expectation that the MM Index should be equal to 1 may be demonstrated using the Gordon [Normal] Growth model. The key variables for this simulation are the expected rate of return for equity investors (10.4%). The company pays out 20% of its earnings per share. The growth rate for dividends is set at 4% per annum. This requires that net sales (NS) increase by 4% per annum and the profit margins remain fixed. The hypothetical firm has 100 million shares issued and outstanding and no further issues are anticipated. The firm enjoys a 40% operating profit margin and a net profit margin of 10%. Year 1 starting values for Net Sales (NS) are \$10 per share. Table 1 contains the result of this 20-year simulation. The 20-year MM Index (= $CV_{EBITDA} \div CV_{MVE}$) is exactly equal to 1.0000. Increases in the payout ratio results in

increased stock price and MVE as expected. Independent or combined changes in any of the start parameters always results in a MMI value of 1.000.

Table 1: Gordon [Normal] Growth Model Simulation Values

All figures per share except market value of equity						
	k =	0.104	Shares		NPM	OPM
	g =	0.04	100M		0.1	0.4
Year	Div	Po	MVE	EPS	EBITDA	NS
1	0.2	3.125	312.5	1	4	10
2	0.21	3.25	325	1.04	4.16	10.4
3	0.22	3.38	338	1.08	4.326	10.82
4	0.22	3.515	351.5	1.12	4.499	11.25
5	0.23	3.656	365.6	1.17	4.679	11.7
6	0.24	3.802	380.2	1.22	4.867	12.17
7	0.25	3.954	395.4	1.27	5.061	12.65
8	0.26	4.112	411.2	1.32	5.264	13.16
9	0.27	4.277	427.7	1.37	5.474	13.69
10	0.28	4.448	444.8	1.42	5.693	14.23
11	0.3	4.626	462.6	1.48	5.921	14.8
12	0.31	4.811	481.1	1.54	6.158	15.39
13	0.32	5.003	500.3	1.6	6.404	16.01
14	0.33	5.203	520.3	1.67	6.66	16.65
15	0.35	5.411	541.1	1.73	6.927	17.32
16	0.36	5.628	562.8	1.8	7.204	18.01
17	0.37	5.853	585.3	1.87	7.492	18.73
18	0.39	6.087	608.7	1.95	7.792	19.48
19	0.41	6.331	633.1	2.03	8.103	20.26
20	0.42	6.584	658.4	2.11	8.427	21.07
			107.41	Std. Dev.	1.37	
			465.28	Mean	5.96	
		MVE	0.2308	CV	0.2308	EBITDA
			MM Index =		1.0000	

The simulations shown in Table 1 assume that the 4% expected growth rate in dividends must be applied to the growth rate of sales. It also assumes that the Operating and Net profit margins that determine operating income (EBITDA) and earnings per share (EPS) remain fixed as sales increase. This assumption requires that the firm's operating characteristics, in particular its production function remain constant. We observe that firms typically add capacity at the margins. It is only when we observe a marked increase or decrease in capacity or capacity utilization that the production function is likely to see a significant change.

The results in Table 1 could only obtain in a deterministic world. Obviously, the market is far from deterministic. It is conceivable that market valuations may break the expected bonds with earnings and behave in random or uncorrelated ways.

Data and Methodology

This study incorporates 4 separate samples for companies in the Standard & Poor 500 Index. Two samples are comprised of yearly data and two samples of quarterly data. The first yearly data sample covers the period from 1977 to 1996 and the second from 1999 to 2008. The yearly data samples are further screened to align fiscal year end (FYE) dates. The 1977-96 dataset is comprised of 282 companies with an FYE of December. The 1999-2008 dataset is comprised of 233 companies.

The quarterly data samples are obtained from the same two Compustat data files. Quarterly data is also aligned by fiscal quarters. This alignment is made necessary to reduce the impact of market trends from adjacent quarters. Companies missing 6 or more consecutive quarters of data are excluded from the final data samples. The first quarterly dataset includes companies with fiscal quarters in March 1987 (n=3), June 1987 (n=45) and September 1987 (n=277) for a total of 325

Echevarria: Testing the MM Risk-Class Property

companies covering 39 quarters. The second quarterly dataset includes companies with fiscal quarters in December 1999 (n=207) and March 2000 (n=177) for a total of 384 companies covering 39 quarters.

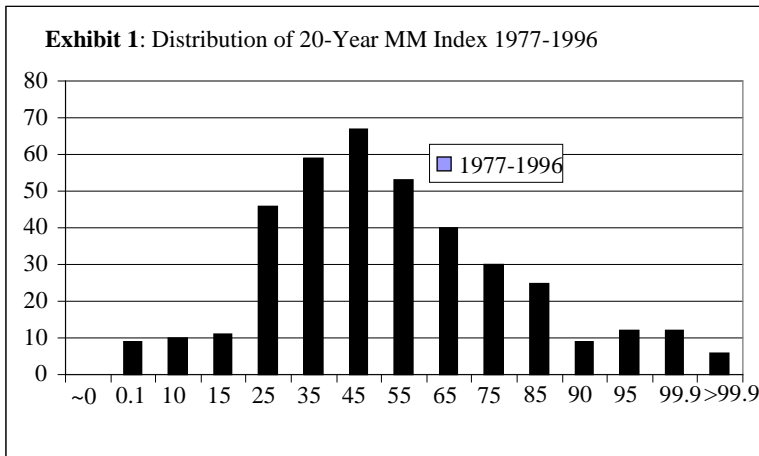
The research hypothesis is first examined by comparing the average MMI in the first 10 year sub-period with the average MMI in the second 10-year sub period for the 1977-1996 dataset. The *a priori* expectations are influenced by the numerous reported observations of a central tendency in prices. Thus we should expect that high index scores in the first period ought to be followed by lower index scores in the second period or vice-versa.

Table 2 contains the Summary Statistics for the first two 10-year sub-periods and the full 20-year period. The summary statistics suggest that market values are more volatile than earnings: MM index values are less than 1.0000. Moreover, the theoretically expected mean MMI value [$\mu_{MMI} = 1$] lies outside the 95% confidence intervals for the standard error of the mean (SEM). The slight difference in mean MMI values is not statistically significant (z-score = -0.7026). Not fully reflected in the summary statistics is the range of MMI values. A simple screen was utilized to limit the effects of extreme outliers. The latter were defined as values greater than +/- 18 standard deviations from the sample mean. Typically, less than 5 MMI observations were excluded from the summary statistics for each period.

Table 2: Summary Statistics MM Index

10- and 20 Year Periods: 1977-1996			
	77-86	87-96	77-96
MEAN	0.8834	0.9192	0.8596
MEDIAN	0.7258	0.7963	0.8156
STDEV	0.6340	0.5766	0.3041
N	282	282	282
SEM	0.0378	0.0343	0.0181
Upper	0.9574	0.9865	0.8951
Lower	0.8094	0.8519	0.8241
Z-score	-0.7026		

In each of the Table 2 period samples, the medians are below the means, suggesting a skewed distribution. Exhibit 1 displays the distribution of MMI values for the 20-year period. These [skewed] distributions are typical of all yearly and quarterly samples.



Summary statistics for the second yearly sample (1990-2008) are contained in Table 3. The 10- and 9-year mean MMI values are statistically different ($Z = -2.7421$). The 00-08 period upper/lower bounds includes the hypothetical mean MMI value ($\mu = 1$) as does the combined 20 year sample. The mean MMI value for the 00-08 period suggests that market values were just slightly less volatile than earnings. Moreover, a non-trivial number of companies had MMI values statistically equal to 1 in contrast to mean MMI values reported in Table 2 and for the period 90-99 in Table 3. Taking a broad perspective, the average MM Index values continued to increase, although not at a steady rate, from the late 1970s until 2008. Further detailed analysis will be necessary to form an explanation for this increase.

Table 3: Summary Statistics MM Index

10, 9 and 20 Year Periods: 1990-2008			
	90-99	00-08	90-08
MEAN	0.8722	1.0405	0.9634
MEDIAN	0.7455	0.8619	0.9236
STDEV	0.6053	0.7151	0.3890
N	233	233	233
SEM	0.0397	0.0468	0.0255
Upper	0.9500	1.1324	1.0134
Lower	0.7945	0.9487	0.9135
Z-score	-2.7421		

Summary Statistics for Rolling Five Year Average MM Indexes

In order to better understand the dynamic behavior of the MM Index over the 20-year time spans, rolling 5-year average MMI values are examined. The 2 datasets overlap for three years and appear to have the same values in 1994, 1995 and 1996. However, 121 companies are contained in both SP500 datasets with FYE in December. Summary statistics for each of the datasets are contained in Tables 4 and 5. Z-score test statistics reflect testing differences between the current 5-year period and the previous 5 year period. The z-score -0.8561 indicates the 82-86 Mean MMI (1.3071) is statistically equal to the 77-81 Mean MMI (1.2019). The results in Table 4 also suggest that the hypothetical mean MM index ($\mu = 1$) is contained within the standard error of period mean MMI for the last two sub-periods.

Table 4: Summary Statistics MM Index

5-Year MM Index values for Period 1977-1996				
	77-81	82-86	87-91	92-96
MEAN	1.2019	1.3071	1.0342	1.1224
MEDIAN	0.9498	0.8182	0.7451	0.7961
STDEV	1.0458	1.7801	1.0372	1.0626
N	282	282	282	281
SEM	0.0623	0.1060	0.0618	0.0634
Upper	1.3239	1.5149	1.1553	1.2466
Lower	1.0798	1.0994	0.9132	0.9981
Z-score		-0.8561	2.2243	-0.9961

Table 5: Summary Statistics MM Index

5-Year MM Index values for Period 1990-2008				
	90-94	95-99	00-04	04-08
MEAN	1.1332	0.9277	1.0619	1.0527
MEDIAN	0.7752	0.6804	0.7702	0.8489
STDEV	1.4207	0.9817	1.0012	1.1656
N	230	233	232	233
SEM	0.0937	0.0643	0.0657	0.0764
Upper	1.3168	1.0537	1.1907	1.2024
Lower	0.9495	0.8016	0.9331	0.9030
Z-score		1.8085	-1.4599	-0.5287

The two datasets have seven years of overlapping data (1990 to 1996). The overlap allows two important observations. The first is the observation that the hypothetical mean ($\mu = 1$) is contained within the standard error of each sub-period mean MMI after 1986. The second observation is medians in each of the eight sub-periods are always less than 1 and hover around 0.80. This suggests that at least 50% of all companies experience market values to be more volatile than earnings. In the next section, the behavior of the MM index will be examined using quarterly data.

Summary Statistics for MM Index using Quarterly Data

In the previous section annual data is examined to determine how the MM Index behaves over several years. A more time-sensitive dataset is examined to determine the sensitivity of the MMI to quarterly earnings reports. The *a priori* expectation is that quarterly data may be more volatile. This expectation is based on casual observations of market responses to quarterly earnings reports. It is frequently observed that stock prices will react rather negatively to earnings reports that match or fall short of expectations. It is not unusual to observe that even when reported earnings exceed analysts' expectations, the stock will still react negatively and vice-versa. The market values utilized in this section are end-of-quarter values and do not reflect the kinds of price (and market value) changes when earnings are reported in the financial media.

In each of the two datasets, the data are aligned by quarters. Thus all companies reporting in March, June, September, or December comprise the test sample data set. The 1977-1996 dataset yields 325 companies with data beginning in March 1987 (n=3), June 1987 (n=45) and September, 1986 (n=277). The 1999-2008 dataset yields 384 companies with data beginning in December 1999 (n=205) and March 2000 (n=177). Both data samples cover 39 quarters.

Two periodicities for mean and median MMI are examined: 10 quarters and 6 quarters. The exception to the 10-quarters is the first sub-period which contains 9-quarters in both samples. Sample summary statistics for 1987-1997 are contained in Table 6 and Table 7 for the 2000-2009.

