

*The*  
**GETTYSBURG  
ECONOMIC  
REVIEW**

**VOLUME 5**

**SPRING 2011**



DEPARTMENT OF ECONOMICS  
GETTYSBURG COLLEGE  
GETTYSBURG, PENNSYLVANIA 17325

The Economics Department and Omicron Delta Epsilon congratulate **Jan Cerny**, winner of the 2011 John Edgar Baublitz Pi Lambda Sigma Award.

The Economics Department and Omicron Delta Epsilon congratulate **Chidochase Munangagwe**, winner of the 2011 Financial Executives International Award.

The Economics Department and Omicron Delta Epsilon congratulate **Niki Shrestha** and **Taylor Smart**, winners of the 2011 Dr. and Mrs. William F. Railing Fellowship for Faculty-Student Research in Economics.

The Economics Department and Omicron Delta Epsilon congratulate **Chris Carrier, Jan Cerny, Tina Cheng, Cara Elliott, Long Le, Stephanie Miller-Hendry, Svetoslav Semov** and **Charles Zange** for their induction into Phi Beta Kappa. Phi Beta Kappa celebrates and advocates excellence in the liberal arts and sciences. Its campus chapters invite for induction the most outstanding arts and sciences students at America's leading colleges and universities.

The Economics Department and Omicron Delta Epsilon congratulate **Jan Cerny, Jacob Hochard, Svetoslav Semov**, and **Drew Shives** for receipt of a 2011 Mellon Grant.

The Economics Department and Omicron Delta Epsilon congratulate **Beth Adams** for receipt of the 2011 Glatfelter Award.

Omicron Delta Epsilon would also like to thank our outgoing officers, **Chris Carrier, Svetoslav Semov** and **Nevena Todorova**.



## CONTENTS

<i>Monetary and Fiscal Policies: Ordinary Recessions and Financial Crises</i> by Svetoslav Semov .....	pg 5
<i>Local Technological and Demographic Effects on Electricity Transmission: A Spatially Lagged Local Estimation of New England Marginal Losses</i> by Jacob Hochard.....	pg 41
<i>“Does It Pay to Be Informed?” Expenditure Efficiency in the US Mutual Fund Industry</i> by Jan Cerny .....	pg 54
<i>Friedrich von Hayek: The Socialist-Calculation Debate, Knowledge Arguments, And Modern Economic Developments</i> by Cara Elliott .....	pg 78
<i>History of Usury The Transition of Usury through Ancient Greece, the rise of Christianity and Islam, And the Expansion of Long-Distance Trade and Capitalism</i> by Cheryl Olechnowicz.....	pg 97



# **Monetary and Fiscal Policies: Ordinary Recessions and Financial Crises**

*Svetoslav Semov*

## Abstract

This paper uses two different samples to study the effects of monetary and fiscal policies on the profiles of recessions and recoveries. Several results emerge from the econometric analysis presented. First, monetary policy during ordinary recessions and banking crises is a powerful tool with lasting effects that extend to recovery growth rates. However, the effect of monetary policy during financial crises is strongly diminished in the case of forbearance – banks left to function despite being technically insolvent. Second, the effectiveness of fiscal policy is reversed – it is a powerful tool during banking crises, but it does not seem to significantly affect recovery growth rates during ordinary recessions. Finally, the policy response during past financial crisis does not seem to be particularly expansionary – on the contrary, fiscal policy is markedly procyclical, while monetary policy is neutral. This is proposed as an alternative explanation to the one usually given for the sluggishness of financial crises.

## **I. Introduction**

The Global Recession of 2008-09 sparked renewed interest in systemic financial crises. A key observation, first documented by Kaminsky and Reinhart, was that recessions associated with financial crises turn out to be particularly severe and protracted (1999). Most of the work on financial crises has concentrated on real-economy variables like output loss, length, depth etc. (Reinhart and Rogoff, 2009; Kannan, 2010; Claessens et al., 2004). The role of monetary and fiscal policies in financial crises has not been extensively studied with the exception of a 15-country study in the latest issue of the World Economic Outlook (IMF, 2009). It is possible that inappropriate monetary and fiscal policies are one reason why recoveries associated with financial crises turn out to be particularly severe. In addition, it might be that in those cases in which monetary policy was appropriately used, its effectiveness was diminished because the monetary policy transmission mechanism was impaired as a result of the stress in the banking system (IMF, 2009a).

This paper attempts to empirically evaluate the effect of monetary and fiscal policies in financial crises on the duration of the recession and the strength of the recovery. It further tries to examine if the effectiveness of monetary policy is dependent on the implementation of financial reforms. As a benchmark, recessions are employed to evaluate the normal impact of monetary and fiscal policies on recovery growth rates. Several results emerge from the econometric analysis presented. First, expansionary monetary policy during ordinary recessions is a powerful tool with lasting effects that extend to recovery growth rates. However, fiscal policy does not seem to affect post-crisis growth. Second, expansionary monetary policy during financial crises still has a positive but insignificant effect on the strength of the recovery, while expansionary fiscal policy has a positive and significant effect. Furthermore, these results are preserved in the analysis of the duration of the recession. Some empirical evidence is provided that explains the ineffectiveness of monetary policy by numerous cases of forbearance – banks left to function despite being technically insolvent. Finally, fiscal policy in financial crises seems to be markedly procyclical – the authorities cut government consumption on average by 2.5 percent of GDP during the duration of the downturn. Monetary policy, on the other hand, seems to be countercyclical – real money market rates are decreased on average by once percent during the duration of the downturn.

Two different samples are used. The first one uses quarterly data for a set of seventy crises in nineteen developed countries to analyze the effect of the policy response on the duration of the downturn and recovery growth rates. The second one employs yearly data for a set of eighty financial crises episodes in different countries to do the same. In addition, the impact of forbearance on monetary policy is also estimated. In both cases, the goal is to use the variation in policy responses and outcomes to find out the relationship between the variables of interest.

Eight sections follow. Section II reviews other cross country studies that examine the profiles of recessions and recoveries associated with financial crises. Section III presents a graphical interpretation of a linearized New Keynesian model with a risk premium. Within this framework, I explain the difference between financial crises and ordinary recessions. Furthermore, I illustrate the importance of monetary and fiscal policy. In addition, in Section III, I provide a concise analysis of the policy response in twelve financial crises and I argue that non-Keynesian policies are the norm rather than the exception. Section IV specifies the econometric model to be used. In addition, it discusses alternative versions of the model that should be estimated to check for the robustness of results. Section V describes the data, on which the analysis will be based and its sources. Section VI presents evidence on the effectiveness of monetary and fiscal policies in both ordinary recessions and financial crises and discusses the role of forbearance on monetary policy effectiveness. Finally, Section VII concludes and summarizes the results.

## **II. Literature Review**

This paper will attempt to add to the literature on cross-country studies of financial countries. Most studies examining recoveries and recessions associated with financial crises look at outcomes (output loss, duration of recession, sluggishness of recovery) without explicitly answering the question what it is that causes financial crises to be such protracted affairs. In addition, they do not include the policy response in the analysis.

For example, Reinhart and Rogoff conduct a comparative historical analysis of the aftermath of systemic financial crises (2008). The countries under consideration are both developed and emerging economies that have experienced financial distress in the after-war period. Reinhart and Rogoff's analysis shows

deep and lasting effects on output and employment. Unemployment rises for five years and output declines last on average for two years following the peak of economic growth. This is substantially more than the length observed during “normal recessions”. However, the authors do not provide any explanations for why this might be the case. Their analysis is merely comparative.

Boysen-Hogrefe et al. use a parametric framework to test whether the size of the bounce-back of GDP following an ordinary recession is larger than that following a recession associated with a banking crisis or housing crisis. The study covers 16 industrialized countries from 1970 to 2006. The results indicate that the output loss during an ordinary recession is completely offset in the following recovery. This is not the case when the recession was triggered by a banking crisis or a housing crisis. Again, this study does not offer explanations for why this might be the case – it simply makes this observation.

Kannan offers one possible reason why recoveries from banking crises might be more protracted (2010). Using a sample of 21 industrialized economies from 1970 to 2004, the author documents that it takes  $5 \frac{1}{2}$  quarters for output to recover following a banking crises, while it takes only 3 quarters following a normal recession. Evidence is presented that stressed credit conditions are an important factor containing the pace of the recovery. Industries that are more reliant on external finance, or more subject to financial frictions, are found not to recover as fast as other industries following all kinds of recession. The author finds strong evidence that the differential growth patterns across industries is much more pronounced in the aftermath of a financial crisis than it is for other recessions.

One potential drawback of this study is the small sample. The author relies on just 15 financial crisis episodes, not all of which are systemic.



There is another strand of literature that attempts to explain why some financial crises are so prolonged. This strand analyzes the effect of financial policies on the depth and duration of recessions. For example, Cecchetti et al. explore a vast array of financial policies (liquidity support, deposit freeze, blanket guarantee, bank holiday, forbearance etc.) and find that establishing an asset management company is associated with shorter recessions (2009). Furthermore, the authors find that forbearance is strongly associated with bigger output losses. Other financial policies do not seem to have a significant effect on length, depth and cumulative output losses during recessions associated with financial crises.

Also, Claessens et al. find that that excessive fiscal outlays delay economic recovery. The fiscal outlay figure includes both fiscal and quasi-fiscal outlays for financial system restructuring, including the recapitalization costs for banks, bailout costs related to the government covering obligations due to depositors and creditors, and debt relief schemes for bank borrowers. Furthermore, better institutional framework, as characterized by less corruption and greater judicial efficiency, does reduce output losses, even when controlling for excessive fiscal outlays.

In summary, the literature on financial policies might explain why some financial crises are so prolonged – if they were not followed by the implementation of the appropriate financial system reforms.

In addition, there is another reason financial crises might turn out to be more sluggish than ordinary recessions – if monetary and fiscal policies were not appropriately used. The effect of monetary and fiscal policies is explored in the most recent World Economic Outlook (IMF, 2009). The authors find that monetary and fiscal policies tend to shorten the duration of all types of recessions. Both increases in government consumption and decreases of interest rates beyond what is warranted by a Taylor rule positively and significantly affect recovery growth

rates. However, when only financial crises are analyzed the effect of monetary policy is found not to be statistically significant. One drawback of this study is that the sample for banking crises is limited to only fifteen episodes in developed countries. This study is also related to the literature on the monetary policy transmission mechanism. If the transmission mechanism is affected, then the way monetary policy works could also be influenced. For example, the interest-rate and the bank-lending channels could be hampered by the stress experienced by the financial system, something that might lead to reduced effectiveness of monetary policy (IMF, 2009a).

This paper attempts to add to the discussion of the sluggishness of financial crises. It will build on previous work on the effects of monetary and fiscal policies during banking crises (IMF, 2009). In particular, a larger sample than used before will be employed to test whether the strength of the recovery and the duration of the recession are affected by the policy response. In addition, the impact of monetary policy will be examined in cases of forbearance. If the lack of financial reforms proves to change the effectiveness of monetary policy, then this might give another explanation why some countries take so long to recover following a banking crisis. Finally, the extent to which fiscal and monetary policies have been used in past financial crises is documented.

### **III. Financial Crises and Past Policy Responses**

Various studies analyze the link between the financial sector and the real economy (Bernanke, 1983; Bernanke and Gertler, 2000; Kiyotaki and Moore, 1997). In this section, I review some of the existing literature that explains how the financial sector can amplify output shocks, making a recession deeper and more prolonged. Furthermore, I use a graphical version of a linearized New Keynesian model that incorporates a risk premium and demonstrates the

difference between financial crises and ordinary recessions. Finally, I propose an alternative explanation for the severity of financial crises – the policy response. I argue that financial crises are often a time of immense political and economic turmoil, something that often leads to the pursuit of non-Keynesian policies. In addition to providing some possible explanations for the contractionary policies countries have undertaken during financial crises, I review, in detail, the policy response in twelve systemic banking crises. The episodes discussed suggest that both developed and developing countries have pursued non-Keynesian policies in the past.

### ***Financial Crises: Why are They Different from Ordinary Recessions?***

Some evidence has been found for Milton Friedman’s “plucking model” which says that cyclical contractions tend to dissipate more quickly the larger the size of the contraction (Sinclair, 2005). However, financial crises do not seem to follow this pattern. They serve as an amplification mechanism that magnifies and accompanies other types of shocks like exchange rate, domestic and foreign debt crises (Reinhart and Rogoff, 2009a). An essential part of this amplification mechanism is the asymmetric information problems that arise during a financial crisis (Bernanke, 1983). Bernanke claims that the loss of confidence in financial institutions and the widespread insolvency of debtors lead to increased cost of credit intermediation, because banks cannot differentiate between good and bad borrowers. Consequently, potential worthy borrowers cannot undertake their projects; also savers have to devote their funds to inferior uses. As a result, there is a contraction in economic activity.

Bernanke and Gertler (2000) formulated a formal model that explains how the financial system serves as an amplification mechanism to negative shocks that hit the economy. The initial output shock leads to a decrease in wealth, which

makes firms more dependent on external financing. A weak banking system cannot provide that financing, leading to a decline in investment. Kiyotaki and Moore trace a similar dynamic in a richer intertemporal model (1997). A collapse in land prices undermines a firm's collateral, something that decreases its credit limit. This causes it to pull back investment in assets and hurts it even more in the next period.

The dynamics described above can be analyzed within an otherwise standard New Keynesian model that includes a risk premium. The model has the following equations (Clarida et al., 1999):

$$\text{AS: } \pi - E_t \pi_{t+1} = \alpha(Y_t - Y_t^n) + u_t$$

$$\text{IS: } Y_t - Y_t^n = -\gamma[i_t - E_t \pi_{t+1}] + E_t(Y_{t+1} - Y_{t+1}^n) + g_t$$

$$\text{TS: } r = f - E_t \pi_{t+1} + \sigma$$

This is a linearized version of a New Keynesian model (Clarida et al., 1999). The AS curve is derived from the Euler equation of firms. It is referred to as the New Keynesian Phillips curve. It shows a positive relationship between prices and output, because an increase in output leads to higher real marginal costs, which in turn make firms increase their prices. The parameters  $\pi$ ,  $\pi^e$ ,  $Y_t$ ,  $Y_t^n$  represent inflation, expected inflation, output and the natural level of output (the level that will arise if prices are perfectly flexible). The parameter  $\alpha$  refers to the fraction of sticky-price firms. The larger this fraction is, the flatter the AS curve, and correspondingly, the smaller change in price level economic fluctuations produce. The last term of the AS curve,  $u_t$ , is referred to as “cost push”, i.e. anything else that might affect marginal costs. In addition, it is a random disturbance term that follows an autoregressive pattern.

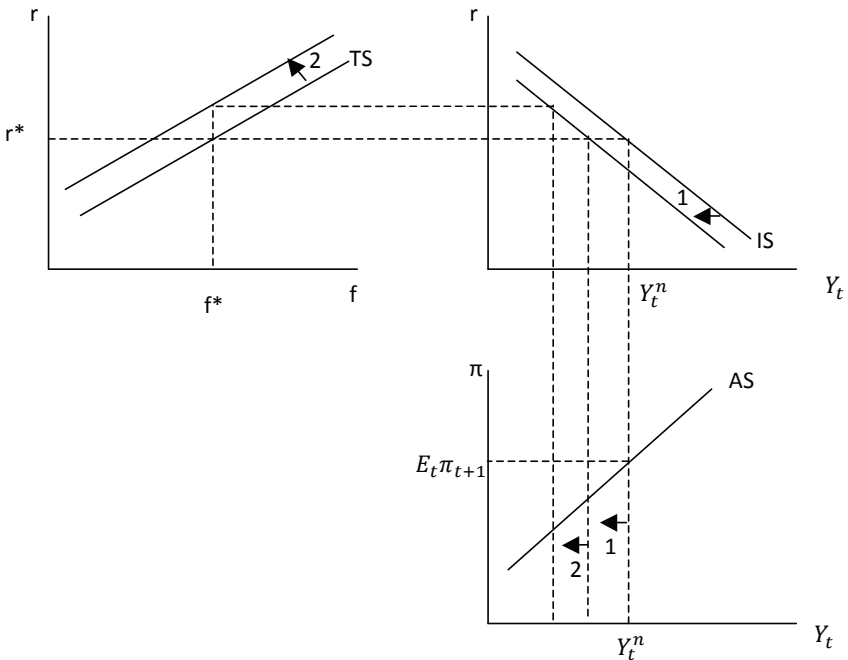
The IS curve is derived from the consumption Euler equations of households, that is the household's optimal saving decision. In this equation the current output gap depends on expected future output,  $E_t (Y_{t+1} - Y_{t+1}^n)$ , and the real interest rate  $-(i_t - E_t \pi_{t+1})$ . Higher expected future output raises the current output, because consumers want to smooth consumption, and, therefore, consume more today. In addition, the negative effect of the real interest rate reflects the intertemporal substitution of consumption. The last term of the IS curve,  $g_t$ , is a function of expected changes in government purchases relative to expected changes to potential output. Since  $g_t$  shifts the IS curve, it is interpretable as a demand shock (Clarida et al., 1999). Also,  $g_t$  is a random disturbance term that follows an autoregressive pattern.

Finally, the TS curve links the real risky rate,  $r$ , and the federal funds rate,  $f$ . The parameter  $\sigma$  is the risk premium. Although, the optimization of the monetary authority's loss function is not a part of the model, it implicitly enters the selection of the appropriate level of the federal funds rate  $f$ . The Fed's stabilizing policy rule makes it offset shocks to the risk premium or to expected inflation.

Recessions associated with financial crises can be analyzed within this model. More importantly, the difference between those recessions and "ordinary" recessions can be illustrated. In the model normal recessions are usually caused by a leftward shift in the IS curve – a demand shock. For example, the demand shock in the financial crisis of 2008 was the collapse of the housing market that caused residential investment and consumption to fall. During times of financial distress there is an additional factor at play – the risk/liquidity premium  $\sigma$ . A jump in its value shifts the TS curve up, raising real interest rates on corporate bonds, mortgages, and other risky assets. This is consistent with Bernanke's claim that higher cost of credit intermediation leads to increased interest rates

or to a curtailment of credit (1983). In the model, the increased interest rates are represented by the risk premium. The shift of the TS curve is also consistent with the lowering of borrowers' credit limits in Kiyotaki's model, something that also leads to higher interest rates (Kiyotaki and Moore, 1997).

For example, at the start of the financial crisis of 2008 there was an uncertainty associated with the solvency of various financial institutions. Also, there was a huge fire sale of risky assets in an effort to raise cash. Such events cause the TS curve to go up (the movement of the curve could be observed in the equations above – as  $\sigma$  increases,  $r$  rises as well). An upward shift in the TS curve leads in turn to a decrease in investment and consumption, causing output to fall even further (illustrated by an upward movement along the IS curve). The graphs below illustrate these dynamics:



In step (1) the economy is undergoing a demand shock often responsible for ordinary recessions. In cases of financial distress, there is an additional force, illustrated in step (2), which is exacerbating the recession.

This model can be further used to illustrate how the policy response can add to the severity of a financial crisis. Expansionary monetary policy is represented by downward movements along the TS curve (the Fed optimizes its loss function, choosing the appropriate level of  $f$ ), which lead to downward movements along the IS curve and correspondingly to higher output. Fiscal policy acts through the IS curve – an increase in government spending shifts the IS curve to the right, leading to an increase in output. Unconventional policies, like measures to calm down financial markets, go through the TS curve. For example, stress tests of the banking system lead to a decrease in  $\sigma$ , the risk premium, and a downward shift of the TS curve. Also, quantitative easing can target the term premium and also shift down the TS curve.

The model specified above is useful for distinguishing between financial crises and ordinary recessions. Furthermore, it illustrates the possible impact of monetary and fiscal policies. However, it does not differentiate between the effectiveness of these policies in different environments. For example, Gali (2005) and Eggertsson and Krugman (2010) demonstrate within a New Keynesian model with heterogeneous agents that during times of financial crises the number of credit constrained agents increases. As a result, government spending is effective in raising the disposal income of those agents, something that makes them spend more. In other words, these studies imply that fiscal policy might be more effective during times of financial crises.

In addition, monetary policy might also have different effectiveness in various environments. For example, if the transmission mechanism is affected

during a financial crisis, then the way monetary policy works might change. The interest-rate and the bank-lending channels could be damaged by the stress experienced in the financial system. (IMF, 2009a). Furthermore, if the economy is in a liquidity trap, as during Japan's 1997 recession, then traditional monetary policy instruments are also not as effective as they would be under normal circumstances.

### ***Why is the Recovery Slow?***

The recovery from a financial crisis is slower than that from an ordinary recession for similar reasons. As Bernanke argues, it takes time to establish new or revive old channels of credit (1983). Furthermore, it takes time to rehabilitate borrowers. This last idea is further developed by Koo (2009). He argues that financial crises are usually connected to “balance sheet” recessions. The last can result from a shock to balance sheets – for example, a bubble burst – that often accompanies financial crises. Then, it takes time for households and businesses to repair their balance sheets. For example, Japan's recovery during the “lost decade” was prolonged as a result of an overhang of corporate debt. Similarly, an overhang of household debt is probably holding down U.S. economic growth right now.

### ***Why has the Policy Response in Past Financial Crises not been Keynesian?***

The divergent policy responses to financial crises have their basis in the fundamental theoretical disagreement about the effects of stabilization policies that exists in the economic profession. Starting in the 1970s there was a shift in economic thinking led in part by Edward Prescott that resulted in the formation of New Classical economics. A main part of this shift was the idea that activist policies to fight the business cycle are undesirable. This was because recessions result from the rational decision of workers to work less when the economic conditions are less favorable and, therefore, are the natural course of events. However, there



were still economists who believed that recessions are caused by demand side of the economy – the New Keynesians. They worked to incorporate enough frictions into the Real Business Cycle models of New Classical economists so that they can bring the two camps closer together.

As Krugman argues, during the period 1980 – 2007 the clash between the New Keynesians and New Classical economists was mainly on the basis of theory and not action, because in the U.S. there was not much need to implement expansionary policies, since recessions were relatively mild over that period. New Keynesians thought that monetary policy was sufficient in managing the business cycle. In contrast, New Classical economists thought that both expansionary fiscal and monetary policy are ineffective, but did not mind the use of monetary policy.

A case in point of why disagreements in the economics profession matter for policy is the recent global financial crisis. Farrell (2011) argues that there were noticeable shifts in the policy debate and implementation in the U.S. starting in early 2010 that are attributable to the sovereign debt crises of Iceland and the Baltic states. In particular, these crises provided conservative policy makers the rhetorical fodder in the debate for more stringent fiscal policy. The intellectual support those policy makers needed was, in turn, sought from economists; and the disunited profession had what to offer. At the time various prominent economists put forward arguments against further extending the stimulus. Examples of such arguments are the work of Alesina and Ardagna (2010) supporting expansionary austerity and the work of Rogoff and Reinhart (2009) on admissible government debt thresholds.

In addition to the theoretical divide responsible for different approaches crisis countries have undertaken in the past, there are some attractive beliefs among policy makers that make them pursue non-Keynesian policies in the face

of financial and economic turmoil. For example, such a belief is the seemingly logical argument that problems of excessive debt, too much private borrowing, cannot be solved by creating even more of it – government borrowing (Krugman and Eggertsson, 2010). During the Asian Financial Crisis the IMF advised some of the crisis-stricken countries to pursue contractionary fiscal policies following a similar argument. The intention was to “restore confidence” by convincing the markets that irresponsible behavior is a thing of the past. For similar reasons, money market rates were increased and unnecessary structural reforms were undertaken (Krugman, 2010). Tightened monetary policy was aimed at convincing the markets that the pegged exchange system will be preserved. Some of the structural reforms had no particular connection to the crisis but they were also aimed at calming down the markets. A typical example of IMF-advised policies was what Korea did in 1997. Money market rates were raised up to 25.6 percent in M1, 1998. Furthermore, there was an initial tightening of fiscal policy to rebuild confidence (for half a year).