Table 6: Summary Statistics MM Index

10-Quarter MM Index values for Period 9/87 - 3/97				
	9/87-9/89	12/89-3/92	6/92-9/94	12/94-3/97
MEAN	0.8561	0.5486	0.5625	0.5114
MEDIAN	0.6063	0.4054	0.4276	0.4463
STDEV	1.1915	0.7947	0.9489	0.5093
N	323	321	318	323
SEM	0.0663	0.0444	0.0532	0.0283
Upper	0.9861	0.6355	0.6668	0.5670
Lower	0.7262	0.4617	0.4582	0.4559
Z-score		3.6844	0.0400	0.7348

The summary statistics for the period between September 1987 and March 1997 in Table 6 support the *a priori* expectation that quarterly market value data vis-à-vis earnings data would prove more volatile than yearly data. The 9-quarter period ending September 1989 is statistically greater than the following quarter ($z = 3.6844$). The next three quarters have mean MMI values slightly greater than 0.5. In each of the four sub-periods, the upper and lower bounds exclude the expected mean value. The mean and median MMI values reveal substantially greater volatility for market values than the yearly data covering the same time period.

The summary statistics for the period March 2000 to September 2009 contained in table 7 tell a significantly different story.

Table 7: Summary Statistics MM Index

10-Quarter MM Index values for Period 3/00 to 9/09				
	3/00-3/02	6/02-9/04	12/04-3/07	6/07-9/09
MEAN	1.8157	1.6955	1.9299	1.3205
MEDIAN	1.1141	1.0794	1.2674	0.9563
STDEV	2.3381	1.8587	2.2946	2.0603
N	379	378	380	380
SEM	0.1201	0.0956	0.1177	0.1057
Upper	2.0511	1.8829	2.1607	1.5276
Lower	1.5804	1.5081	1.6992	1.1133
Z-score		0.7832	-1.5459	3.8526

Mean and median MMI values are significantly greater than 1; see upper and lower bounds. Compared to the first dataset, mean and median MMI values are 2 to 3 times greater. This suggests that for the majority of companies in this second time period, market valuations were significantly less volatile than earnings. Why this period proved to have market values much less volatile than earnings will require a detailed study of earnings expectations, earnings announcement dates and changes in prices. End-of-quarter financial reports always appear after the end of the quarter and it is possible that some

price reaction occurred before these reports were widely available. Sensitivities of the MM Index are examined utilizing 6-quarter averages.

The summary statistics in Table 8 clearly show the tendency for mean and median MM Index values to run substantially below 1.00 for the period between September 1987 and March 1997. Table 9 displays a substantially different mean and median MMI trend for the period from June 2001 and September 2009. In the latter time frame, mean and median MMI values are significantly greater than 1.00.

Six-Quarter mean and median MM Index values for the period beginning September 1987 and ending June 1996 follow a pattern of values similar to the 10-quarter summary statistics in Table 6. Mean and Median MMI values are statistically less than 1.00 and support the a priori expectations that quarterly market value data would prove to be more volatile than the annual data statistics in Table 4. The 6-quarter period ending June 1991 was significantly more volatile than the prior 6-quarter period ending June 1990 (Z-score = 3.232).

Table 8: Summary Statistics MM Index

6-Quarter MM Index values for Period 9/1987 - June 1996						
	Dec-88	Jun-90	Dec-91	Jun-93	Dec-94	Jun-96
MEAN	0.8250	0.6833	0.4691	0.5666	0.5856	0.5401
MEDIAN	0.5830	0.5117	0.3838	0.4369	0.4309	0.4438
STDEV	1.0091	0.8776	0.8336	0.7592	0.7633	0.6147
N	322	318	318	307	320	322
SEM	0.0562	0.0492	0.0467	0.0433	0.0427	0.0343
Upper	0.9352	0.7798	0.5607	0.6515	0.6692	0.6072
Lower	0.7147	0.5869	0.3774	0.4817	0.5020	0.4729
Z-score		1.7953	3.232	-1.541	-0.2706	0.8114

The 6-quarter results for the period beginning September 1999 and ending September 2008 are contained in Table 9. Similar to the 10-quarter results reported in Table 7, mean and median MM index values are statistically greater than one in every instance save the first 6-quarter period ending June 2001. Z-score are statistically significant for the rise in mean MMI for the period ending December 2002 and for the decline in the period ending December 2008. Similar to the 10-quarter results for this period, further detailed analysis will be require to determine causal factors for the decline in market value volatilities vis-à-vis earnings.

Table 9: Summary Statistics MM Index

6-Quarter MM Index values for Period 12/1999 - 9/2008						
	Jun-01	Dec-02	Jun-04	Dec-05	Jun-07	Dec-08
MEAN	1.4611	1.9241	1.7815	2.0904	1.9983	1.3874
MEDIAN	0.9461	1.1886	1.0866	1.3495	1.2966	0.9824
STDEV	2.1719	2.5109	1.9746	2.4655	2.0968	1.9890
N	379	376	381	379	380	375
SEM	0.1116	0.1295	0.1012	0.1266	0.1076	0.1027
Upper	1.6797	2.17792	1.9798	2.3387	2.2091	1.5887
Lower	1.2424	1.6703	1.5832	1.8422	1.7875	1.1860
Z-score		-2.7092	0.8680	-1.9061	0.5546	4.1079

Conclusions

This study began with the objective of determining the feasibility or usefulness of the Modigliani & Miller risk-class concept without the impediment of requiring positively correlated (substitutable) earnings streams. It is quite obvious that the risk class property displayed in the empirically derived MM indexes requires some modification. Rather than point estimates, ranges of values may be necessary to fix the risk-class discount rate. Unhappily, the range values may only be useful in one time period.

Two important conclusions may be drawn from this study. First, the results suggest that market valuations are, on average, 20 to 50% more volatile than earnings and that the magnitudes are sample calendar time and periodicity sensitive. Second, MM index scores display a considerable amount of temporal instability. However, temporal instability does not suggest that central tendency is active. A non-trivial number of companies exhibit valuation volatilities consistently greater than earnings volatilities. In like manner, a smaller number of companies exhibit market value volatility less than earnings volatility. Further study will be required to ascertain the causal factors.

Notes

1. Modigliani, Franco and Merton H. Miller. The Cost of Capital, Corporation Finance and the Theory of Investment. *American Economic Review*, Vol. 48, No. 3, page 268.
2. Ibid, at 269.

References

- Bhattacharya, Sudipto. 1988. "Corporate Finance and the Legacy of Miller and Modigliani." *The Journal of Economic Perspectives*, Vol. 2, No. 4: 135-147.
- Kleidon, Allan W. 1986. "Anomalies in Financial Economics: Blueprint for Change?" *Journal of Business*, Vol. 59, No. 4: S469-S499
- Miller, Merton H. 1988. "The Modigliani-Miller Propositions After Thirty Years." *Journal of Economic Perspectives*, Vol. 2, No. 4: 99-120.
- Modigliani, Franco and Merton H. Miller. 1958. "The Cost of Capital, Corporation Finance and the Theory of Investment." *American Economic Review*, Vol. 48, No. 3.
- Modigliani, Franco. 1988. "MM--Past, Present, Future." *Journal of Economic Perspectives*, Vol. 2, No. 4: 149-158.
- Morck, R., A. Shleifer, R. W. Vishny. 1990. "The Stock Market and Investment: Is the Market a Sideshow?" *Brookings Papers on Economic Activity*, Vol. 1990, No. 2: 157-215.
- Ross, Stephen A. 1988. "Comment on the Modigliani-Miller Propositions." *Journal of Economic Perspectives*, Vol. 2, No. 4: 127-133.
- Shiller, Robert J. 2003. "From Efficient Markets Theory to Behavioral Finance." *Journal of Economic Perspectives-Volume 17, Number 1*: 83-104.
- Shiller, Robert J. 1981. "The Use of Volatility Measures in Assessing Market Efficiency." *Journal of Finance*, Vol. 36, No. 2: 291-304.
- Shiller, Robert J. 1987. "The Volatility of Stock Market Prices." *Science*.235: .33-37.

A Nonparametric Kernel Method Test of the Symmetry of Security Returns

Gary S. Moore, Qin Shao, and Haifeng Zhang, University of Toledo

Abstract

Significant asymmetry in stock return data exists in several years according to both the skewness test and kernel function test. However, the degree of asymmetry in the CRSP value weighted index and to a lesser degree the S&P 500 data is substantially less than that suggested by the parametric tests. This allows some degree of confidence that the results of decades of earlier research which incorporated an assumption of symmetry remain valid at least as a first approximation.

Introduction

A significant amount of financial research has been conducted under the assumption of a symmetric distribution for stock returns. In particular, many of the concepts in the theoretical and empirical literature that have been developed in the past half century rest on the assumption of normal distributions. However, controversy concerning the distributional properties of security returns has occupied a predominant position in the finance and statistics literature for decades. Initially, researchers used normality as a working approximation of the distribution of security prices. The first attack on the assumption that returns or prices were properly modeled by a normal distribution was launched by Benoit Mandelbrot (1963). After rejecting normality for a distributional model for asset returns, Mandelbrot suggested that financial returns are more properly modeled by non-normal stable distributions. Fama (1965) extended Mandelbrot's research, which resulted in the popularity of a stable non-normal symmetric characterization. Non-Gaussian stable distributions are often called "stable Paretian" distributions.

A major problem with the stable Paretian distribution was that the distribution has an infinite variance. Given that variance was accepted as the appropriate measure of risk in a diversified portfolio by Markowitz (1952), many financial theorists shunned the framework in favor of the normal distributions. Markowitz theory posits that only expected return and risk are important in portfolio selections and that the risk of a portfolio of stocks that were not perfectly positively correlated would be less than the weighted average of the risk of the individual stocks comprising the portfolio. The portfolio predictions concerning diversification do not follow in a Paretian distribution theoretical model. However, the Markowitz framework seemed to have many practical implications. On the one hand, "normality" remained the cornerstone of the leading theories used in finance despite evidence from numerous studies like Fama (1965); on the other hand, asset pricing models incorporating asymmetry and, in particular, skewness have been around since the mid 1970s. Kraus and Litzenburger (1976) were among the first to incorporate skewness into a capital asset pricing model. The problem was that many researchers have often ignored the issue when conducting research.

Increasing evidence of asymmetry places distributional assumptions such as the normal or Paretian into question. Studies that have investigated the matter, such as Harvey and Siddique (2000), have not in general found convincing support for the symmetric distributions such as the normal or Paretian. They find that the unconditional skewness of returns is negative for the period they considered. In general, we think the tenor of the argument has developed into a question as to whether, despite the statistical empirical evidence against symmetry, the assumption of symmetric distributions such as the normal or Paretian continues to make sense as a tenable simplification of reality. This article examines additional evidence concerning the issue of whether symmetric distributions such as the Gaussian continue to be useful characterizations of security returns.

Since a substantial number of empirical applications of modern finance theory assume a symmetric distribution, it is important to know not only if there are statistically significant violations of asymmetry, but also the relative incidence of these violations. This is crucial to our understanding of the usefulness of the results of the empirical methods which remain in use. In particular, we must realize that practitioners continue to use methods which assume symmetry. Therefore, we must be interested in the actual incidence of asymmetry in the distribution of returns to assess the practical deficiencies of the state of the profession.

We attack this task in a number of ways. First, we apply the traditional skewness test, which has a Gaussian assumption incorporated into the test. Next, we apply a nonparametric test to measure the incidence of asymmetry in various indexes. We compare and contrast traditional measures of skewness incorporating the Gaussian assumption and our selected kernel method test of asymmetry. Our finding is that although there is statistically significant asymmetry in the data, the incidence is not as high as traditional skewness tests would suggest. This may partially explain the lack of a greater movement to more

sophisticated financial and statistical models among some practitioners. Moreover, this allows some degree of confidence that the results of earlier research studies, representing decades of research which incorporated an assumption of symmetry, remains valid at least as a first approximation.