Furthermore, there might be institutional reasons for some countries’ inability to pursue Keynesian policies – Kaminsky et al. argue that developing countries face credit constraints during bad times that prevent them from borrowing (2004). Furthermore, developing countries tend to also follow procyclical policies during good times, meaning that they do not have the necessary cushion to fight recessions.

Advanced countries are not immune to institutional problems. Central banks in some developed countries have become increasingly conservative in the past two decades, focusing too much on inflation, and this might have its consequences during severe recessions (Krugman, 2010). A recent IMF study of 25 severe recessions in advanced economies finds that prolonged periods of economic weakness are associated with falling inflation rates (Meier, 2010). However, it also

finds that as the inflation rate goes toward zero, it becomes sticky. This means that a severely depressed economy can still have a positive inflation rate – most likely because of downward nominal rigidities and well-anchored inflation expectations. A central bank that is overly focused on inflation might miss the urgency of the situation and not act as aggressive as necessary (Krugman, 2010).

Finally, there is an additional reason why some countries cannot further stimulate a depressed economy – the liquidity trap. Such an environment was observed in Japan in the 1990s and is currently the reality in U.S.

#### IV. Modeling

The effect of monetary policy during recessions on the ensuing recoveries is first analyzed on the background of countries, experiencing “ordinary” recessions. This is meant to serve as a benchmark. Then, the effect of monetary policy is analyzed in countries undergoing banking crises.

##### *Monetary Policy in Ordinary Recessions*

The goal is to see if recovery growth rates after ordinary recessions are significantly affected by monetary policies. For that purpose a fixed effects model is used (IMF, 2009a). The reason for this is to capture the effect of any unobservable country-level characteristics that pertain to the recessions and recoveries experienced. For example, a country with an export-oriented economy might be able to faster drag itself out of a recession. Such an occurrence would be captured by the fixed-effects model assuming that throughout the period under consideration the export industry has held a similar role. In particular, the model estimated is:

$$RecGrowth_{i,t} = c_0 + c_1 * Amplitude_{i,t} + c_2 * Duration_{i,t} + c_3 * MP_{i,t} + c_4 * CAGC_{i,t} + e_{i,t} \quad (1)$$

The dependent variable is the recovery growth rate one year after the trough of the recession. I control for the amplitude and duration of the recession – these are characteristics of the business cycle itself that might differ within a certain country over time. For example, there is nothing to make us believe that external shocks which hit an economy should be of the same size. Milton Friedman’s “plucking model”, which has been empirically verified, suggests that the coefficient estimate on *Amplitude*,  $c_1$ , should be positive – the deeper the recession, the stronger the recovery. Furthermore, I expect that prolonged recessions have slower recoveries. As a result, the coefficient estimate on *Duration*,  $c_2$ , is conjectured to be negative. Duration is measured in quarters.

The monetary policy response over the recession period, the variable *MP*, is measured as the sum of the residuals of a monetary policy rule over each quarter over the recession period. I expect that countries that increased interest rates above what is warranted by a Taylor rule experienced slower recoveries. As a result, the coefficient estimate on *MP*,  $c_3$ , is conjectured to be negative.

As mentioned, monetary policy shocks are identified from the residuals of a monetary policy rule. For that purpose, following the methodology of the latest issue of the *World Economic Outlook*, a Taylor rule of the following form was estimated for each country:

$$i_t = b_0 + b_1 * dummy\_85 + b_2 * \pi_t + b_3 * gap_t + u_t$$

where  $i_t$  is the nominal interest rate, *dummy\_85* is a dummy for periods after 1985 (to allow for a shift in equilibrium rates),  $\pi_t$  is the inflation rate and *gap<sub>t</sub>* is a measure of the output gap (potential GDP is measured using the Hodrick-Prescott filter). Using the real interest rate as the dependent variable does not change the results, since inflation is included in the model.

The fiscal policy response over the recession period, the variable  $CAGC$ , is cyclically adjusted government consumption. I expect that an increase in this variable will lead to a higher growth rate of the recovery. Therefore,  $c_4$  should be positive. Again, the methodology of the *World Economic Outlook* is used. First, the elasticity of government consumption with respect to the business cycle is estimated:

$$\ln gc_t = d_0 + d_1 * gap_t + d_2 * trend + e_t$$

As above,  $gap_t$  is a measure of the output gap.  $Trend$  is a time trend. Second, the cyclically adjusted government consumption is computed as:

$$CAGC_t = gc_t (1 - d_1 * gap_t)$$

In addition to the above estimations, a check for the robustness of the results is performed. The duration of the recession is used as dependent variables in some of the estimations to see if the effects of the monetary and fiscal policies change.

### *Monetary Policy in Financial Crises*

The effects of monetary and fiscal policies on the duration of the recession and the strength of the recovery following financial crises are analyzed. In addition, monetary policy is examined in cases of forbearance - banks left to function despite being technically insolvent, and prudential regulations (such as for loan classification and loan loss provisioning) suspended or not fully applied. If forbearance has a negative effect on the effectiveness of monetary policy, then insufficient use and diminished effectiveness of an otherwise powerful tool for stimulating recoveries might be responsible for the sluggishness of some financial crisis episodes. Data for forbearance is available only for about thirty five countries (Laeven and Valencia, 2010), while the sample of all financial crises includes about eighty countries. Consequently, two different estimations are performed. The model that includes forbearance is:

$$RecGrowth = c_0 + c_1 * Amplitude + c_2 * Duration + c_3 * MP + c_4 * MPForb + c_5 * Forbearance + c_6 * GDP(-1) + e \quad (2)$$

The variables *RecGrowth*, *Amplitude*, and *Duration* are the self-explanatory. They are measured as the recovery growth rate one year after the trough of the recession, the sum of GDP growth rates during the recession, and the duration of the recession in quarters.

Monetary policy is measured as the change in real money market rates over the course of the recession. If money market rates are not available, then their closest substitute is used. The reason for the difference from before is the usage of yearly data. Estimating residuals from a Taylor rule would be too imprecise with yearly data. A decrease in interest rates would mean that there is a negative change in real money market rates. Therefore, we are testing if  $c_3$ , the coefficient estimate on the monetary policy measure, is negative. Note that the dependent variable is the growth rate, or the output gap, in the recovery phase, which is at least one year after the implementation of monetary policies; this would eliminate any endogeneity problems.

In addition to the measure for monetary policy, the regression equation includes an interaction term between the changes in real interest rates and forbearance. Forbearance is a dummy variable that indicates whether or not there is regulatory forbearance during the years  $[t, t+3]$ , where  $t$  denotes the starting year of the crisis. This variable is based on a qualitative assessment of information contained in IMF Staff Reports (Laeen and Valencia). As part of this assessment, information is collected on whether or not banks were permitted to continue functioning despite being technically insolvent, and whether or not prudential regulations (such as for loan classification and loan loss provisioning) were suspended or not fully applied during the first three years of the crisis.

The interaction term is supposed to estimate whether in cases of forbearance the effect of monetary policy is reduced. Previous experience has

suggested that this might be the case. For example, many observers consider the policies undertaken by Sweden in the early 1990s to have been highly effective in restoring the health of the financial system and paving the way for a strong recovery through extensive use of expansionary policies (IMF, 2009). In contrast, in Japan, slow recognition of the bad-loan problem contributed to a sluggish recovery from the financial crisis, even though interest rates were at the zero bound. The effectiveness of monetary policy might be disrupted if the interest rate and credit channels of the monetary policy transmission mechanism are not properly working during a banking crisis.

In other words, the coefficient estimate on  $MPForb$  is expected to be positive. The marginal impact of  $MP$  is given by  $c_3+c_4*Forb$ . We expect that  $c_3$  would be negative. Therefore, if forbearance diminishes the effect of monetary policy it should be making the whole term bigger (“less negative”). This would mean that we are testing whether the coefficient estimate on  $MPForb, c_p$ , is positive.

Finally, forbearance is also included in the model. We would expect that if the authorities do not address and act on failing banking institutions, then this would have a direct negative effect on the economy. However, it is not particularly clear how long lasting this deleterious impact might be. Generally, we would expect that recovery growth rates might be negatively affected by forbearance. Therefore, we are testing to see if the coefficient estimate of *Forbearance*,  $c_5$ , is negative.

The estimated model without forbearance would look like:

$$RecGrowth=c_0+c_1*Amplitude+c_2*Duration+c_3*MP+c_4*GC+c_5*GDP(-1)+e \quad (3)$$

The definition of the variables is the same as above. The only difference is that the government consumption variable is added. In particular, fiscal policy is proxied by the percentage change in government consumption. This measure is used instead of the fiscal balance, because the last would cause endogeneity. A

change in output affects the fiscal balance (it is a fraction of output) and a change in the fiscal balance affects output. An increase in government consumption during the recession phase is expected to positively affect recovery growth rates. Therefore, we are testing if  $c_{4p}$ , the coefficient estimate on the change in government consumption, is positive.

In addition to the above estimations, a couple of robustness checks are performed. The duration of the recession and the output gap one year after the end of the recession are used as dependent variables.

## **Data**

In getting a better understanding of the recovery that will follow the 2008-09 recession through the lenses of historical experience we have two choices. We can either draw conclusions from the financial crises that occurred during the 1930s, or look at the ones that have plagued the world in the past forty years. The reason for this is the striking pattern of occurrence of financial crises worldwide. From the 1940s up to the early 1970s, there were almost no banking crises in the world.<sup>1</sup> However, with the financial and international capital account liberalization of the 1970s, banking crises have re-emerged (Reinhart and Rogoff, 2009a).

This paper focuses on the period 1970-2005 and it uses two distinct data sets. The first one consists of data on recessions and recoveries in a set of advanced countries. The countries are those identified in the Statistical Appendix of the 2010 issue of the World Economic Outlook as advanced. Then, subject to data availability the monetary and fiscal policy responses during all of the recessions since 1970 in the selected countries are analyzed. Quarterly data is used. To

---

<sup>1</sup> Reinhart and Rogoff (2008) argue that this calm might be partly explained by booming world growth, but perhaps more so by the repression of the domestic financial markets (in varying degrees) and the heavy-handed use of capital controls followed for many years after WWII.



measure the stance of monetary and fiscal policies money market rates and government consumption are employed. The main source of the data is the International Financial Statistics database of the IMF. Given that only a few of the countries have data going back before 1977 this limits the sample.

The procedure for identifying business cycles is an algorithm called BBQ (Bry and Boschan procedure for quarterly data; see Harding and Pagan, 2002). A MATLAB version of a program that imitates the algorithm can be found at [www.ner.edu.au](http://www.ner.edu.au). It uses quarterly output data to identify peaks and troughs. A complete cycle goes from one peak to the next peak with its two phases the contraction phase (from peak to trough) and the expansion phase (from trough to peak). The algorithm requires that the minimum duration of the complete cycle and each phase must be at least five and two quarters, respectively. Table1 in the appendix shows the recessions (peak-to-trough) identified by this algorithm.

The second dataset consists of eighty financial crisis episodes in both developed and developing countries. Laeven and Valencia identify 124 systemic banking crises between 1970 and 2007 (2008). Data on real GDP, inflation, government consumption and interest rates is collected from the International Financial Statistics database of the IMF. Eighty of the 124 crisis episodes had output data available. Furthermore, of those eighty countries not all have both government consumption and interest rates data available. As a result, the sample is limited to less than eighty countries in the various regressions below. To measure the stance of monetary and fiscal policy money market rates and government consumption are used. Wherever money market rates are unavailable, their closest substitute is used. Data on forbearance is taken from Laeven and Valencia (2010) and it is limited to about 35 countries for which the authors provide information on various financial policies undertaken. All data is yearly.

Figures 1 and 2, in the appendix, describe the output dynamics and duration of recessions associated with financial crises. Those recessions are particularly severe – the amplitude of the recession is on average about four percent of real GDP and the mean duration is about five quarters. In addition, seventy percent of the crisis periods considered have a duration of one year or more. The policy response in those crises is illustrated in Figures 3 and 4. Monetary policy, proxied by the change in money market rates, seems to be expansionary – real interest rates have declined during both the first year and the whole duration of the crisis. However, these declines can be mainly explained by the inflationary dynamics in the countries and not by the explicit behavior of policy makers. Figure 3 shows that nominal interest rates have actually increased slightly. On the contrary, fiscal policy, proxied by the change in government consumption, seems to be markedly procyclical. Approximately one half of the crisis episodes were characterized by a negative change in government consumption during the duration of the recession. Note that government consumption data is available for 78 out of the 80 countries under consideration. That number for interest rates is 70.