Symmetry as a First Approximation

Symmetry of the return distribution has long been considered as a useful theoretical assumption for building financial models. Fama (1976) examined the distribution properties of the monthly returns of the Dow Jones 30 stocks during the January 1951 through June 1968 period. Although he observed some positive skewness in the returns, he concluded: “the right skewness of the frequency distribution of simple monthly returns is slight, and we can be comfortable with the assumption of symmetry as a working approximation.” (Fama 1976, p. 41).

It is interesting to note that Fama did not favor the normal distribution but was a proponent of the stable Paretian. For decades, financial researchers have used Fama’s assumption of symmetry as a working approximation. Most of the workhorse models of finance, including the CAPM and the Black-Scholes Option Pricing Model, implicitly incorporate symmetry as an underlying assumption. More recently, many studies have started to question the usefulness of this assumption as a working approximation. Many studies have suggested that some significant degree of unconditional skewness exists in some stock returns distributions. It is interesting to note that some of the original studies using monthly data, such as Fama (1976), tended to find positive skewness, while later studies such as Campbell and Siddique (2000) have often found negative skewness. Moreover, the sign is not constant from year to year. Also, the literature reports that individual stocks tend to have positive skewness (some of these studies use monthly data); while indexes tend to have negative skewness (most of these studies use daily data).

A number of economic rationales have been proposed as explanations for the documented asymmetry. The leading theories concerning unconditional skewness are the leverage effect hypothesis, the volatility feedback effect, and an option effect. The earliest explanation of unconditional asymmetry in stock returns is called the “leverage effect.” Black (1976), Christie (1982), French, Schwert, and Stambaugh (1987), and Nelson (1989) all document that stock volatility is negatively related to stock returns. Of particular interest is the fact that an unexpected increase in volatility is associated with negative returns. This negative relation is predicted by two hypotheses. First, since the firms in market portfolio possess financial leverage, a drop in the relative value of stocks versus bonds increases the volatility of the stocks (Christie 1982). Black (1976) suggests that operating leverage should have a similar effect. Second, if increases in predictable volatility increase the discount rates of future cash flows to stockholders, but not the expected cash flows, then unexpected increases in volatility will cause a drop in stock prices.

The volatility feedback effect (Campbell and Hentschel, 1992) is a leading explanation of the observed left-skewness of unconditional return distributions. Volatility feedback amplifies the impact of bad news but dampens the impact of good news, although an increase in volatility is associated with all kinds of news. Bollerslev, Litvinova, and Tauchen (2006) point out that the fundamental difference between the leverage effect and the volatility is the direction of causality. They say that with volatility being priced into the stock, any anticipated increase in volatility should raise the equilibrium return, which should cause an immediate stock price decline. The leverage effect suggests a low return leads to a higher volatility, whereas the volatility feedback effect explains why an upward movement in volatility can yield a negative return. Consequently, the causality in the volatility feedback effect runs from volatility to prices, as opposed to leverage effect, which suggests causality in the opposite direction (French et. al, 1987).

Studies comparing the relative magnitude of the two effects have suggested that the volatility feedback effect is stronger than the leverage effect (Bekaert and Wu, 2000) and (Wu, 2000). Nelson (1991), Engle and Ng (1993), and Glosten et. al. (1993) have discovered that variability in stock returns is greater following negative returns.

The existence of limited liability in firms is also proposed as a cause of asymmetric returns by Black (1976) and Christie (1982). Limited liability is thought to create an option like asymmetry in returns since returns are bounded on the down side. Formal tests of the limited liability hypothesis are absent from the literature.

Conditional variance and conditional skewness have gained prominence in the finance literature (Harvey and Siddique, 2000). The conditional variance and conditional skewness literature is a vast and growing body of literature and will not be reviewed here. Overall, the exact role of asymmetry and the connected issue of skewness (both conditional and unconditional) in portfolio theory remains one of disagreement and controversy. The area is certainly a fertile ground for future research.

Still, a very basic question remains. Can an assumption of unconditional symmetry be upheld as a working approximation, or is the assumption violated so badly that we should abandon it – and with it many of our most useful asset pricing models and option pricing models? The issue is one of pragmatism and can only be decided by examining the recent data and the possible cost and benefits to retaining the assumption of symmetry as a first approximation.

To address these issues we examine stock return data with an emphasis on the post 1987 crash period. The emergence of the issue of “crash-a-phobia” suggests that unconditional asymmetry takes on even more importance, particularly in the option pricing literature (Jackwerth and Rubinstein, 1996). Crash-a-phobia is the phenomenon that options, like puts, became unusually “expensive” after 1987 (Rubinstein, 1994).

Because significant asymmetry implies that the distribution is non-normal, it seems especially useful to analyze the issue with a non-parametric test of asymmetry. Consequently, we will want to compare and contrast the traditional parametric test of symmetry with a non-parametric test.

Data Description

The data are the daily returns to securities included on the Center for Research in Security Prices (CRSP) NYSE/ Amex/ NASDAQ file. Consequently, the analysis is focused on common stocks of domestic firms. The sample period is January 1980 through December 2005. We use three major indexes to measure daily stock returns: S&P 500, NASDAQ, and CRSP value-weighted index.

The S&P 500 represents the large capitalization stocks. Typically large capitalization stocks are over the 10 billion dollar range, and this group usually represents over 70 percent of the market capitalization of the domestic stock market. The NASDAQ represents a collection that includes many smaller stocks and consequently would be expected to have a higher expected return, higher volatility and possibly more skewness. The CRSP value weighted index is a broad based index that would include both of these groups.

Parametric Test of Skewness with Normal Distribution

Suppose Z is the return. Let $Y = \ln(1 + Z)$ and then center the data by $X = Y - \bar{Y}$. In the following analysis, we will focus on the data $\{X_i : i = 1, 2, \dots, n\}$, the log transformed percentage change of the returns.

Let τ be the skewness. The hypotheses are

$$H_0 : \tau = 0 \text{ vs } H_a : \tau \neq 0.$$

Let $\hat{\mu}_r = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^r$. Define the test statistic as follows:

$$\hat{\tau}_n = \frac{\hat{\mu}_3}{(\hat{\mu}_2)^{3/2}}. \tag{1}$$

If $\{X_i : i = 1, 2, \dots, n\}$ are independent and identically normally distributed, then $\sqrt{n}\hat{\tau}_n$ is asymptotically normally distributed, i.e.

$$\sqrt{n}\hat{\tau}_n \rightarrow N(0,6) \text{ as } n \rightarrow \infty. \tag{2}$$

Note the p -value of the above parametric test on skewness can be obtained by (2), and thus the validity of test relies on the assumption of normality.

We calculate the test statistics in (1) along with their p -values by (2) for all the 26 years, and summarize the results in Table 1. Statistically significant skewness is found in about half of the years. According to this parametric test, about half of the daily data are skewed. This is extremely high incidence and should give us pause as to whether the methods of modern finance implicitly assuming symmetry are practically justified.

However, there are a number of problems with this parametric approach in this context. First, since the p -value is obtained under the normality, the conclusion is only appropriate when the true distribution of $\{X_i : i = 1, 2, \dots, n\}$ is normal. However, this true distribution in practice is usually unknown. The conclusion could be misleading with a non-normal distribution. Second, another drawback of this method is that skewness equal to zero does not imply the symmetry of the distribution; i.e. the skewness of some asymmetric distributions is zero. Lack of skewness is a necessary but not sufficient condition for symmetry. Consequently, the absence of skewness doesn't prove things either way concerning asymmetry. Consequently, to overcome the awkward assumption on normality, in the following we apply a nonparametric test to the hypothesis of symmetry. Instead of specifying the distribution, we estimate and test the distribution solely by data.

Table 1: Parametric Skewness Test

	CRSP		NASDAQ		S&P 500	
	$\widehat{\tau}_n$	<i>p</i> -value	$\widehat{\tau}_n$	<i>p</i> -value	$\widehat{\tau}_n$	<i>p</i> -value
1980	-0.364	0.018*	-1.615	0.000***	-0.220	0.152
1981	-0.335	0.030*	-1.178	0.000***	-0.055	0.717
1982	0.506	0.001***	-0.059	0.697	0.573	0.000***
1983	-0.176	0.254	-0.757	0.000***	-0.021	0.887
1984	0.839	0.000***	0.661	0.000***	0.802	0.000***
1985	0.322	0.034*	-0.126	0.414	0.422	0.006**
1986	-1.132	0.000***	-1.696	0.000***	-0.966	0.000***
1987	-4.472	0.000***	-2.881	0.000***	-4.962	0.000***
1988	-1.070	0.000***	-0.916	0.000***	-1.024	0.000***
1989	-2.233	0.000***	-1.286	0.000***	-1.803	0.000***
1990	-0.322	0.036*	-0.559	0.000***	-0.163	0.287
1991	0.065	0.671	-0.409	0.007**	0.173	0.259
1992	-0.077	0.615	-0.115	0.452	0.055	0.719
1993	-0.705	0.000***	-1.01	0.000***	-0.177	0.247
1994	-0.424	0.005*	-0.364	0.018*	-0.289	0.060
1995	-0.411	0.007*	-0.653	0.000***	-0.074	0.630
1996	-0.701	0.000***	-0.622	0.000***	-0.608	0.000***
1997	-1.112	0.000***	-0.893	0.000***	-0.672	0.000***
1998	-0.734	0.000***	-0.703	0.000***	-0.616	0.000***
1999	-0.036	0.810	-0.313	0.041*	0.061	0.691
2000	-0.038	0.802	0.139	0.365	0.001	0.996
2001	0.008	0.956	0.410	0.008**	0.020	0.896
2002	0.416	0.007*	0.428	0.005**	0.422	0.006**
2003	0.023	0.880	0.016	0.914	0.053	0.730
2004	-0.223	0.147	-0.154	0.316	-0.109	0.478
2005	-0.075	0.624	-0.036	0.813	-0.015	0.919

Note: The asterisks, *, **, and *** denote statistic significance at the 5%, 1%, and .01%, respectively.

Nonparametric Test of Symmetry with Kernel Function

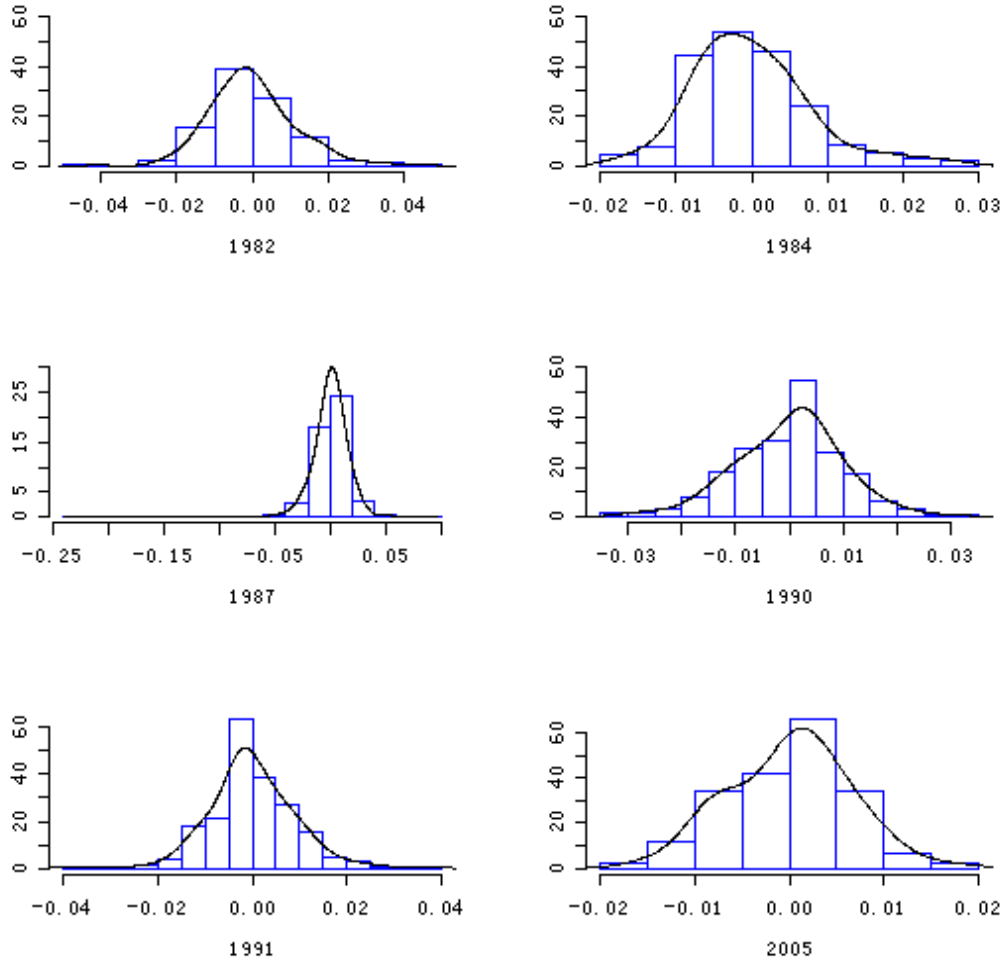
The Kernel method is one of the commonly used nonparametric methods to estimate the density function in statistics. With the kernel method, the density function $f(x)$ of X is estimated as follows:

$$\widehat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right), \tag{3}$$

where $K(x)$ is the kernel function and h is the bandwidth. Generally the kernel functions are assumed to be some symmetric probability density function with $\int_{-\infty}^{\infty} K(x)dx = 1$. Silverman (1985) provided the details on this method with several kernel functions. Especially, in the following discussion, we use the standard normal density function as the kernel function:

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \tag{4}$$

The bandwidth $h = \widehat{\sigma} \left(\frac{4}{3n}\right)^{1/5}$ selected in our analysis is the tradeoff between the complexity of calculation and mean square error. Figure 1 displays the histograms along with the kernel estimates for the daily S&P 500 of the years marked in Table 3. Since the kernel estimates and histograms coincide in Figure 1, the kernel estimation method works well for the data.



Note: The years included here exhibit significant asymmetry according to the nonparametric test J_{2n} at significance level 5% in Table 3.

Next, we consider a formal statistical test on the symmetry of the distribution. Note that if the distribution is symmetric, then its density function satisfies $f(x) = f(-x)$. Therefore, the hypotheses we are interested in are

$$H_0 : f(x) = f(-x) \text{ vs } H_a : f(x) \neq f(-x). \quad (5)$$

We follow Ahmad and Li's (1997) approach. Specifically, define

$$\hat{I}_n = \int_{-\infty}^{\infty} [\hat{f}(x) - \hat{f}(-x)]^2 dx, \quad (6)$$

where $\hat{f}(x)$ is the estimated density function obtained by (3). It is obvious that if the distribution is symmetric, \hat{I}_n in (6) should be close to zero. Moreover, if there are many observations, then under H_0 , the distribution of the following two standardized versions of \hat{I}_n can be well approximated by the standard normal distribution:

$$J_{2n} = \frac{na^{1/2}}{2\hat{\sigma}} \hat{I}_{2n} \text{ and } J_{2n} = \frac{na^{1/2}}{2\hat{\sigma}} (\hat{I}_{2n} - C_n), \quad (7)$$

where $C_n = \frac{1}{nh\sqrt{2\pi}}$ and $\hat{\sigma}^2 = \frac{1}{2\sqrt{\pi}} \int \hat{f}_n(x) dF_n(x)$ with $F_n(x)$ being the empirical distribution function.

Asymptotically J_{2n} and J_n are equivalent, and their asymptotical p -values can be calculated by the standard Normal distribution. The difference between J_{2n} and J_n is that J_{2n} has a smaller bias for a small sample size as shown by simulation study. We calculated \hat{I}_n in (6), J_{2n} as well as J_n in (7) and summarize the results for all the 26 years in Table 2 and Table 3, respectively.

Table 2: Nonparametric Test with J_n

	CRSP		NASDAQ		S&P 500	
	J_n	p -value	J_n	p -value	J_n	p -value
1980	0.676	0.499	5.632	0.000***	-0.333	0.7384
1981	-0.192	0.847	6.535	0.000***	-0.980	0.3269
1982	2.099	0.035*	-0.852	0.394	2.285	0.022*
1983	-0.691	0.489	1.520	0.128	-0.787	0.431
1984	2.090	0.036*	0.765	0.444	2.347	0.018*
1985	-0.515	0.606	-0.924	0.355	0.145	0.884
1986	0.694	0.487	4.829	0.000***	0.042	0.965
1987	2.211	0.027*	5.506	0.000***	1.519	0.128
1988	-0.901	0.367	2.224	0.026*	-0.915	0.359
1989	0.076	0.939	2.473	0.013*	-0.763	0.444
1990	1.412	0.158	3.643	0.000***	1.788	0.073
1991	0.219	0.826	-0.648	0.516	0.869	0.384
1992	-0.983	0.325	-1.100	0.271	-0.821	0.411
1993	-0.508	0.611	2.299	0.021*	-0.833	0.404
1994	0.466	0.640	1.836	0.066	-0.586	0.557
1995	-0.936	0.349	0.713	0.475	0.046	0.963
1996	0.245	0.806	-0.051	0.958	0.040	0.967
1997	-0.628	0.529	0.411	0.680	-0.891	0.372
1998	0.801	0.418	1.489	0.136	-0.404	0.686
1999	-1.084	0.278	1.238	0.215	-1.032	0.302
2000	-0.362	0.717	-0.773	0.439	-0.774	0.438
2001	-0.673	0.500	-0.210	0.833	-0.651	0.514
2002	-0.617	0.536	0.870	0.383	-0.652	0.514
2003	-0.764	0.444	-1.018	0.308	-0.875	0.381
2004	0.324	0.745	0.988	0.323	-0.141	0.887
2005	0.791	0.428	0.713	0.475	1.149	0.250

Note: The asterisks, * and *** denote statistic significance at the 5%, and .01%, respectively.

Table 3: Nonparametric Test with J_{2n}

	CRSP		NASDAQ		S&P 500	
	J_{2n}	p -value	J_{2n}	p -value	J_{2n}	p -value
1980	1.857	0.063	6.701	0.000***	0.861	0.389
1981	0.982	0.326	7.639	0.000***	0.210	0.8329
1982	3.246	0.000***	0.331	0.7405	3.432	0.000***
1983	0.501	0.616	2.689	0.007**	0.406	0.684
1984	3.235	0.000***	1.909	0.056	3.495	0.000***
1985	0.677	0.498	0.280	0.779	1.337	0.181
1986	1.801	0.071	5.921	0.000***	1.158	0.246
1987	3.147	0.000***	6.402	0.000***	2.461	0.013*
1988	0.166	0.868	3.312	0.000***	0.171	0.863
1989	1.166	0.243	3.586	0.000***	0.337	0.735
1990	2.566	0.010**	4.746	0.000***	2.959	0.003**
1991	1.362	0.173	0.502	0.615	2.017	0.043*
1992	0.179	0.857	0.091	0.927	0.347	0.728
1993	0.633	0.526	3.443	0.000***	0.310	0.756
1994	1.596	0.110	2.966	0.003**	0.564	0.572
1995	0.221	0.824	1.861	0.062	1.195	0.232
1996	1.392	0.164	1.097	0.272	1.194	0.232
1997	0.494	0.620	1.565	0.117	0.248	0.803
1998	1.907	0.056	2.599	0.009**	0.701	0.482
1999	0.120	0.904	2.440	0.014*	0.177	0.859
2000	0.815	0.415	0.417	0.676	0.390	0.696
2001	0.485	0.627	0.954	0.340	0.509	0.610
2002	0.566	0.571	2.083	0.037*	0.529	0.596
2003	0.420	0.674	0.172	0.863	0.306	0.759
2004	1.509	0.131	2.187	0.028*	1.045	0.295
2005	2.001	0.045	1.919	0.054	2.353	0.018*

Note: The asterisks, *, **, and *** denote statistic significance at the 5%, 1%, and .01%, respectively.

Discussion

Analyzing the degree of asymmetry of the broad index, the CRSP value-weighted index using the traditional parametric skewness test would cause a substantial degree of alarm. Nineteen of the twenty-six years are statistically significantly skewed at the .05 level of significance. Ten are significant at the .001 level. This would lead us to wonder if we should throw out a substantial number of our old research journals in which the research design was based upon an assumption of a symmetric distribution. Luckily, the nonparametric tests indicate that the degree of asymmetry is not as great as the parametric test might imply.

Indeed, we suspect that a relatively high degree of positive kurtosis in stock returns may bias the parametric skewness test. As shown in Tables 3 and 2, the tests based on J_{2n} and J_n for the CRSP data indicates significant asymmetry in only four years and three years, respectively. In Table 2 we can see in the J_n test, using the CRSP data, the significance level is only .05, none are rejected at the .01 or .001 level, as compared to the ten that were rejected at the .001 level of significance in the parametric test. However, in the J_{2n} test, three of the years are significant at the .001 level. Overall, this seems to imply that as a first approximation the domestic market as whole is relatively symmetric; on the other hand, the differences among the tests of skewness and asymmetry are not as great when examining the NASDAQ stocks. Eighteen of the most recent 26 years are statistically skewed when measured by the parametric skewness test. Eight of the 26 years analyzing the NASDAQ distribution are still statistically asymmetric using the nonparametric J_n test. Fourteen of the 26 years using the NASDAQ distribution are statistically significant using the nonparametric J_{2n} test. Asymmetry will be more of a concern when making conclusions about the NASDAQ stocks.

The S&P 500 was statistically significantly skewed in 11 of 26 years. Using the J_n test there was significant asymmetry in only two years and, using the J_{2n} test, there was significant asymmetry in six years. Thus, large capitalization stocks as whole have some degree of asymmetry, but the incidence is still half of what is suggested by the parametric test. Given that a great deal of financial research has been conducted with large capitalization stocks, we can probably put more stock in financial studies that have combined large capitalization stocks with assumptions of symmetry.

Conclusions

Significant asymmetry exists in several years according to both the skewness test and the kernel function test. However, the degree of asymmetry in the CRSP value weighted index and to a lesser degree the S&P 500 data is substantially less than that suggested by the parametric tests.

Our results support referencing previous finance articles and current papers. Even though absolute symmetry is clearly violated, as a first approximation, symmetry seems to be a reasonable assumption. But do read those old articles with a degree of skepticism. Armed with knowledge of the exact degree of asymmetry, you should be in a better position to evaluate the relative value of those older findings. In some cases you may want to replicate the studies using some of today's more powerful statistical tools. In most cases, we suspect the results will not be substantially different.

However, traders using trading strategies that are based upon an assumption of symmetric returns should be wary. In the competitive world of the stock or option trader, a first approximation may not be precise enough.

Acknowledgment

We would like to Dr. Steven Jones, of Samford University, for his helpful comments at the 37th Annual meeting of the Academy of Economics and Finance.