The start of the financial crises themselves is taken from Laeven and Valencia (2008). The peaks of the recessions are identified using a one-year window around the start of the financial crisis. In this way, it is ensured that the recessions under consideration are, in fact, associated with the financial crises. Note, however, that in some of the crisis periods there was no output loss – in those cases, following Cecchetti et al. (2010) the duration and the amplitude of the recession are set equal to zero. Table 2, in the appendix, shows the start of each of the banking crises under consideration (Laeven and Valencia, 2008).

## V. Empirical Evidence

The effects of monetary and fiscal policies during recessions on the ensuing

recoveries are first analyzed in the sample of advanced countries, experiencing “ordinary” recessions. Then, the effects of monetary and fiscal policies are analyzed in the sample of countries undergoing banking crises.

*Monetary and Fiscal Policies in Ordinary Recessions*

The table below shows the regression results from estimating the fixed effects model specified above – regressions (3) and (4). In addition to this model, one is estimated with a dependant variable the duration of the recession – regressions (1) and (2). The results of the two sets of models are largely consistent. The same set of variables is statistically significant in both of them. Also, the estimations without fixed effects in both cases have a much smaller explanatory power than the ones with fixed effects.

Table 3 – Results for the severity and sluggishness of **ordinary recessions**.

	(1)	(2)	(3)	(4)
Duration			-0.024 (-.13)	-0.077 (-0.46)
Amplitude	-.109 (-0.91)	-.133 (-3.00)***	-.488*** (-3.19)	-.130 (-1.47)
RealRate	5.39 (2.60)**	4.93 (2.72)***	-4.207 (-2.43)***	-1.005 (-.62)
GC	.134 (1.50)	.006 (0.22)	.120 (.373)	.008 (.18)
Fixed Effects	Yes	No	Yes	No
Observations	66	66	74	66
R-squared	0.41	0.30	0.40	0.04

Notes: unbalanced panel with country fixed effects. t-statistics are in brackets. \*, \*\*, \*\*\* denote level of significance indicating 10%, 5% and 1% respectively. Robust standard errors used. The dependent variable in (1) and (2) is the duration of the recession. The dependent variable in (3) and (4) is the recovery growth rate one year after the trough of the recession.

Looking at regression numbered (3) we see that the recession amplitude has a statistically significant effect on the growth rate *in the recovery phase*. The coefficient estimate is statistically significant in difference from zero at the 1 percent level of significance. Note that amplitude measures the percentage decline in GDP during the recession phase – peak to trough. This result suggests that

the steeper the recession, the faster the recovery. Furthermore, as expected the coefficient estimate on the RealRate is also statistically significant in difference from zero – at the five percent level of significance. As previously noted the RealRate is the sum of the impulses relative to the policy rule for each quarter over the recession period. In other words, RealRate represents the sum of the residuals from an estimated monetary policy rule. This would mean that an increase in the RealRate corresponds to an increase in money market rates above what a policy rule warrants. Therefore, as expected the coefficient estimate is negative. However, government consumption does not significantly affect recovery growth rates. One reason for this occurrence might be that the estimation does not account for the level of government debt, something found to be important for the effectiveness of fiscal policy (IMF, 2009). Furthermore, as Krugman and Eggertsson (2010) argue, the effect of fiscal policy is the biggest when there are credit constrained agents in the economy – as during a financial crisis.

The coefficient estimates agree in magnitude with those estimates in previous studies (IMF, 2009). In addition, the R-squared of the fixed effects model is pretty high, 40 percent, given that the dependent variable is growth rates one year after the recession has occurred. However, the R-squared of the model without the fixed effects is rather low – less than 4 percent of the variation of the dependent variable is explained by the independent variables included.

The estimation that has the duration of the recession as the dependent variable (regression equation (1) also suggests that monetary policy significantly affects the length of the recession.

### *Monetary Policy in Financial Crises*

The table below shows the regression results from estimating the model for financial crises. The dependent variables in (1), (2) and (3) are the duration of

the recession, the recovery growth rate and the output gap one year after the trough of the recession. The results of these three models are consistent with each other. In models (1), (2) and (3) increases in government consumption have a significant negative effect on the duration of the recession and a significant positive effect on recovery growth rates. The growth rate of GDP prior to the recession also seems to matter (regression equations (1) and (2)). Countries with higher prior growth rates tend to have stronger recoveries and shorter recessions.

Table 2 – Results for severity and sluggishness of **banking crises**.

	(1)	(2)	(3)
Duration		-0.058 (-.62)	-.0003 (-4.82)***
Amplitude		-.066 (0.33)	
Real GDP (-1)	-.428 (-1.77)*	.232 (1.86)*	-.0003 (0.86)
RealRate	0.011 (0.32)	.011 (0.50)	-.0003 (-.79)
Cum.Gov.Con.	-.125 (-3.74)***	.066 (2.06)**	.001 (3.23)***
Observations	66	65	66
R-squared	0.17	0.14	0.40

Notes: unbalanced panel with country fixed effects. t-statistics are in brackets. \*, \*\*, \*\*\* denote level of significance indicating 10%, 5% and 1% respectively. Robust standard errors used. The dependent variables in (1), (2) and (3) are the duration of the recession, the recovery growth rate and the output gap one year after the trough of the recession.

However, monetary policy does not significantly affect recovery growth rates, output gaps and duration of the recession. This might be due to the reduced effectiveness of monetary policy transmission mechanism during times of financial distress (IMF, 2009). The same results hold whether real or nominal rates are used. Furthermore, it does not make a difference if the cumulative change in interest rates over the whole duration of the recession is used or the change in the first year of the crisis. There have been reversals of policy, especially in the

crises that involve currency upheavals like the ones in the Asian Financial crisis so such a check makes sense.

As explained before, the effectiveness of monetary policy might be affected by the extent to which financial reforms were implemented in the affected countries. To formally test this hypothesis, a smaller sample of countries is used, for which data on forbearance is available. Forbearance is the qualitative assessment of whether banks were permitted to continue functioning despite being technically insolvent. The regressions below try to assess the impact of forbearance. Again, three dependent variables are used – the duration of the recession, the recovery growth rate and the output gap one year after the trough of the recession.

Table 3 – Results for the effect of monetary policy in the case of forbearance during **banking crises.**

	(1)	(2)	(3)
Duration		-0.052 (-0.40)	-0.004 (-2.98)***
Amplitude		-0.018 (-0.18)	
Real GDP (-1)	-0.060 (-0.27)	.006 (0.964)	-0.0008 (-0.37)
RealRate	-0.034 (0.59)	.004 (0.92)	-0.0026*** (-4.34)
RealRate*Forb.	.06 (0.92)	.092 (1.30)	.020 (2.70)**
Forb.	-1.35 (2.07)**	2.42 (-1.16)	-0.020 (-0.55)
Observations	30	30	30
R-squared	.03	0.27	0.47

Notes: t-statistics are in brackets. \*, \*\*, \*\*\* denote level of significance indicating 10%, 5% and 1% respectively. The dependent variables in (1), (2) and (3) are the duration of the recession, the recovery growth rate and the output gap one year after the trough of the recession. Robust standard errors used.

The coefficient estimates in the model that has recovery growth rates as a dependent variable are statistically insignificant in difference from zero. However, in the estimation using the output gap as the dependent variable, the change in

money market rates is statistically significant. An increase in money market rates leads to a decrease in the output gap. Furthermore, forbearance dampens the effect of monetary policy as indicated by the negative coefficient estimate on *RealRate\*Forb*. Also, this diminishing effect seems to be quite significant as it is bigger in magnitude than the positive effect of monetary policy on recovery growth rates. The lack of explanatory power of the independent variables in the model with recovery growth rates as the regressand agrees with previous studies which find that recovery growth rates are harder to predict than output gaps (IMF, 2009). Furthermore, we can see that the estimation with duration as the dependent variable also lacks statistical significance. This might be explained with the fact that forbearance is defined over the three years following the beginning of the recession. As a result, its effect might not be felt during the duration of the recession.

In summary, the empirical results suggest differences between the effects of monetary and fiscal policies on the duration of recessions and the strength of recoveries in ordinary recessions and in systemic financial crises. During ordinary recessions expansionary monetary policy seems to be a powerful tool, generating significant increases in recovery growth rates. During recessions associated with financial crises, expansionary monetary policy still has a significant effect on the strength of the recovery. However, this effect is dependent on the implementation of prompt financial policies, and in particular, on intervention with insolvent financial institutions. The effectiveness of fiscal policy is reversed – it is a powerful tool during banking crises, but it does not seem to significantly affect recovery growth rates during ordinary recessions.

## **VI. Conclusions**

Non-Keynesian policies in the face of a financial crisis are not a thing of the past. A number of advanced economies have pursued contractionary policies in the most recent financial crisis. This has certainly been the case in Europe. Many countries there embraced austerity in the face of a slumping economy – France, Britain and Ireland, for example. In addition, some EU members had to settle with insufficiently expansionary monetary policies, because of the outsized influence of Germany over the European Central bank and the better performance of the German economy. While policies in the U.S. have been more favorable towards sustaining a recovery, this has not come without much debate. Ideas and arguments supporting fiscal retrenchment have abounded. This is exemplified in the work of some prominent economists like that of Alesina and Ardagna (2009). The political climate has also been antagonistic towards some of the actions policy makers have tried to undertake. For example, there was a huge backlash against the quantitative easing program the Fed started to implement in late 2010 – something that can have a particularly deleterious effect when the economy is in the midst of a liquidity trap and when the Fed’s credibility in influencing the public’s expectations is the main tool out (Mankiw and Weinzierl, 2011).

In this paper, I use two different samples with data from 1970 to 2005 to study the effects of monetary and fiscal policies on the profiles of recessions and recoveries. In other words, I ask whether pursuing non-Keynesian policies has mattered during past financial crises. Several results emerged from the econometric analysis. First, monetary policy during ordinary recessions and banking crises is a powerful tool with lasting effects that extend to recovery growth rates. However, the effect of monetary policy during financial crises is strongly diminished in the case of forbearance – banks left to function despite being technically insolvent. Second, the effectiveness of fiscal policy is reversed – it is a powerful tool during



banking crises, but it does not seem to significantly affect recovery growth rates during ordinary recessions. Finally, the policy response during past financial crisis does not seem to be particularly expansionary – on the contrary, fiscal policy is markedly procyclical, while monetary policy is neutral. In summary, the results suggest it is possible that inappropriate fiscal and monetary policies and the lack of financial reforms could be one reason why recoveries associated with financial crises turn out to be particularly protracted.

Reinhart and Rogoff (2009b) argue that both advanced and developing countries suffer from the “this time is different” syndrome when it comes to financial crises, because policy makers and the public tend to believe that they are immune from a crisis due to some circumstances that make them special.<sup>2</sup> This paper suggests that this syndrome should be avoided when it comes to policy as well. This time is not different and expansionary policies should be pursued. Past financial crises have been so protracted in part because of the embracement of austerity and in part because of the lack of realization that financial crises are inherently more severe.