References

- Ahmad Ibrahim A. and Qi Li. 1997. "Testing Symmetry of an Unknown Density Function by Kernel Method." *Journal of Nonparametric Statistics* 7: 279-293.
- Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Heiko Ebens. 2001. "The Distribution of Stock Return Volatility." *Journal of Financial Economics* 61: 43-76.
- Bekaert, Geert and Guojun Wu. 2000. "Asymmetric Volatility and Risk in Equity Markets." *The Review of Financial Studies*, 13: 1-42.
- Black, Fisher. 1976. "Studies of Stock Price Volatility Changes." *Proceedings of the 1976 Meeting of the American Statistical Association, Business and Economical Statistics Section*: 177-181.
- Bollerslev, Tim, Julia Litvinova, and George Tauchen. 2006. "Leverage and Volatility Feedback Effects in High-Frequency Data." *Journal of Financial Econometrics*, 4: 353-384.
- Campbell, John Y., and Ludger Hentschel, 1992. "No News Is Good News: An Asymmetric Model of Changing Volatility in Stock Returns." *Journal of Financial Economics* 31: 281-318.
- Christie, Andrew A. 1982. "The Stochastic Behavior of Common Stock Variances: Value, Leverage, and Interest Rate Effects." *Journal of Financial Economics* 10: 407-432.
- Fama, Eugene F. 1965. "Portfolio Analysis in a Stable Paretian Market." *Management Science*: 11: 404-419.
- Fama, Eugene F. 1976. *Foundations of Finance*, New York: Basic Books.
- French, Kenneth R., G. William Schwert, and Robert F. Stambaugh, 1987. "Expected Stock Returns and Volatility." *Journal of Financial Economics* 19: 3-29.
- Harvey, Campbell R., and Akhtar Siddique. 2000. "Conditional Skewness in Asset Pricing Tests." *Journal of Finance* 55: 1263-1295.
- Glosten, Lawrence R., Ravi Jagannathan, and David E. Runkle. 1993. "On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks," *Journal of Finance*, 48: 1779-1801.
- Jackwerth, Jens C. and Mark Rubinstein, 1996. "Recovering Probability Distributions from Contemporary Security Prices." *Journal of Finance*, 51: 347-369.
- Kraus, Alan and Robert H. Litzenburger. 1976. "Skewness Preference and the Valuation of Risk Assets." *Journal of Finance* 31: 1085-1100.
- Markowitz, Harry M. 1952. "Portfolio Selection." *Journal of Finance* 7: 77-91.
- Mandelbrot, Benoit B. 1963. "The Variation of Certain Speculative Prices." *Journal of Business* 36: 394-419.
- Nelson, Daniel B. 1991. "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica* 59: 347-370.
- Rubinstein, Mark. 1994. "Implied Binomial Trees." *Journal of Finance*, 49: 771-818.
- Silverman, Bernard W. 1985. "Density Estimation for Statistics and Data Analysis." *Chapman and Hall*: New York.
- Wu, Guojun. 2001. "The Determinants of Asymmetric Volatility," *Review of Financial Studies* 14: 837-859.

Socioeconomic Factors and Educational Attainment: Evidence from South Carolina

Muhammad Mustafa and Rahul Singh, South Carolina State University

Abstract

The primary objective of this study is to analyze the impact of socioeconomic factors on the educational attainment in South Carolina. SAT scores (math and verbal), and percentage of high school graduate rate are used to measure educational attainment. The results indicate that per capita income, dollar spent per student and average teachers salary have significant positive impacts on SAT scores and percentage of population below the poverty level has a significantly negative effect on educational attainment.

Introduction

There is growing concern with education in the United States as a result of reported educational outcomes and their implications for U.S. competitiveness in the world economy. The educational attainment of young adults in the U.S. has not increased in the last 15 years, according to 2000 census data. Attainment levels vary geographically, especially within urban areas. One in seven students does not finish high school, and fewer than one in four earns at least a college degree. Since education is provided at the state and local level, all states and local school districts have been constantly trying to improve the quality of education. Education improves the skills, knowledge and overall learning abilities of the state citizens. This, in turn, increases individual productivity and quality of labor force. States and communities are interested in improving the quality of their public education in order to enhance community competitiveness in attracting new industry and, hence, jobs.

To measure educational attainment, most of the previous studies employed SAT scores, and high school graduation rates. The SAT originated in 1926 to provide colleges with information that was not evident from applicants' high school grades or achievement tests. Colleges wanted to identify students who have learning potential by measuring their innate comprehension and reasoning ability. To factor out the quality differences among high schools, the SAT included questions on material that was either taught in all high schools or taught in none. There is widespread dissatisfaction with the quality of public education in South Carolina as the state ranked last versus other Southeastern states in SAT verbal and SAT math in 2009. South Carolina had the lowest-percentage of high school graduates (66%) in 2009. To improve the quality of education, the South Carolina legislators passed the Education Improvement Act (ETA) in 1984. Among several improvement measures, this act included teacher salary and per pupil spending increases. The relative decline of student scores on the Scholastic Aptitude Test (SAT) during the past two decades has received attention in both the popular and academic press. Several studies examined the educational outcomes. The empirical evidences in literature are mixed and suggestive.

Literature Review

Several studies have been conducted to examine the impact of different socioeconomic factors on educational attainment. Some of these studies are reviewed briefly.

Jackson. (2009) showed that Meaningful relationships between early-life health and educational attainment raise important questions about how health may influence educational success in young adulthood and beyond, as well as for whom its influence is strongest. Using the data from the National Longitudinal Survey of Youth 1997, this study examined. how adolescents' health and social status act together to create educational disparities in young adulthood, focusing on two questions in particular. First, does the link between adolescent health and educational attainment vary across socioeconomic and racial/ethnic groups? Second, what academic factors explain the connection between adolescent health and educational attainment?

Luisa and others. (2004) assessed the prevalence of periodontitis by education and income levels among US adults with data from the third National Health and Nutrition Examination Survey. They found out that Blacks with higher education and income levels had a significantly higher prevalence of periodontitis than their white and Mexican-American counterparts and that the relationship between income level and periodontitis was modified by race/ethnicity. Neves (2010) investigated the role of the urban bias in the Brazilian development on the educational attainment process. His results are based on a national-wide probability sample survey developed by Brazilian Institute of Geography and Statistics. Ordinary Least Square (OLS) regression models are estimated to assess the hypotheses.

Currie and others (2003) examines the effect of maternal education on birth outcomes using Vital Statistics Natality data for 1970 to 1999. He also assesses the importance of four channels through which maternal education may improve birth outcomes: use of prenatal care, smoking, marriage, and fertility. In an effort to account for the endogeneity of educational attainment, he use data about the availability of colleges in the woman's county in her seventeenth year as an instrument for maternal education.

Card and Krueger (1992) indicate that a decrease in pupil-teacher ratio has significant effect on the rate of return from education. Sander (1992) examines educational attainment outcomes in Illinois for the 1989-1990 school. The results suggest that an increase in average teacher's salary increases ACT scores and the percentage of college bound and an increase of pupil-teacher ratio reduce the graduation rate and the percentage of college bound. States that operated segregated school systems before 1954 typically had lower pupil/teacher ratios, longer term lengths, and higher teacher salaries in white schools than in black schools (see Card and Krueger 1992).

Sander and Krautman (1991) did not find any significant relationship between school expenditures and ACT scores. On the other hand, several studies find evidences that suggest that expenditures can have substantial effects on achievement. Tennessee State Department of Education (1991) finds a reduction in the pupil-teacher ratio significantly increases test scores- particularly for minority students. Hanushek (1986) indicate that the quantitative characteristics of schools and teachers are not strongly related to academic achievement. His study suggests that there is no systematic relationship between school expenditures and student performance.

Kurth (1987) concludes that teacher unionization is the most significant factor contributing to the decline in SAT performance. Kurth's model relies on the traditional monopoly view of teachers' unions, where-by union-induced productivity declines are translated into less student learning and, hence, lower exam scores. Miller (1974) concluded that regardless of the extent of such potential bias, minority students are more likely to be measured as below average or disadvantaged than are white students. In Miller's opinion, Ebert and Stone are correct that below average students perform less well in a unionized setting. Their study expects to find a negative racial effect between teacher unionization and, student achievement test scores.

Krein et.al (1988) control variables include mother's education and age at child's birth, and child's age, gender and race. He concluded that children who ever lived in a female-headed family are about 12 percent less likely to have completed high school by ages 18-20 than children who never did; controlling for income, they are 9 percent less likely.

McLanahan (1985) found that growing up in a single parent family also appears to affect educational outcomes, negatively, although much of this effect appears to result from the lowered family income associated with single parenthood. Blacks have substantially lower level of completed education than whites, but the independent effect of race often disappears when controls are introduced for a variety of socioeconomic background characteristics. From the results of various studies it was concluded that being poor and in welfare family for three of the four years during ages 12 to 15 decreases the probability of graduating from high school to .73, an 11percent reduction from the base probability of .82. Blake et. al (1981) research showed that increased family size is negatively related to educational attainment because it reduces or dilutes the time and income inputs available to reach each child. There is a significant negative effect of number of years spent in a single-parent family on educational attainment for all groups but white women. Family income during the high school years has a positive effect on the education of all groups, but the effect is significant only for white males. Mother's education has a large positive impact on the education of children for all race-sex groups. Freeman et.al (1974) concluded that in recent studies, the findings have consistently shown that those who ever lived in female-headed families complete fewer years of schooling than those who always lived in two-parent families.

Haveman and Wolfe (1984) concluded that the failure to complete high school is a primary characteristic of those experiencing persistent poverty and living in the urban underclass. In addition, graduation from high school is associated with productivity in a variety of household and other nonmarket activities. Educational attainment is perhaps the most common and most basic outcome on which researchers have concentrated, on the presumption that schooling is an important contributor to a wide variety of subsequent behaviors and attainments.

It is clear from Table A1, South Carolina is at the bottom among Southern States in high school graduate rates, SAT – Verbal and Math scores. There is no similar study done for South Carolina.

Methodology

Educational outcomes are affected by several factors including teacher salary, teacher experience, teacher certification, poverty and income of the family, pupil-teacher ratio, expenditure per pupil, qualified teachers, and parents attending conferences. Descriptive statistics and multiple regression models are used to analyze the influence of socio economic factors. It may be mentioned that educational production studies using regression models have two limitations. First, educational attainment is difficult to measure because students have innate abilities. Outcome may occur because of student abilities rather than educational system. Second, educational attainment variables are measured at point in time. It is quite possible that some of the variables changed from one year to next year. The Standardized Aptitude Test (SAT) is a test that

is frequently taken by college bound students to predict success in the freshman year in college. SAT does not predict other types of achievement such as success in the labor market. For this reason, high school graduation rates and the percent planning to attend college are also estimated because education attainment is an important human capital variable that has a significant and substantial effect on earnings.

Data

This study used annual pooled cross-section data of the 46 counties and 92 school districts of South Carolina for the period (2001-2008). SAT scores, high school graduation rates, and parents attending conferences are estimated from school districts data. Average teacher salary, expenditure per pupil, pupil-teacher ratio, percentage of high qualifies teachers, student teachers ratio, dollar spent per student data were collected from the South Carolina Educational Department Profile. Per capita income, percent of population below poverty level are obtained from South Carolina Statistical Abstract.

Statistical Model

To evaluate the factors that influence educational attainment, three multiple regression models will be estimated using a linear and log forms. The explanatory variables are selected based on the previous empirical research on educational attainment. The empirical models are specified as follows:

$$SATV = \alpha_1 + \alpha_2 PCI_{jt} + \alpha_3 PLBPL_{jt} + \alpha_4 PAC_{jt} + \alpha_5 STS_{jt} + \alpha_6 STR_{jt} + \alpha_7 DSPTS_{jt} + \alpha_8 HQT_{jt} + \varepsilon_{1t} \quad (1)$$

$$SATM = \beta_1 + \beta_2 PCI_{jt} + \beta_3 PLBPL_{jt} + \beta_4 PAC_{jt} + \beta_5 STS_{jt} + \beta_6 STR_{jt} + \beta_7 DSPTS_{jt} + \beta_8 DSPTS_{jt} + \beta_9 HQT_{jt} + \varepsilon_{1t} \quad (2)$$

$$HSGR = \gamma_1 + \gamma_2 PAC_{jt} + \gamma_3 PCI_{jt} + \gamma_4 PLBPL_{jt} + \gamma_5 STR_{jt} + \gamma_6 DSPTS_{jt} + \gamma_7 ATS_{jt} + \gamma_8 STS_{jt} + \gamma_9 HQT_{jt} + \varepsilon_{1t} \quad (3)$$

where,

- i** = the suffix for county (i = 1,2...46);
- j** = the suffix for district (j = 1,2...92);
- t** = the suffix for time (t = 2001 to 2008) for model 1;
- $\alpha_1 - \alpha_8$** = parameters to be estimated for model 1;
- $\beta_1 - \beta_9$** = parameters to be estimated for model 2;
- $\gamma_1 - \gamma_9$** = parameters to be estimated for model 3;
- HSGR** = High School Graduation Rate;
- SATV** = SAT Verbal Scores;
- SATM** = SAT Mathematic Scores;
- PCI** = Per Capita Income (in dollars);
- PLBPL** = People Living Below Poverty Level (in percentage);
- STR** = Student/Teacher Ratio;
- DSPTS** = Dollars Spent Per Student (in dollars);
- ATS** = Average Teacher Salary (in dollars);
- HQT** = High Qualified Teachers (in percentage);
- STS** = Spent on Teacher Salary (in percentage);
- PAC** = Parents Attending conferences (in percentage); and
- ε_1** = Disturbance Term.