---

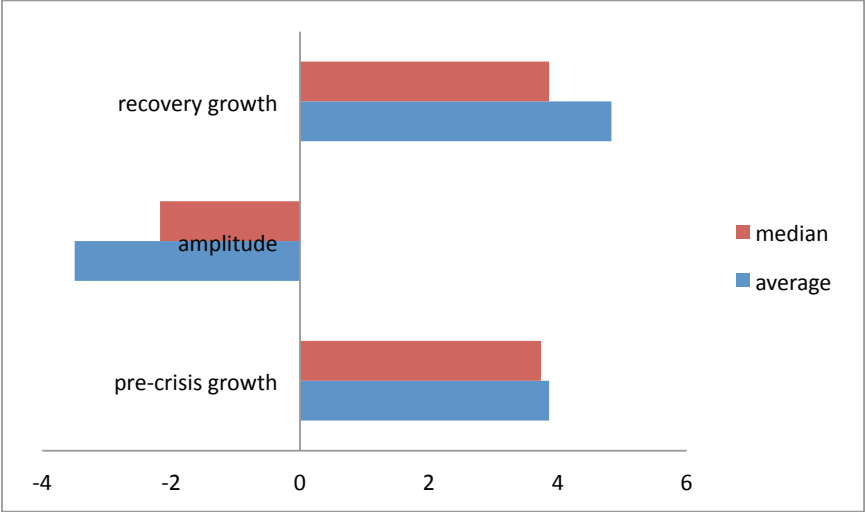
2 An example of that line of thinking involves the securitization process of mortgage backed securities in the U.S. prior to the most recent recession. People thought that these new “synthetic” products have managed to practically eliminate all risk from the economy. With the benefit of hindsight, we know that this was not the case.

## **References**

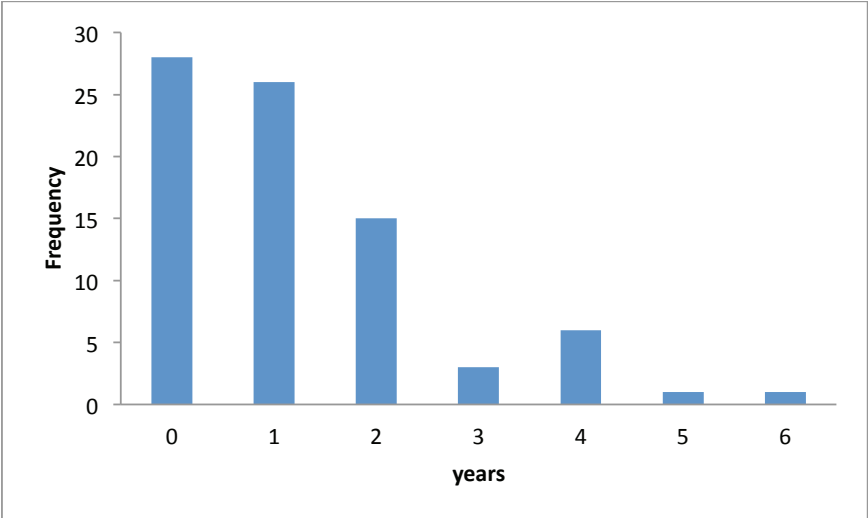
- Alesina, A. and Ardagna, S. (2009). "Large Changes in Fiscal Policy: Taxes versus Spending", *NBER Working Paper*
- Bernanke, Ben. 1983. "Non-Monetary Effects of the Financial Crisis in the Propagation of The Great Depression", *NBER Working Papers 1504*.
- Bernanke, Ben, and Mark Gertler. 2000. "Financial Fragility and Economic Performance", *NBER Working Paper*, No 2318.
- Boysen-Hogrefe Jens, Nils Jannsen, and Carsten-Patrick Meier. 2010. "The Ugly and the Bad: Banking and Housing Crises Strangle Output Permanently, Ordinary Recessions Do Not", *Kiel Working Paper*, 1586
- Calvo, Guillermo, Alejandro Izquierdo, and Ernesto Talvi. "Sudden Stops, the Real Exchange Rate, and Fiscal Sustainability: Argentina's Lessons", *NBER Working Paper 9828*.
- Cecchetti Stephen, Marion Kohler and Christian Upper. 2009. "Financial Crises and Economic Activity." *Financial Stability and Macroeconomic Policy Symposium*, Federal Reserve Bank of Kansas City.
- Cerra Valerie and Sweta Saxena. 2005. "Growth Dynamics: The Myth of Economic Recovery", *IMF Working Paper*, 147.
- Claessens, Stijn, Daniela Klingebiel, and Luc Laeven. 2004. "Resolving Systemic Financial Crises: Policies and institutions", *World Bank Policy Research Working Paper*, 3377.
- Claessens, Stijn, Daniela Klingebiel, and Luc Laeven. 2001. "Financial Restructuring in Banking and Corporate Sector Crises: What policies to Pursue", *NBER Working Paper*, 8386.
- Clarida, R., Jordi Gali, and Mark Gertler. 1999. "The Science of Monetary Policy: A New Keynesian Perspective". *Journal of Economic Literature*, Vol. XXXVII, pp. 1661-1707.
- Dominguez, Kathryn, and Linda Tesar. 2005. "International Borrowing and Macroeconomic Performance in Argentina", *NBER Working Paper 11353*.
- Eggertsson, Gauti and Paul Krugman. 2010. "Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo approach", *preliminary draft*.
- Englund, Peter. 1999. "The Swedish Banking Crisis: Roots and Consequences", *Oxford Review of Economic Policy*, Vol. 15, No. 3.
- Farrell, Henry. 2011. "Networks and Ideational Power: The Rise and Decline of Keynesianism during the Economic Crisis", *George Washington University paper*.
- Gallego, Sonsoles, Alicia Herrero, and Jesus Saurina. 2002. "The Asian and European Banking Systems: The Case of Spain in the Quest for Development and Stability", *KOBE Research Project*.
- Gallo, Andres. 2006. "The Role of Political Institutions in the Resolution of Economic Crises: The Case of Argentina 2001-05", *Oxford Development Studies*, Volume 34, Number 2, June 2006.
- Haug, D. P. Ollivaud and D. Turner 2009. "The Macroeconomic Consequences of Banking Crises in OECD Countries." *OECD Economics Department Working Paper No.683*, Paris.
- Honkapohja, Seppo, and Erkki Koskela. 1999. "The Economic Crisis of the 1990s in Finland", *Economic Policy*, Oct.
- Jimenez, Angel, and Jose Alegre. 1996. "Analysis of Fiscal Policy in Spain: A macroeconomic Perspective". Banco de Espana.

- Jonung, Lars, Jaakko Kiander, and Penttievartia. 2008. "The Great Financial Crisis in Finland and Sweden", European Commission Economic Papers.
- Joung, Kiander and Vartia, ed. 2008. "The Crisis of the 1990s in Finland and Sweden", Edward Elgar.
- IMF (International Monetary Fund). 2009. "From Recession to Recovery: How Soon and How Strong", *World Economic Outlook*. Washington D.C.
- IMF. 1998. "Argentina: Recent Economic Developments".
- IMF. 2003. "Lessons from the Crisis in Argentina", Prepared by the Policy Development and Review Department.
- IMF. 2000. "Argentina: 2000 Article IV Consultation and First Review Under the Stand-By Arrangement", *IMF Staff Country Report*, No. 00/164.
- IMF. 2001. "Argentina: Third Review Under the Stand-By Arrangement".
- IMF. 2003. "Argentina. Request for Stand-By Arrangements".
- IMF. 2004. "Evaluation Report. The IMF and Argentina, 1991-2001."
- Kaminsky, Graciela, Amine Mati, and Nada Choueiri. 2009. "Thirty Years of Currency Crises in Argentina: External Shocks or Domestic Fragility?" *NBER Working Paper 15478*.
- Kaminsky, Graciela and Carmen Reinhart. 1999. "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems", *American Economic Review*.
- Kaminsky, Graciela, Carmen Reinhart, and Carlos Vegh. 2004. "When It Rains, It Pours: Procyclical Capital Flows, and Macroeconomic Policies", *NBER Macroeconomics Annual*.
- Kannan, Prakash. 2010. "Credit Conditions and Recoveries from Recessions Associated with Financial Crises", *IMF Working Paper*, WP/10/83.
- Krugman, Paul. 2010. "How Did Economists Get It So Wrong", *New York Times Magazine*.
- Laeven, Luc and Fabian Valencia. 2008. "Systemic Banking Crises: A New Database", *IMF Working Paper 08/224*
- Mankiw, Gregory, and Matthew Weinzierl. 2011. "An Exploration of Optimal Stabilization Policy".
- Moe, Thorvald, Jon Solheim and Ben Vale. 2004. "The Norwegian Banking Crisis", Norges Bank Occasional Papers, No. 33.
- Nyberg, Peter, and Vesa Vihriala. 1993. "The Finnish Banking Crisis and Its Handling", *Bank of Finland Discussion Papers*.
- Reinhart, Carmen and Kenneth Rogoff. 2008. "Banking Crises: An Equal Opportunity Menace", *NBER Working Paper No. 1458*.
- Reinhart, Carmen and Kenneth Rogoff. 2009a. "The Aftermath of Financial Crises". *American Economic Review*, 99, 466-472.
- Reinhart, Carmen and Kenneth Rogoff. 2009b. "This Time is Different: Eight Centuries of Financial Folly".
- Roubini, Nouriel, and Giancarlo Corsetti. 1996. "Budget Deficits, Public Sector Solvency and Political Biases in Fiscal Policy: A Case Study of Finland", *Finnish Economic Papers*, Vol. 9, Number 1.
- Sinclair, Tara. 2008. "Asymmetry in the Business Cycle: Friedman's Plucking Model with Correlated Innovations". *Studies in Nonlinear Dynamics & Econometrics*: Vol. 14, No 1, Article 3.
- Weise, Charles. 2007. "A Simple Wicksellian Macroeconomic Model".

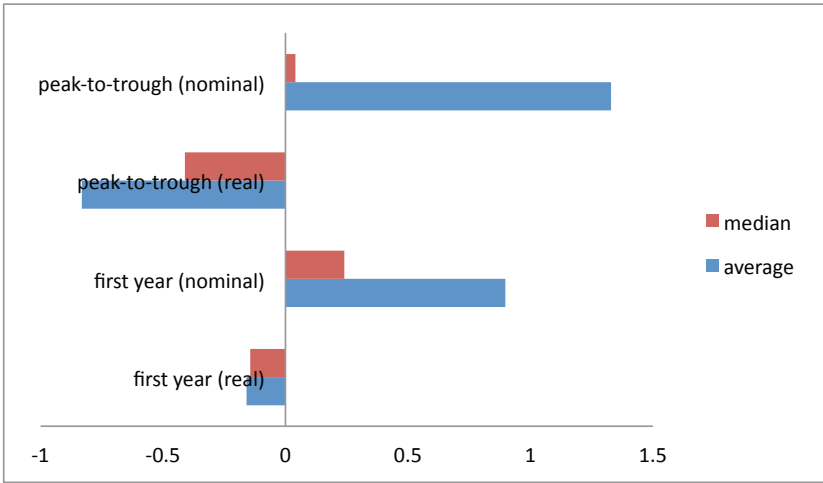
**Figure 1: Output Dynamics in Banking Crises (growth rates)**



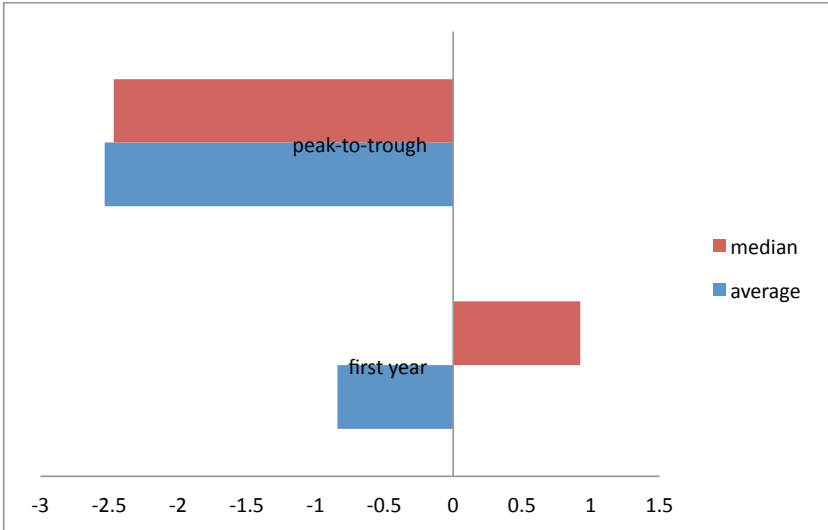
**Figure 2: Duration of Banking Crises (in years)**



**Figure 3: Change in Interest Rates** (first year of crisis and peak-to-trough)



**Figure 4: Change in Government Consumption** (first year of crisis and peak-to-trough)



**Table 1: Recessions in Advanced Countries (peaks and troughs)**

<b>Country</b>	<b>Peak</b>	<b>Trough</b>
Australia	1960 Q3	1961 Q3
Australia	1965 Q2	1966 Q1
Australia	1971 Q3	1972 Q1
Australia	1975 Q2	1976
Australia	1977 Q2	1978
Australia	1981 Q3	1983 Q2
Australia	1990 Q2	1991 Q2
Austria	1978	1979 Q1
Austria	1981	1982 Q1
Austria	1984	1985 Q1
Austria	2001	2002 Q1
Belgium	1982	1983 Q1
Belgium	1992	1993 Q1
Belgium	1998	1998 Q3
Canada	1980 Q1	1980 Q3
Canada	1981 Q2	1983
Canada	1990 Q1	1991 Q1
Denmark	1977 Q3	1978 Q1
Denmark	1980	1981 Q1
Denmark	1988	1988 Q3
Denmark	1993	1993 Q3
Denmark	1995	1995 Q3
Denmark	2002	2003 Q1
Denmark	2005	2006 Q1
France	1974 Q3	1975 Q1
France	1992 Q3	1993 Q2
Germany	1962 Q3	1963 Q1
Germany	1966 Q3	1967 Q2
Germany	1974 Q1	1975 Q2
Germany	1978	1978 Q2
Germany	1980 Q1	1982 Q3
Germany	1992 Q1	1993 Q1
Germany	1995 Q3	1996 Q1
Germany	2002 Q3	2003 Q2
Italy	1981	1981 Q3

Italy	1982 Q1	1983
Italy	1992 Q1	1993 Q3
Italy	1996 Q1	1997
Italy	2001 Q1	2002
Italy	2003	2003 Q2
Italy	2004 Q3	2005 Q1
Japan	1974	1975 Q1
Japan	1993 Q1	1993 Q3
<b>Japan (financial crisis)</b>	<b>1997 Q1</b>	<b>1999 Q1</b>
Japan	2001 Q1	2002
Netherlands	1980 Q1	1981 Q3
Netherlands	1982 Q1	1983 Q1
Netherlands	2001 Q2	2002 Q1
Portugal	1978	1978 Q2
Portugal	1981 Q2	1982 Q1
Portugal	1983	1984 Q1
Portugal	1992 Q1	1993 Q1
Portugal	2002	2003 Q1
Portugal	2005 Q2	2007 Q1
Spain	1975	1975 Q2
<b>Spain (financial crisis)</b>	<b>1978 Q3</b>	<b>1979 Q1</b>
Spain	1981	1981 Q2
Spain	1992 Q1	1993 Q2
Switzerland	1981 Q2	1983
Switzerland	1986 Q2	1987
Switzerland	1990 Q3	1991 Q2
Switzerland	1992 Q1	1993 Q1
Switzerland	1996 Q1	1996 Q3
Switzerland	2002 Q2	2003 Q2
United Kingdom	1961 Q2	1962
United Kingdom	1973 Q2	1974 Q1
United Kingdom	1974 Q3	1975 Q3
United Kingdom	1979 Q2	1981 Q1
United Kingdom	1990 Q2	1991 Q3
USA	1969 Q3	1971
USA	1974	1975 Q1
USA	1980 Q1	1980 Q3
USA	1981 Q3	1982 Q1
USA	1990 Q2	1991 Q1

**Table 2: Financial Crises included in the sample (Laeven and Valencia, 2008)**

Country	Start of Financial Crisis				
		Dominican Republic	2003	Norway	1991
		Ecuador	1998	Panama	1988
Argentina	2001	Ecuador	1982	Paraguay	1995
Argentina	1995	Finland	1991	Peru	1983
Bangladesh	1987	Ghana	1982	Philippines	1983
Benin	1988	Guinea-Bissau	1995	Philippines	1997
Bolivia	1986	Guyana	1993	Poland	1992
Bolivia	1994	Hungary	1991	Russia	1998
Brazil	1990	India	1993	Senegal	1988
Brazi	1994	Indonesia	1997	Slovak Republic	1998
Bulgaria	1996	Jamaica	1996	Spain	1977
Burkina Faso	1990	Japan	1997	Sri Lanka	1989
Burundi	1994	Jordan	1989	Swaziland	1995
Cameroon	1987	Kenya	1985	Sweden	1991
Cameroon	1995	Kenya	1992	Thailand	1997
Cape Verde	1993	Korea	1997	Thailand	1983
Central African Rep.	1995	Kuwait	1982	Togo	1993
Chad	1992	Latvia	1995	Tunisia	1991
Chile	1981	Lithuania	1995	Turkey	2000
Chile	1976	Madagascar	1988	Uganda	1994
China, P.R.	1998	Malaysia	1997	United States	1988
Colombia	1998	Mali	1987	Uruguay	1981
Colombia	1982	Mexico	1994	Uruguay	2002
Congo, Republic of	1992	Mexico	1981	Venezuela	1994
Costa Rica	1987	Morocco	1980	Vietnam	1997
Costa Rica	1994	Mozambique	1987	Yemen	1996
Croatia	1998	Nepal	1988	Zambia	1995
Czech Republic	1996	Nicaragua	2000	Zimbabwe	1995
		Nigeria	1991		



# **Local Technological and Demographic Effects on Electricity Transmission: A Spatially Lagged Local Estimation of New England Marginal Losses**

*Jacob Hochard*

## **Abstract:**

Electricity transmission is subject to distribution losses and congestion costs. Economists have prior theorized that these transmission imperfections could create divided markets with electricity generating spatial oligopolists. This concern has been largely dismissed because of recent technological advances in electricity transmission. The effects of local technological and demographic indicators on electricity transmission costs remains both commonly accepted as negligible and spatially untested. This analysis employs a spatially lagged local estimation of New England’s marginal electricity losses with respect to both technological and demographic indicators. The results of this analysis are consistent with the widely accepted notion that technological advances have mitigated the effect of local distribution losses and local congestion costs on electricity prices.

**Keywords:** Transmission grid losses, locational marginal prices, New England ISO, technological indicators.

## **Introduction**

Electricity markets and other networked goods like water, oil, cable television and railways have become an enjoyable research pastime for economists concerned with market structure issues. Electricity markets are embedded with unique commodity specific and trade specific considerations that complicate the analysis of this market. The two most important electricity specific considerations are (1) the physical laws that electricity must abide by and (2) the limitations of electricity transmission via infrastructural constraints.

Electricity, in the form of electrons transmitted along a transmission cable, is subject to a certain degree of “resistance”. This resistance is defined in the form

of two foundational laws of electricity (1) Ohm's Law and (2) Kirchoff's Law. Ohm's law connects the three concepts of current, voltage and resistance. Current can be defined simply as electrons moving through a transmission cable between two points. Voltage is the force that allows these electrons to move between two points and resistance is the measurable inhibiting force between those two points caused by electron transit friction (Kostiner, 1994). Kirchoff's Law states that the "sum of all currents entering and exiting a node must be equal to zero". In essence, this law states that electrons will travel the path of least resistance. These two laws have direct applications to electricity trade markets that must be taken into consideration (Kostiner, 1994).

Under Kirchoff's law, unlike most traditional commodities, electricity cannot be stored for future consumption. The transmission grid is therefore always in a state of perfect production-consumption equilibrium. Ohm's law also states that with "resistance" as a function of voltage capacity (the size of the transmission cable) and current strength there will be inherent distribution losses of electricity. These distribution losses are also a function of the distance that electrons travel from the point of production to the point of consumption and the local weather (Robertazzi, 2007). The existence of "resistance" on the transmission grid also makes possible transmission grid "congestion" to occur. This congestion occurs when the current from point A to point B on the transmission grid reaches the voltage capacity of that respective transmission line. Under Kirchoff's law, however, we know that electrons will take the path of least resistance and congestion into and out of any particular node will be equal.

The physical properties of electricity also have unique benefits. In theory, a transmission grid without congestion (a high enough voltage and low enough current) can transfer electricity across long distances, instantly. It has been argued, for example, that electricity produced from solar resources in Northern

Africa could be traded competitively in the European electricity markets with proper transmission voltage capacity expansions (Bauer et. al., 2008). The demand schedule of electricity is predictable but inconstant. For example, peak-demand typically occurs during a warm day in the early evening when electricity consumption is high. The transmission grid's unique range may allow distant competitive suppliers of electricity to exploit far-reaching peak-demand markets during their off-peak hours (Shakourig et. al., 2009).

Conventional theory recognizes price increases as a result of distribution losses and congestion but the impossibility of predicting where produced electricity will be consumed has overshadowed the potential influence of citing generation sources in close proximity to electricity demand centers. Furthermore, the spatial analysis of these grid losses, on a local scale, may have lost its appeal because advances in transmission capacity are occurring rapidly and better integrating larger regions. In theory, this would make local indicators less important and spatial demographic and technological indicators less predictable. The purpose of this analysis is to estimate locational marginal price losses, resulting from electricity distribution losses and congestions, using local and demographic indicators.

## **Literature Review**

Two and a half decades ago Benjamin Hobbs (1985) predicted that the deregulation of electricity generation would create spatial oligopolists resulting from network barriers. Hobbs conducted a theoretic Nash-Bertrand spatial equilibrium to predict the price variation of electricity in New York's regional markets. His results showed that transport costs and significant scale economies would yield generator spatial oligopolists. The spatial oligopolists would cause regional price increases and the ability for generators – with natural

barriers caused by transport costs - to exercise market power. Over the past 25 years, however, the scale of these economies has increased drastically. The grid is interconnected extremely well and technological advances have allowed for less costly electricity transmission across further distances. These changes may have removed the natural barriers that Hobbs envisioned in the mid-1980s. Additionally, Hobbs' analysis is conducted using a theoretic price equilibrium calculated using mathematically linear programs to obtain local spatial price equilibria. This theoretic concept deserves attention using spatial analysis.

The Independent Electricity Market Operator (IEMO) in charge of operating the electricity grid in Ontario released a PowerPoint in January, 2004 outlining historical nodal pricing analysis on their grid. This operator references spatial analysis and its relevance to the impact of congestion and relative losses on the electricity market. The presentation uses locational marginal prices (LMPs) that include a congestion and loss component<sup>1</sup>. This analysis found that losses, not congestion, contributed the most to pricing variation on the local grid. These system operators have perfect information and were able to determine which transfers incurred the highest losses. In this case, the highest rate of congestion occurred along the East-West Transfer interfaces, whereas the highest losses from distribution occurred between the Northwest and Northeast regions of the grid. No spatial analysis was considered to determine if generators' proximity to demand centers influenced grid losses.

Ostergaard (2004) examines critically the geographic distribution of electricity generation in relation to grid losses in Denmark. The Danish example is particularly interesting because over 40% of consumed electricity is covered by scattered sources as a result of large scale wind turbine investments. Ostergaard

---

<sup>1</sup> In theory, an electricity system (in this case IEMO) will have one theoretic price across the entire system – any pricing deviations occur as a result of incurred congestion and distance of travel losses.

adopts basic assumptions to map the distribution of generated electricity in order to conduct a consumption-production spatial analysis. In order to model transmission grid flows, Ostergaard uses EnergyPro GRID a complex model founded on an algorithm designed to predict transmission flows throughout a grid. His analysis concludes that it is essential for Denmark to control generation not only at an aggregate level but also at a local level to prevent congestion from occurring. A suggestion is given to coordinate local and regional electricity production to ensure a fair balance without the inefficient alternative of electricity traveling far distances, incurring distribution costs.

Baban and Perry (2000) used spatial analysis to determine the optimized locations of new wind farm investments in England. These clean electricity generators were determined based upon geographic constraints (including topography, land use, wind direction, wind speed, population, road access, hydrology and historical and cultural land marks). The only factor that was considered with regard to transmission compatibility was a constraint that the wind farm is within 10 KM of the transmission grid. On the demand side, the only consideration with regard to population was actually a 2 KM buffer on large settlements. This type of consideration speaks volumes of traditional electricity generator citing ideologies. The cost in distribution losses, transmission losses and congestion losses are not considered carefully when citing an electricity generator in distant proximity from its intended consumers of electricity.