The models are estimated using pooled cross-sectional school district level data for the years 2001-08. This study does not include teacher experience, percentage taking SAT, teacher turnover ratio and schooling of parents due to lack of

available data. It is reasonable to expect that per pupil expenditure would have a positive effect on educational attainment at least up to some point. However, additional expenditures on items such as transportation and administrative bureaucracy may have negative impact on attainment. Teacher salary is expected to have positive impact on educational attainment as teacher salary increases teachers will be motivated. Also, better teachers will be recruited with better salary.

The demographic variables are included in the model to account for environmental factors that may influence student performance. The percentage of population below poverty level (BPL) and percent of black population (PEP) in a county may have negative impact on educational attainment. The households with low income may not provide necessary support to facilities for student performance in schools. The pupil-teacher ratio is expected to have negative impact on educational attainment. Teachers with bigger class-size cannot pay close attention to students needs.

Results

Table 1 reports the descriptive statistics. The correlation between SAT Verbal and SAT Math is very high (0.88). However correlation between student /teachers ratio (STR) and high school graduate rate is negative (-0.083). (A correlation matrix is available upon request from the authors).

Table 2 presents the parameter estimates of model 1, where the dependent variable is SAT-Verbal. In this model adjusted R2 is 0.5478, which is relatively reasonable for this type of model. The F-Value is statistically significant for 1 percent level. The coefficient of per capita income (PCI) with expected sign is significant at the 1 percent level. The coefficient of percentage of people below poverty level (PLBPL) and dollar spent per student (DSPS) are significant with expected signs. The coefficient of parents attending conferences (PAC), dollar spent on teachers salary (STS), and high qualified teachers (HQT) have expected sign but are not significant.

Table 3 reports the estimate of model 2, where the dependent variable is SAT-Math. In this table, R2 is 0.455, which is not relatively high. The overall model is good as the F-Value is significant at 1 percent level and there is no autocorrelation. The coefficient of people living below poverty level, dollar spent per student and average teachers salary have expected signs and are significant. The coefficient of per capita income, parents attending conferences, student teachers ratio, dollars spent on teacher salary and highly qualified teachers have expected sign and are not significant.

Table 4 provides the results of high school graduation rates (HSGR). In this model adjusted R2 is 0.1576 which is very low. The coefficients of most of the variables have unexpected signs and are not significant except the coefficient of people living below poverty level which has expected sign and is significant.

Summary, Conclusion and Limitations

This study used pooled cross-sectional data of 46 counties and 92 school districts of South Carolina to analyze the factors that affect educational attainment. The results of this study find some support of the hypothesis that expenditures on teacher salary and per pupil expenditures can have significant effect on educational attainment. In the estimates of SAT-Math scores, teacher salary is significant. The per-pupil expenditure is positive and significant in both SAT-Math and SAT-Verbal Scores. These results are in agreement with Hanushek and Sander and Krautman and in sharp disagreement with Sander. The pupil-teacher ratio is positive and insignificant in all estimates except in the college bound estimate, where it is positive and significant. This finding is in conflict with Card and Krueger and Sander, but it is in agreement with Hanushek.

The demographic variables such as percentage of the population below poverty level have predicted negative impact on educational performance in all models. This indicates that environmental factors, which are beyond the control of school board, need to be changed if student performance is to improve. This study suggests that until some of the social problems are dealt with, increased findings for school education may not have the impact that school administration and tax payers hope for.

It may be mentioned that the use of SAT scores as a measure of educational attainment has a few limitations. One limitation of SAT scores as an indicator of the quality of education is that students taking the test are unrepresentative of all high school students because they tend to be academically superior. It may tell a great deal about excellence in education, but reveal very little about the average quality of education. College entrance exam scores fall when the proportion of students taking the exam increases. This happened during the 1960s when baby-boomers swelled the ranks of high school graduates by 50 percent and a larger proportion of graduates had college aspirations.

Table 1: Descriptive Statistics

	SAT_V	SAT_M	HSGR	PCI	PLBPL	PAC	STS	STR	DSPS	ATS	HQT
Mean	463.548	472.793	78.107	24146.86	17.255	80.7	55.801	24.3	6792.93	40889.67	89.008
Median	467	477.5	78.3	23383	16.3	84.116	55.862	24.656	6498	40892.15	90.26
Maximum	553	565	97.15	45427	38.3	100	68.125	38.4	20341.67	50268	100
Minimum	366	342	50.882	16084	9	21.2	26.85	5.833	3605.5	29436.33	44.2
Std. Dev.	33.4196	35.627	8.657	4662.107	5.205	14.386	4.3603	4.188	1598.207	3630.664	6.619
Skewness	-0.44	-0.637	-0.185	1.008	1.09	-1.019	-0.909	-0.542	2.391	-0.28358	-2.476
Kurtosis	2.537	3.025	3.062	4.557	4.523	3.7808	8.436	4.742	17.734	3.307245	13.65
Jarque-Bera Probability	15.193 0.0005	24.934 0.000004	2.171 0.337	99.626 0	108.575 0	73.093 0	503.845 0	64.58 0	3679.605 0	6.379722 0.041	2116.681 0
Sum	170586	173988	28743.55	8886046	6350.1	29697.63	20534.98	8942.507	2499798	15047398	32755.23
Sum Sq. Dev.	409891.1	465834.3	27510.51	7.98E+09	9942.868	75958.69	6977.58	6437.405	9.37E+08	4.84E+09	16083.32
Observations	368	368	368	368	368	368	368	368	368	368	368

Table 2: Dependent Variable: SAT Verbal Scores (SATV)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	481.8937	30.85586	15.61758*	0.0000
PCI	0.000813	0.000338	2.403389**	0.0168
PLBPL	-4.282855	0.282601	-15.15513*	0.0000
PAC	0.138321	0.097621	1.416929	0.1574
STS	0.129706	0.303503	0.427365	0.6694
STR	0.048411	0.402880	0.120162	0.9044
DSPS	0.001725	0.000848	2.032753**	0.0428
ATS	-3.88E-05	0.000386	-0.100460	0.9200
HQT	0.069954	0.217436	0.321723	0.7478
R-squared	0.547755			
F-statistic	54.35230			

Significant at .05 level, ** Significant at .01 level.

Table 3: Dependent Variable: SAT Mathematics Scores ((SATM)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	442.5228	35.76117	12.37439*	0.0000
PCI	0.000357	0.000393	0.908023	0.3645
PLBPL	-4.129433	0.319636	-12.91917*	0.0000
PAC	0.154101	0.122039	1.262721	0.2075
STS	0.262276	0.356986	0.734694	0.4630
STR	0.068320	0.424352	0.160997	0.8722
DSPS	0.002060	0.000920	2.239179**	0.0258
ATS	0.001030	0.000453	2.276224**	0.0234
HQT	0.090484	0.231294	0.391208	0.6959
R-squared	0.455637			
F-statistic	37.56083			

Significant at .05 level, ** Significant at .01 level.

Table 4: Dependent Variable: High School Graduation Rates (HSGR)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	115.8223	9.620698	12.03886*	0.0000
PCI	-0.000436	0.000109	-4.005040*	0.0001
PLBPL	-0.614436	0.098715	-6.224346*	0.0000
PAC	-0.036398	0.035646	-1.021103	0.3079
STS	0.004657	0.100314	0.046420	0.9630
STR	-0.076837	0.117305	-0.655020	0.5129
DSPS	-0.000153	0.000311	-0.492648	0.6226
ATS	-0.000226	0.000154	-1.474094	0.1413
HQT	-0.019662	0.071228	-0.276040	0.7827
R-squared	0.157629			
F-statistic	8.397260			

Significant at .05 level, ** Significant at .01 level.

References

- Blake, J. (1981). Family Size and the Quality of Children. *Demography*, 18, 421
- Borrell Luisa N, Burt Brian A, Neighbors Harold W, and Taylor George W. (May 2004). *American Journal of Public Health*. Washington: May 2004. Vol. 94, Iss. 5; pg. 748-755.
- Breland, H.M. (1977). Family Configuration Effects and the Decline in College Admissions Test Scores: A Review of the Zajonc Hypothesis. College Entrance Examinations Board.
- Card, D. and A. Krueger (1992). Does School Quality Matter? *Journal of Political Economy*, 100, 1-40.
- Currie, Janet, Enrico Moretti. (2003) Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings. *The Quarterly Journal of Economics*. Cambridge: November 2003. Vol. 118, Iss. 4; pg. 1495
- Freeman, Richard B. (1976). *Black Elite- The New Market for Highly Educated Black Americans*. McGraw- Hill, Book Company, 41-45.
- Hanushek, E.A. (1986), *The Economics of Schooling*. *Journal of Human Resources*, 24, 1141-1177.
- Haveman, R., and B. Wolfe (1984). Schooling and Economic Well-Being: The Role of Non-Market Effects. *Journal of Human Resources*, 19, 377-407.
- Jackson, Margot I. Understanding Links Between Adolescent Health And Education Attainment. *Demography*, Silver Spring: November 2009. Vol. 46, Iss. 4; pg. 671-685.
- Krein, Sheila F.(1988). Growing Up in a Single Parent Family: The Effect on Education and Earnings of Young Men. *Family Relations*, 35, 161-168.
- Kurth, Michael (1987). Teachers' Unions and Excellence in Education: An. Analysis of The Decline in \$AT Scores. *Journal of Labor Research*, 8, 351+.
- McLanahan, S. (1983). The Effects of the Absence of A Parent on The Educational Attainment of Offspring. Institute For Research on Poverty Discussion Paper, 72083.
- Miller, LaMar (1974). *The Testing of Black Students: A Symposium*. Englewood Cliffs, NJ: Prentice Hall.
- Neves, Jorge Alexandre. (2010). Urban Bias in Development and Education Attainment in Brazil. *The Journal of Developing Areas*, Vol. 43, November 2010. pg. 271-288
- Sander, W. and A.C. Krautman (1991). Local Taxes, Schooling and Jobs in Illinois. *Economics of Education Review*, 10, 111-121.
- Sander, William (1992). Expenditures and Student Achievement in Illinois. *Journal of Public Economics*, 403-416.
- Tennessee State Department of Education (1990). Student/Teacher Achievement Ratio.

Student Performance in the Economics Classroom

Richard Vogel and Wisdom Akpalu, Farmingdale State College

Abstract

Using students' final numerical course grade across a sample of six economics courses we assess the impact of several factors on student performance, including seat location, student college classroom experience, cumulative gpa, employment status, hours spent working, average hours spent studying, gender, student major, and age. We found that cumulative gpa, age and attendance impact student performance but seat location preference and employment status were not statistically significant. However, employment status as well as cumulative gpa determine the choice of seat location in the classroom.

Introduction

Student performance in the economics classroom is affected by a wide range of factors including ability, class size, previous course work in the subject, mathematics background, attendance and motivation, learner personality and characteristics. The conclusions drawn from previous studies regarding the importance and impact of these factors on performance have been mixed. Understanding this relationship is important to institutions and instructors alike, in order to better provide both with greater guidance for allocating and devoting instructional resources.