## **Methodology**

I retrieved the source data for LMP nodes across New England for the year of 2008. This data was created by the Independent System Operator of New England (ISO-NE). I used Google Earth to locate the coordinates of each of these nodes and converted this coordinate data to a point data shapefile. This

pricing data is in terms of \$/Kilowatt Hour (KWH) and is valid for June 11<sup>th</sup>, 2008 for electricity trade from 6:00PM – 7:00PM<sup>2</sup>. I retrieved source data for electricity generators, with their respective generating capacities, present in New England valid for the year of 2008. I used the Environmental Protection Agency's (EPA) facility registration database to locate each of these generator's respective addresses. I then used Google Earth to locate the coordinates of these addresses and created a point data shapefile with these coordinates (Figure 1). United States (U.S.) spatial demographic data valid for the year of 2000 was retrieved from the U.S. census, to create demographic indicators (population and population density). Finally, transmission grid spatial data, including individual line voltage capacities, was obtained from the Federal Emergency Management Agency (FEMA) valid for 1993.

I created 587 Thiessen polygons around my 813 LMP nodes (Figure 2). In some cases, there was more than one node located at an identical location. These prices were averaged because under Ohm's and Kirchoff's laws and the framework of the transmission network, electricity prices at identical locations are by identity, equal. I estimated the population and population density of each Thiessen polygon by converting my census block-level population data to a raster file and then using zonal statistics to sum population. I then calculated the geometry of these polygons and conducted a simple field summation to determine estimated population density. I also use field calculations to estimate electricity generation capacity, transmission capacity/per capita, and total transmission length within each polygon.

I employed a spatially-lagged ordinary least squares (OLS) model to estimate the effect of these spatial and technological indicators on the nodal marginal

---

<sup>2</sup> This date and time was chosen because it was one of the warmest days recorded during the summer of 2008. This year was also the most recent year that LMP data was available.

electricity losses. In addition to the focal indicators, I included a dummy variable that was present (1) if my dependant variable spatial accuracy was to the street level and not-present (0) if the dependant variable was only accurate to the town level.

$$Y_i = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6Y_{Ni} + u_i$$

Y = Marginal Loss Component (\$/KWH)

X<sub>1</sub> = Length of grid (Miles)

X<sub>2</sub> = Generator Capacity (Kilowatts)

X<sub>3</sub> = Voltage Capacity / Per Capita (Kilovolts)

X<sub>4</sub> = Population Density (PP/KM<sup>2</sup>)

X<sub>5</sub> = Precision Dummy Variable (1,0)

Y<sub>Ni</sub> = Spatially Lagged Neighborhood Weights of Marginal Loss Component

Lastly, a breusch-pagan test was employed to test our estimates for the presence of heteroskedastic errors.

## Results

None of the focal explanatory variables had an estimated coefficient that was statistically significant in difference from zero (Table 2). The coefficient estimate on the dummy variable for dependant variable precision (at the street level) was negative and statistically significant. This dummy variable suggests that my flawed data reporting accuracy does affect my overall estimates. This dummy variable coefficient is relatively intuitive as it would appear that estimated effects on a marginal loss price component would be less in an area that the node may not actually exist. These nodal centers are likely to have higher population densities and transmission grid presence. Flawed accuracy may discount this estimation.

The spatially lagged estimated coefficient is positive and statistically significant. This is expected as most of our chosen variables are inherently spatially-autocorrelated (Table 1). The worst spatial autocorrelation exists within our dependant variable with a positive Moran's I coefficient of 0.79 (Table 1). The coefficient estimates do not appear to have heteroskedastic error terms but despite the spatially lagged variable the coefficient estimates still suffer from spatial autocorrelation.

## Conclusions

These results do not support my null hypothesis that local technological and demographic indicators affect electricity's marginal loss component at LMP nodes. These results support multiple conclusions. In a perfectly operating transmission network there would be no variation across our spatially lagged variable. That is, the distribution losses would be constant and minimal across spatial units. The estimated model only explains about three-quarters of the variation in our loss component. We can conclude, therefore that there are technological and demographic negative influences on the New England transmission network causing distribution losses.

Since we have variation in our distribution loss prices but that variation cannot be explained with local indicators, we can conclude that the distribution losses are being incurred at locations beyond the reaching of our spatial "Thiessen polygon" units. This may support that electricity is being produced in distant locations from where it is being consumed. This conclusion is a success story for the New England transmission grid. A distant spatial relationship between supply and demand of electricity supports that there is little congestion mitigating distant trade. This conclusion is also supported by the Federal Energy Regulatory Commission (FERC) stating recently that New England is a transmission system with close to no transmission congestion.

This study does suffer from many limitations. This study does not account for a potential "edge effect". I was not able to obtain import and export data for New England or spatial data for the neighboring New York Independent System Operator (NYISO). New York City is a major demand center in close proximity to Connecticut. This may be one reason why the left-hand side observations in this area have such a high Moran's I coefficient (Figure 3). Also, limited resources and funding have forced me to use imperfect data. My transmission grid data was created by FEMA for national security impact assessment not transmission grid analysis. This particular dataset is also two decades invalid. Finally, despite using a spatially-lagged model, my regression estimates still suffer from the presence of autocorrelated errors. I chose not to pursue this problem any further because the

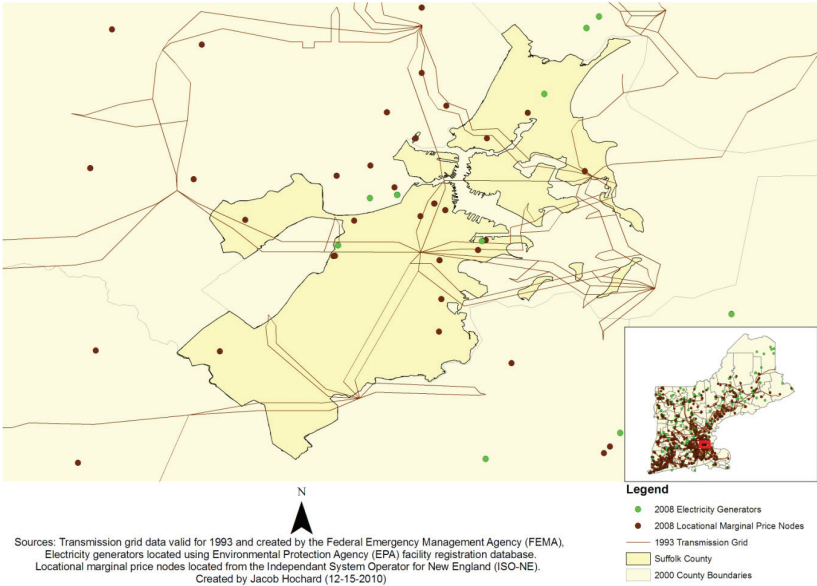


spatially autocorrelated errors preserve unbiased but inefficiency estimates. The relatively low z-scores of my estimates indicate that even with corrected errors the coefficients would likely remain statistically insignificant.

### Appendix

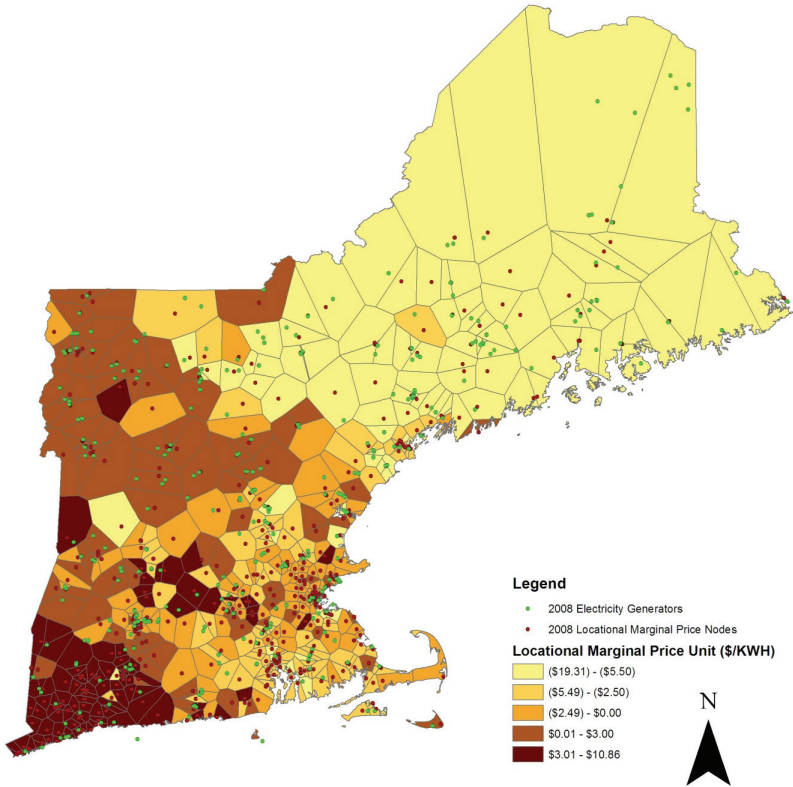
**Figure 1** – Electricity structure in Suffolk County, Boston, MA including 2008 electricity generators, 2008 locational marginal prices, 1993 transmission grid, and the local county boundaries.

Figure 1: Suffolk County, Boston, MA Electricity Infrastructure



**Figure 2** – New England electricity marginal losses (\$/KWH) for June 11<sup>th</sup>, 2008 from 6:00PM-7:00PM. This map includes generators and locational marginal price (LMP) nodes.

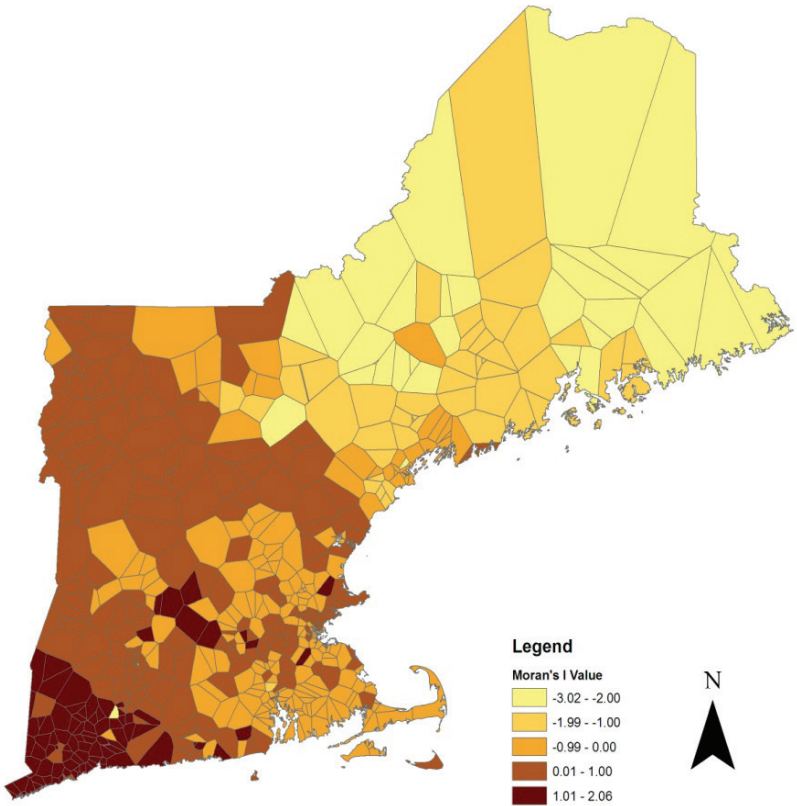
**Figure 2: New England Electricity Marginal Losses (\$/KWH)  
(June 11th 6:00 PM-7:00 PM, 2008)**



Sources: Electricity generators located using Environmental Protection Agency (EPA) facility registration database.  
 Locational marginal price nodes located from the Independent System Operator for New England (ISO-NE).  
 Created by Jacob Hochard (12-15-2010)

**Figure 3** – Thiessen polygon-level Moran’s I values for New England Electricity Marginal Losses on June 11<sup>th</sup>, 2008 from 6:00 PM-7:00 PM.