Recent studies related to the question of seat location include Armstrong and Chang's (2007) analysis of large multi-section biology classes and Benedict and Hoag's (2004) study of principles of economics courses. The conclusions from the study of biology classes suggest that seat location does not have an impact on student performance. Benedict and Hoag's analysis arrives at exactly the opposite conclusion seat location does impact student performance. The specific variables used to measure seat location differ slightly between the two studies, where one uses actual seat location in science courses (both student self-selected and randomly instructor assigned) and the other used student seat location preferences. Additionally, the study on biology classes looks at grades from machine scored exams, while Benedict and Hoag's study uses course grade as the dependent variable. Underlying the analysis of both studies is the question of whether seat location affects performance, i.e. students in the back of the room may face certain disadvantages as a result of their specific location in the classroom, or a behavioral issue, e.g. more motivated students sitting closer to the front of the classroom.

Park and Kerr (1990) conclude that the key determinants of academic performance are cumulative GPA and percentile rank on ACT/entrance exams, with other factors such as attendance, and students' value of the course having a positive but lesser impact on outcomes. GPA and standardized entrance exam scores in their analysis are indicators of intelligence and effort. Their findings also suggest that variables such as age, gender, and hours worked out of school have little impact on student performance.

Borg and Stranahan (2002a; 2002b) suggest that personality type as measured by the Meyers-Brigg instrument is an important factor in student academic performance in economics. Introverted personality types had better grades than extroverts. They also found that race and gender did matter, except that it was not a simple or direct relationship but instead combined with temperament/personality type.

Taking up the often studied issue of class size, Arias and Walker (2004) conclude that class size does have an impact on performance in the economics classroom, reinforcing the results of analysts such as Lopus and Maxwell (1995). This is in direct contrast to studies such as those by Siegfried and Kennedy (1995), Kennedy and Siegfried (1997), and Siegfried and Walstad (1998), which suggest that class size does not matter especially when looking at the principles class.

Cohn and Johnson (2006) look specifically at the relationship between class attendance and student performance as measured by exam scores. One of their chief concerns is the issue of ability and student motivation versus an attendance model. In other words, is there an indirect or direct relationship between attendance and performance? Their analysis suggests that performance is related to attendance across the whole semester. Stanca (2006) concludes though that class attendance, effort and motivation are interrelated and that while attendance was statistically significant, it is open to debate how important it is to overall performance.

In a recent study Wenz and Yu (2010) investigated the impact of term-time employment on academic performance of undergraduates and found employment to have a modest negative impact on students' grades. Furthermore, students who worked for financial reasons were found to earn higher grades than their counterparts who have a general desire to work but lower grades than those who work for career specific skills.

Our study contributes to this evolving body of literature by investigating the factors that are likely to impact student performance in economics classes. While initially we were most concerned with the question of a student's choice of seat location and performance, we incorporate a number of other factors as well including student experience, prior academic

performance (cumulative GPA as a measure of intelligence), students' use of time (academic and non-academic such as employment status), attendance, gender, and class size. While some studies have focused on scores from standardized exams, we focus on students final numerical grades. Instructors can control or implement many different strategies in the classroom that impact learning outcomes. If factors such as seat location and attendance do significantly affect classroom performance, then strategies such as the use of assigned seating and mandatory attendance policies may lead to improved performance.

The rest of this paper is organized as follows. Next we present an overview of the student population group and classes that are used in the analysis followed by the econometric model and results of our analysis. The conclusions of our analysis are presented after the econometric model.

Students and Courses Studied

Farmingdale is a public college of technology with approximately 6,500 students. It is primarily a commuter campus, with dormitory space to accommodate only about 10 percent of the student population. Many of the degree programs are technical or career oriented, and the college's single largest program is the B.S. in Business Management (formerly named Management Technology), with over 1000 students enrolled. In the last year, the college has added several new programs in the Arts and Sciences including a B.S. in Applied Economics and another in Applied Psychology.

The analysis is conducted across six sections of economics that were taught during the spring semester of 2009. These courses include three sections of Principles of Macroeconomics, one section of Intermediate Macroeconomics, one section of Engineering Economics, and a section of Economics of Global Natural Disasters. The principles courses serve as both required courses for programs such as Applied Economics and Business Management, as well as part of the college's General Education offerings in Social Sciences. All students in the SUNY system are required to take six credit hours in Social Sciences, which may be courses in anthropology, economics, geography, political science, psychology, and sociology.

Intermediate macroeconomics is a required course for the Applied Economics major and also had a substantial number of students taking it as a general elective – including students from both Applied Mathematics and Business Management. Engineering economics is a required course for students in the School of Engineering Technology, with students coming from programs such as architectural, electrical, and mechanical engineering technology, as well as construction management. The course may also be taken by other students including applied economics majors, as a general elective. Economics of global natural disasters is an upper division Arts and Science elective that is taken by students from across the campus and schools. Both the intermediate microeconomics and economics of disasters courses have principles of economics as prerequisites.

Our study is composed of regularly scheduled courses with no prior selection process of students or the use of seating charts, pre and post exams, or standardized exam instruments. Economics was one of the subjects evaluated during the current three year general education cycle, but that process was undertaken during the fall semester of 2008. During the last four weeks of the semester, students in the six courses were surveyed and asked to provide the following information; gender, seat location preference (front - middle - back), age, years in college/status, part-time or full-time student, major, employed (full or part-time), average number of hours work per week, average number of hours study per week, average hours spent on entertainment/leisure per week, and cumulative GPA. Each student completed a short questionnaire in class to provide the required information. The only anticipated bias is the reported cumulative GPA. However, since final grades for the courses are technically independent of the GPA, there is no incentive in misrepresenting the GPA. Moreover, if all GPA are uniformly over exaggerated, the effect on the regression coefficients will be negligible.

The student provided information was then matched up with each individual student's final numerical course grade and semester attendance. Individual names and identifiers were then stripped from the data set, thus insuring that individual students and their grades could not be identified. Only students that actually completed the courses, and fully completed the survey were included in the analysis. Total registered enrollment in the six courses was 196 students. The total number of students surveyed was 158.

Unlike some of the institutions from previous studies, Farmingdale's class sizes tend to be of moderate to small size. Very few classrooms on campus are able to hold more than 40 students, and general labor-management practices on the campus (faculty are covered under the campus' and state wide collective bargaining agreements) have kept classes in general below 48 students. Within the economics department, introductory courses are capped at 40 students, and upper level courses are generally capped at between 32 and 36 students. The three principles courses included in the analysis had beginning total registered enrollments of 39, 32, and 40 students respectively, while the three upper division courses had 22 (intermediate), 38 (engineering), and 25 (economics of disasters) students enrolled. The courses were also spread throughout the day and week, with two of the principles courses meeting twice a week at 8 and 9:30 AM, the other principles course and intermediate course also meeting twice a week at 11 AM and 12:30 PM, the engineering economics course held one afternoon a week from 2 to 4:45, and the disasters course held one evening a week from 6-8:45.

Students were instructed to provide their habitually preferred seating choice for the particular class as opposed to simply the one that they were actually seated in on the day of the survey. It is possible that the student came in late on the particular days that the surveys were administered and consequently were unable to sit where they normally did. Over the course of the semester, we observed that students appeared to be quite consistent and one might even say territorial, in that they tended to sit in the same seat location that they originally selected at the start of the semester.

Farmingdale enrolls a wide range of students, from first-time in college, to mid-career individuals returning to college. Thus we collected both age and college experience data. A priori, it is possible that adults returning to college to complete a degree or embarking upon a degree program after being in the labor market for a number of years may be more serious in their studies than younger students fresh in from high school. On the other hand, younger students may have greater study skills as they have either been continuously in school or had a shorter break from school than mid-career learners. The college experience data essentially captures how many years students have been studying at the college level, with students responding from 1 year to some student responses of 6 years or more. Number of years of college may differ significantly from the student's actual designation as a freshman, sophomore, junior, or senior, especially with the number of part-time students enrolled on campus, as well as the number adult-learners on campus.

The cumulative GPA (at the start of the spring semester) represents a measure of the student's pre-existing effort and intelligence. These are self-reported GPA's, and it is possible that this information may suffer from the Lake Wobegone syndrome (see Maxwell and Lopus 1994) that they are inflated. This is an issue which we plan to address in further study in the future. The other variables all tend to measure students' leisure, work, and study choices. A large percentage of the student population at Farmingdale work either full-time or part-time, and the amount of time spent at work, leisure, and study are all important factors that may affect their final grade outcome.

Attendance was summarized for each student from instructor records and is measured as percentage of the semester's classes that were attended. No attempt is made to discern between excused and unexcused absences or lateness. While presumably, greater class attendance would lead to better overall performance, some of the studies discussed previously do not fully support this thesis. Certainly some type of attendance factor may be embedded in the overall individual instructor course grading mechanism such as a class participation grade, or in class quizzes, etc. that students cannot receive credit for unless they are actually present in the class. Two of the courses here do include a component of this nature as a part of the overall course grade, but it is only a small part of the course grade, with more emphasis placed on other aspects of the course such as out of class written assignments and a multipart course project. The grades in the other four courses in this study, rely primarily on exam scores, but also do include a smaller component of implied credit derived from attendance. Classroom attendance though does not guarantee student learning, as there is also the question of individual student motivation, student preparation for class, and other factors.

Course grade evaluation processes varied between the six courses. In two of the courses, course grades were primarily based upon a series of weekly problem sets and written assignments, and a multi-part course project. The weekly problem sets/written assignments were conducted by the students outside of the classroom, and submitted and discussed in class. The multi-part course project was also completed by students outside of class, submitted in separate parts at different times in the semester, and required students to demonstrate their mastery of the each portion of the subject material presented in class up to that point in time. The other four courses relied upon in-class examinations and weekly assignments. One instructor utilized the online course management system as an adjunct to the course, using it to provide students with course readings, lecture notes, and specific assignment instructions (assignments were submitted to the instructor in the classroom though). The other instructor regularly emailed students copies of lecture notes and assignments.

Theoretical Framework

Score Production Function

Suppose the score or grade that a student obtains at the end of the course depends on his or her seat location preference in class and other attributes (i.e., A), e.g. innate ability. Let the distance from any seat in the class to the professor's desk be measurable on a continuous scale. If the distance to the farthest seat is normalized to one, let $x_i \leq 1$ be the seat preference of student i in a particular class. In addition, let the score obtained depend on the average number of hours that the student studies for the course (i.e., s). The student i 's score (grade) production function is

$$g = g(x, s, A), \text{ with } g_x \leq 0, g_s \geq 0, g_{ss} < 0, g_{sx} < 0 \text{ and } g_{xx} \geq 0 \quad (1)$$

where $g_x (\leq 0)$ is the marginal score/grade productivity with respect to the normalized distance from the professor's desk. In addition, we have hypothesized that this marginal productivity is non-increasing in increased distance from the desk and that the attributes include the age of the student, his or her cumulative GPA, class attendance rate, employment status, the number of years the student has spent in the college and his or her gender.

Seat Location Preference

There is sufficient empirical evidence that returns to college education is generally positive in the labor economics literature (Card, 1999; Katz and Murphy, 1992). As a result, suppose the student derives utility indirectly from the score that he/she obtains in a course. Furthermore, it is possible that a student who sits at the rear row of a class is more likely to engage in leisure rewarding activities, e.g. text, sleep or read literature that is unrelated to the topic under discussion-relative to a student who sits in the front row. As a result, the in-class leisure depends on $(1 - x_i)$. Using equation (1), the corresponding utility function for the student is

$$u = u(g(x, s, A), l) \quad (2)$$

where $l = l(1 - x) = 1 - x$, is the distance to the rear of the classroom. Suppose the individual has a fixed amount of time that he can allocate to working (k) and studying (s). If the time endowment is normalized to one, then $k = 1 - s$. Let the marginal opportunity cost of the time spent on studying be fixed at w . If the individual derives utility from income obtained from working then we can re-specify the utility function as

$$u = u(g(x, s, A), l(1 - x), w(1 - s)) \quad (3)$$

Maximizing equation (3) with respect to x and s gives

$$\frac{\partial u}{\partial x} = u_g \cdot g_x + u_l \cdot l_x = 0 \Rightarrow g_x = u_l / u_g = mrs_{l,g} \quad (4)$$

$$\frac{\partial u}{\partial s} = u_g \cdot g_s - u_s = 0 \Rightarrow g_s = (u_s / u_g) = mrs_{s,g} \quad (5)$$

From equation (4), the student will chose a seating position for which, in equilibrium, the marginal gain in grade will equate his or her marginal rate of substitution (*MRS*) between in-class leisure in class and his or her grade. Equation (5) indicates that, in equilibrium, the student will equate his or her marginal gain in the score to the marginal rate of substitution between the proportion of time allocated to studies and the score/grade.

From equation (4), we can derive the comparative statics of x with respect to the composite index A . Thus

$$(u_{gg} g_x g_A) dA - (u_{gl} \cdot g_x - u_{gg} \cdot g_x^2 - u_g \cdot g_{xx} - u_{ll} + u_{lg} g_x) dx = 0 \quad (6)$$

$$\frac{dx}{dA} = \frac{u_{gg} g_x g_A}{(2u_{gl} \cdot g_x - u_{gg} \cdot g_x^2 - u_g \cdot g_{xx} - u_{ll})} \leq 0 \quad (7)$$

From equation (6), given that an attribute improves scores, a student with such an attribute is more likely to sit in the front row relative to the rear row. This is empirically tested.

Empirical Model and Estimations

From the score production function, we can specify a Cobb-Douglas production function of the form

$$g_i = \alpha_0 + \alpha_1 x_i + \alpha_2 s_i + \beta A_i + \varepsilon_i \tag{8}$$

where α_0 , α_1 , α_2 and β are constants and ε_i is a normally distributed error term (i.e. $\varepsilon_i \sim N(0, \sigma)$). The dependent variable (i.e., grade) as well as the continuous explanatory variables e.g. cumulative GPA and age are logged. To estimate equation (7), we administered a questionnaire to 158 students taking principles of economics, intermediate macroeconomics, sports economics and engineering economics. The descriptive statistics of the data collected is presented in Table 1. The mean percentage score of the students is 79.05% and the mean cumulative GPA is above 3.0. On the average, the students spent less than a quarter of the time spent on working on studies. Furthermore, 40% of the students sit in front. Our sample has more females (59%) than males (41%) and less than 8% were unemployed or did not work. Of the number employed, more than a third work fulltime. Furthermore, only a quarter of the respondents were freshmen and the average age is approximately 22 years. In addition, approximately a quarter of the students had class attendance rates of 95% or higher.

Table 1: Descriptive Statistics of Variables

Variable	Mean	Standard Dev.
grade (percentage score)	79.0453	12.003
Cumulative GPA	3.1645	0.4362
Front seat(=1, 0 otherwise)	0.4118	0.4938
Ratio of hours studied to hours worked	0.2215	0.4166
Attendance more than 95% (=1, 0 otherwise)	0.2595	0.4398
Age of students (continuous)	21.7468	5.375
Female (=1, 0 otherwise)	0.5886	0.4937
At least Sophomores (=1, otherwise)	0.7468	0.4362
Employed (=1, otherwise)	0.9241	0.2658
Employed fulltime (=1, otherwise)	0.3380	0.4747
Observations	158	

Table 2 has the results of the estimated score production function using the ordinary least square (OLS) method with robust standard errors. Note that the number of observations used for the regression is 130, not the full sample of 158 due to missing observations in the data set. The R-squared indicates that about 18% of the variability of the dependent variable is explained by the regressands, an indication that the line is a good fit for a cross sectional data of this nature. The variables that are statistically significant in explaining the scores are the CGPA, attendance rate and age. From the elasticity coefficients, a 10% increase in the CGPA, on the average, increases the score mark by 4.3%. This implies that, all other things being equal, innate ability is important in determining the overall course grade. The coefficient is statistically significant at 1% level of significance. Secondly, students who attended the classes more than 95% of the time have a modest higher course grade, albeit statistically significant only at 10%. Furthermore, relatively older students performed better than younger ones. Finally, seating in front is positively related to higher performance, as expected, but the coefficient is not statistically significant. This result adds to the inconclusive evidence in the literature that seat location preferences may not necessarily explain student performance. In addition, impact of employment status on academic achievement is not significant. Although this contrasts the finding of Wenz and Yu (2010), they found a very low impact of about 0.007 points per work hour. As a result, the relationship between employment status and student performance needs to be further researched. Finally, we controlled for the possibility of differences in mean performance across the courses sampled for the study but found no significant effects.

Table 2: Ordinary Least Squares Estimation of Determinants of total marks for Spring 2009 Semester

Explanatory Variables	Regression 1	Regression 2	Elasticity
log (CGPA)	0.419(0.105)***	0.43(0.098)***	0.4301
Front seat (=1)	0.01 (0.029)	0.006 (0.027)	0.0006
Study more hrs than work	0.042 (0.032)		
Attendance more than 95% (=1)	0.049 (0.027)*	0.052(0.025)**	0.0028
log (Age of Student)	0.125 (0.058)**	0.125(0.056)**	0.1253
Female	0.008 (0.033)		
At least Sophomore (=1)	0.021 (0.037)		
Employed (=1)	0.05 (0.044)		
Constant	3.403 (0.184)***	3.467(0.149)***	
Observations	130	130	
R-squared	0.19	0.18	

*significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors in parentheses

Furthermore, we specify the empirical relationship between the *demand* for a seat in class and its determinants. The empirical equation is

$$x_i = \delta + \theta A_i + v_i \tag{9}$$

where δ and θ are constants and v_i is a normally distributed error term (i.e. $v_i \sim N(0, \sigma_v)$). The estimated results are presented in Table 3 below. From the results, the seat location preference is determined by the employment status of the student and his or her CGPA. Students who are employed fulltime prefer the front seats to the rear ones and the probability that a student who is employed fulltime will seat in front is 0.2 more than those who are not fully employed. Moreover, a 10% increase in CGPA increases the probability that a student will sit in front by 24%. The Pseudo R-Square indicates that about 10% of the variability of the decision to choose a front seat or otherwise is explained by the explanatory variables.

Table 3: Logit regression for the Choice of Front Seat

Explanatory Variables	Coefficients	Elasticity
log (Age of Student)	-0.014 (0.01)	0.0083
Employed fulltime(=1)	1.224 (2.75)***	0.2722
log (CGPA)	4.019 (2.51)**	2.3973
Constant	-5.418 (1.58)	
Observations	118	
Pseudo R-Squared	0.10	

** significant at 5%; *** significant at 1%. Robust z statistics in parentheses

Conclusions

We find that the principal determinants of student performance are innate ability measured by gpa, attendance and age. While seat location was not statistically significant in the student performance equation, a separate logit analysis of seating preferences indicated a positive relationship with student gpa and fulltime employment. Innate ability and motivation are important determinants of student performance.

One factor that we were unable to control for in this study was physical classroom space, as the six classes used for the analysis were held in rooms that may have varied in configuration and physical size. In future work, we plan on expanding the sample size to a larger and broader range of economics courses, and refining the survey instrument to obtain some additional information from students. The results of our analysis are not strong enough to suggest enforced attendance policies, the use of assigned seating, and other similar strategies.

Acknowledgement

The authors would like to acknowledge participants of conference sessions at the Academy of Economics and Finance and the New York State Economics Association Annual Meeting. We also acknowledge and thank both Dean Frank Pellegrini and the State University of New York Joint Labor Management Individual Development Awards Program for their support in conducting and presenting this research.

References

- Arias, J.J. and Douglas M. Walker. 2004. "Additional evidence on the relationship between class size and student performance." *Journal of Economics Education* (Fall): 311-329.
- Armstrong, Norris and Shu-Mei Chang. 2007. "Location, location, location: does seat location affect performance in large classes." *Journal of College Science Teaching* 37(2): 54- 58.
- Arulampalam, Wiji, Robin A. Naylor and Jeremy Smith. 2007. "Am I missing something? The effects of absence from class on student performance." Warwick Economics Research Papers No 820.
- Benedict, Mary Ellen and John Hoag. 2004. "Seating location in large lectures: are seating preferences or location related to course performance?" *Journal of Economics Education* (Summer): 215- 231.
- Borg, Mary O. and Harriet A. Stranahan. 2002a. "Personality type and student performance in upper-level economics courses: The importance of race and gender." *Journal of Economics Education* (Winter): 3- 14.
- Borg, Mary O. and Harriet A. Stranahan. 2002b. "The effect of gender and race on student performance in principles of economics: the importance of personality type." *Applied Economics* 34: 589-598.
- Card, David. 1999. "The causal effect of education on earnings." In *Handbook of Labor Economics*, vol. 3A, edited by David Card and Orley Ashenfelter, Amsterdam; New York and Oxford: Elsevier Science, North-Holland.
- Cohn, Elchanan and Eric Johnson. 2006. "Class attendance and performance in principles of economics." *Education Economics* 14(2): 211-233.
- Eskew, Robert K. and Robert H. Faley. 1988. "Some determinants of student performance in the first college level financial accounting course." *The Accounting Review* 63(1): 137- 147.
- Grant, Darren. 2007. "Grades as information." *Economics of Education Review* 26(2007): 201-214.
- Katz, Lawrence and Kevin Murphy. 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics* 107: 35-78.
- Kennedy, P.E. and J. Siegfried. 1997. "Class size and achievement in introductory economics: Evidence from the TUCE III data." *Economics of Education Review* 16 (4): 385-94.
- Kherfi, Samer. 2008. "Economic education in the Middle East: are the determinants of success in introductory economics any different?" *Journal of Economics Education* (Winter): 22- 40.
- Maxwell, N. L., and J. S. Lopus. 1994. "The Lake Wobegon effect in student self-reported data." *American Economic Review. Papers and Proceedings* 84(2): 201-05.
- Park, Kang H. and Peter M. Kerr. 1990. "Determinants of academic performance: A multinomial logit approach." *Journal of Economics Education* (Spring): 101-111.
- Pyne, Derek. 2007. "Does the choice of introductory microeconomics textbook matter?" *Journal of Economics Education* (Summer): 279- 296.
- Rich, Steven P.. 2006. "Student performance: does effort matter?" *Journal of Applied Finance* (Fall/Winter): 120-133.
- Romer, David. 1993. "Do students go to class? Should they?" *Journal of Economic Perspectives* 7: 167-174.
- Siegfried, J. and P.E. Kennedy. 1995. "Does pedagogy vary with class size in introductory economics?" *American Economic Review. Papers and Proceedings* 82(2): 347-51.
- Siegfried, J. J., and W. B. Walstad. 1998. Research on teaching college economics. In W. B. Walstad and P. Saunders, eds.. *Teaching undergraduate economics: A handbook for instructors*, 141-66. New York: Irwin/McGraw-Hill.
- Stanca, Luca. 2006. "The effects of attendance on academic performance: panel data evidence for introductory microeconomics." *Journal of Economics Education* (Summer): 251- 266.
- Stinebrickner, Todd R. and Ralph Stinebrickner. 2007. "The causal effect of studying on academic performance." University of Western Ontario, CIBC Human Capital and Productivity Working Papers 200702. University of Western Ontario, CIBC Human Capital and Productivity Project. found at <http://ideas.repec.org/p/uwo/hcuwoc/20072.html>.
- Trine, J. Arden and Michael H. Schellenger. 1999. "Determinants of student performance in an upper level corporate finance course." *Proceedings of the Academy of Educational Leadership* 4(1): 91-99.
- Wenz, M. and Wei-Choun Yu. 2010. "Term-Time Employment and the Academic Performance of Undergraduate Students." *Journal of Education Finance* 35(4): 358-373.