**Figure 3: New England Electricity Marginal Losses  
Moran’s I Values (June 11th 6:00 PM-7:00 PM, 2008)**



Sources: Moran's I values calculated using Geoda. Electricity price data obtained from the Independent System Operator of New England (ISO-NE). Map created by Jacob Hochard (12-15-2010).

**Table 1** – Estimations from spatially-lagged ordinary least squares (OLS) regression as well as summary statistics and tests for heteroskedastic and spatially autocorrelated error terms.

Variable	Coef.	Std. Error	Z-score	Prob >  z
Lagged marginal loss component (\$/KWH)	0.8545	0.0226	37.85	0.0000
Precision Dummy	-0.8103	0.2433	-3.33	0.0009
Length (Miles)	-0.0522	0.1096	-0.48	0.6337
Generating Capacity (KWH)	-0.0304	0.0318	-0.96	0.3384
Grid Capacity/PP (KV/PP)	0.1182	0.3604	0.33	0.7430
Pop Density (PP/KM <sup>2</sup> )	0.0003	0.0003	1.02	0.3073
Constant	-0.2295	0.4526	-0.51	0.6122

Summary Statistics					
R-Squared	0.7679		Likelihood Ratio Test	p=0.0000	
Breusch-Pagan Test	p=0.49				

**Table 2** – Univariate Moran’s I coefficients for each variable.

Variable	Moran’s I Coefficient
<b>Locational Marginal Prices (\$/KWH)</b>	<b>0.7940</b>
<b>Length of transmission grid (Miles)</b>	<b>0.3188</b>
<b>Generation capacity (KWH)</b>	<b>-0.0029</b>
<b>Capacity/Per Capita (KV/PP)</b>	<b>0.3729</b>
<b>Population Density (PP/KM<sup>2</sup>)</b>	<b>0.5140</b>

## Works Cited

- Baban, Serwan M.J., and Tim Parry. "Developing and Applying a GIS-assisted Approach to Locating Wind Farms in the UK." *Renewable Energy* 24 (2001): 59-71. Print.
- Bauer, Nico, Ottmar Edenhofer, Michael Jacob, Sylvie Ludig, and Michael Luken. "Electricity Trade among World." Proc. of Potsdam-Institute for Climate Impact Research, Germany. 14 Nov. 2008. Web. 11 Dec. 2010. <[http://www.nccr-climate.unibe.ch/conferences/climate\\_policies/working\\_papers/Bauer.pdf](http://www.nccr-climate.unibe.ch/conferences/climate_policies/working_papers/Bauer.pdf)>.
- Federal Emergency Management Agency. Transmission Grid Data with Kilovolts. 1993. Spatial data. Washington, DC.
- Hobbs, Benjamin F. "Network Models of Spatial Oligopoly with an Application to Deregulation of Electricity." *Operations Research* 34.3 (1986): 395-409. Print.
- "Independant Electricity Market Operator - Spatial Analysis." *Market Evolution Program*. Proc. of Historical Nodal Pricing Analysis, Ontario. 2004. Print.
- Independant System Operator of New England. Electricity generators operating in New England. 2008. Raw data. ISO-NE, Holyoke, MA.
- Independant System Operator of New England. Electricity locational marginal prices and nodes. 2008. Raw data. ISO-NE, Holyoke, MA.
- Kostiner, Barry. "Spatial Market Equilibrium for Resistive Electric Networks." *Library Archives*. Massachusetts Institute of Technology - Operations Research Center, 12 Oct. 1994. Web. 11 Dec. 2010. <<http://dspace.mit.edu/handle/1721.1/37544>>.
- Ostergaard, Poul A. "Modelling Grid Losses and the Geographic Distribution of Electricity Generation." *Renewable Energy* 30 (2005): 977-87. Print.
- Robertazzi, Thomas G. *Networks and Grids Technology and Theory*. New York: Springer, 2007. Print.
- Shakourig, H., M. Eghlimi, and D. Manzoor. "Economically Optimized Electricity Trade Modeling: Iran-Turkey Case." *Energy Policy* 37.2 (2009): 472-83. Print.
- U.S. Census Bureau. 2000 Census Data. 2000. Spatial Tiger/line data. Washington D.C.

**“Does It Pay to Be Informed?”**  
**Expenditure Efficiency in the US Mutual Fund Industry**  
*Jan Cerny*

**Introduction**

The mutual fund industry would like us to believe that fund expenses are justifiable by their extensive management expertise, security analysis and the consequent delivery of returns that exceed the market performance. Management know-how is costly and thus it drives up the expenditure of actively managed mutual funds and potentially lowers their net returns. Nevertheless the fund managers argue that their contributions to the returns fully outweigh their costs and in general their trading strategies add value to the investors. On the other hand many academics hold that such claims are fundamentally misleading and actively managed funds cannot continuously outperform a market index (See: Carhart 1997, Jensen 1968, Malkiel 2003, Sharpe 1964).

This study aims to provide additional insight into the debate by examining the performance of US equity mutual funds over the period of 2002 – 2010. I carry out empirical analysis to evaluate relative performance of the funds and test whether managers can justify their expenses and fees by higher risk adjusted returns. This provides valuable implications about the validity of the Efficient Market Hypothesis (EMH) as developed by Sharpe (1964) and is beneficial to the broad public that engages in various fund-picking strategies.

The preponderance of studies regarding mutual fund performance indicates that the topic is of crucial importance to the academics, practitioners and general public. Unlike most of the previous works, I do not focus on individual characteristics of funds that could be used for prediction of future returns. Instead,

I examine the performance of the mutual fund market as a whole and test the efficiency of resource expenditure across the industry over a period that is yet to be fully covered by the researchers. I explore the concept of costly information in financial markets and hypothesize that the market is in informational equilibrium where resources are spent efficiently.

For fund expenditure to be “efficient”, it must satisfy the equilibrium condition that the marginal cost equals marginal benefit. Return maximization is the proclaimed primary goal of mutual funds and so the funds should generate new expenses only if such expenses are offset by resulting higher returns. Thus, in theory, any extra research and trading may take place only if they add value to the fund. If such activities that are inevitably costly do not add enough value to outweigh their cost, the industry does not spend their resources efficiently. Such finding would suggest investors should focus on funds that minimize their expenditure to the point where the marginal cost of their activities equal their marginal benefit.

In addition to examining the EMH, analyzing performance of the funds and efficiency of their expenditure, this study provides insight into the controversial assumption of perfect investor rationality. Theoretically, assuming perfect information and rational consumers, investors would not pay high fees to mutual fund managers unless the managers could deliver (or create credible expectation of) returns that would exceed the management costs. In other words, if active trading did not add value, rational consumers would adjust in a long run and seek alternative investment strategies which would diminish the demand for actively trading funds.

In the first section of this paper, I review past literature and examine its contributions and shortcomings. In the second section I discuss the theoretical

background of informational equilibrium and resource expenditure efficiency. Next I introduce the Capital Asset Pricing Model (CAPM) and develop its extended version that I employ in my panel data analysis. Subsequently I discuss my data in the third section and present the empirical evidence in the fourth section. Lastly, I point out the limitations of my work and draw conclusions regarding my hypotheses.

## **Literature Review**

The ability of mutual fund managers to earn excess risk-adjusted returns<sup>1</sup> has been of great interest to researchers for years. Prior to the 1990s the general consensus of academics was that investors are not able to earn such returns and no fund characteristics could substantially aid them in predicting which managers will become the next winners or losers. Nonetheless, numerous studies after 1990 arrived at opposite conclusions, claiming that returns on mutual funds and underlying securities are predictable to a certain degree. These researchers concluded that some types of analysis and trading activity allow for superb returns, which supports the case of “skilled managers” (Malkiel 1995). The literature on the performance of asset management strategies and mutual funds that is relevant to this study can be divided into three general categories:

### 1.) Testing the efficient market hypothesis (EMH).

Eugene Fama gave birth to the EMH in the 1960s claiming that, under the semi-strong version of the hypothesis, security prices instantly reflect all available public information. Consequently there is no information that the traders could employ to outsmart or time the market. Thus any charting or fundamental analysis will fail to generate substantial risk-adjusted excess

---

<sup>1</sup> Returns in excess of the risk free rate on Treasury Bills are generally referred to as “excess returns”



returns. The weak form of the EMH holds that there are no patterns in stock prices and so the active management is likely going to significantly increase expenses while only marginally contributing to returns. Therefore, no one is able to systematically benefit from the inefficiencies of the market and no research or expertise can enhance the fund returns over several years.

Ever since its creation, EMH has been tested by hundreds of empirical studies that aimed to determine the extent to which markets are efficient. The early tests focused on charting and technical analysis, that are often associated with active trading, and found most such techniques utterly worthless in predicting future price movements (Karz 2010). However, the professionals practicing these arguably futile methods have not been driven out of the market and so, under the assumption of consumer rationality, their service must be considered valuable. In reality most financial institutions continue to spend billions to support their technical analysis departments.

Academics have also identified several anomalies and patterns that would allow active traders to capture substantial risk-adjusted excess returns, such as the “size effect” or the “January effect”. Nevertheless, many studies concluded that once such patterns are documented and made public, the investors exploit these new opportunities to the extent that the patterns disappear or become unprofitable. The academic research also points out the paradox of EMH. This paradox states that, if all investors believed that markets were efficient, no one would spend resources on thorough asset analysis, and so the market would effectively become inefficient. Thus, the fund managers who do not believe in market efficiency and carry out asset research in pursuit of outperforming the benchmarks play a crucial role in actually making the market efficient. Overall, academic research and back-testing provided a relatively strong support for the validity of EMH across different periods and diverse markets. Therefore, high expenses generated

by active management research are not likely to be sufficiently offset by their higher returns.

2.) Existence of manager stock picking or market timing skills.

Many researchers designed empirical studies to test the existence of skill or talent of portfolio managers that would enable the funds to pick the winning stock or to properly time their market entry and exit points. Generally, the scholarly literature refutes the concept of superior stock picking skills as a determinant of fund returns (See the renowned study by Carhart 1997 or Henriksson 1984). Nonetheless, a limited number of studies argue that some managers do possess exceptional skills that allow them to exceed the market returns with some level of persistence (See: Gray and Kern 2010). Hendricks and Zeckhauser (1991), for example, examined the period of 1975-88 and found that extensive research and active management strategy of mutual funds could yield an excess return of 3% to 4% every year net of expenses. If such skills did exist my analysis should indicate that at least some funds were able to significantly outperform the market after expenses.

3.) Persistence in mutual fund performance.

A large body of literature focuses on the persistence in mutual fund performance claiming that, if there were outstanding actively trading mutual fund managers, it would be likely that their excess returns would display some level of continuity. Said differently, good players would be expected to win more often than others. Nonetheless, past research does not support the existence of long term persistence in mutual fund returns and the higher the expenses the shorter the persistence in positive returns usually is. For example Carhart (1997) documented that, even though some evidence for short term persistence of returns can be found, future performance of mutual funds is almost impossible to predict.

In addition to this wealth of academic literature, thousands of investment practitioner articles discuss approaches that should allow investors to select the best mutual funds based on their history or characteristics. This literature often holds turnover and expense ratios to be substantial determinants of fund returns. Some suggest that low turnover (buy-and-hold strategy) and low expenses are desirable while others believe that high turnover (active trading) may be an indicator of sound strategy and that higher transaction costs of frequent trading are fully offset by increased returns. This segment of the literature fails to reach a consensus regarding the role of expenses in determination of returns and often suffers from severe methodological problems such as omission of survivorship bias, which leads to false sense of return predictability (Peterson et al. 2002)

## **Methodology**

This study builds on the theories developed by Grossman and Stiglitz (1980) and Ippolito (1989) who introduce the concept of costly information into the debate over the validity of EMH. Grossman and Stiglitz assert that EMH cannot hold since prices cannot reflect all available information, because if they did, traders who spent resources on obtaining such information would not receive any compensation (1980). I apply this framework on the equity mutual fund market and focus on the role of expenditure that is associated with the acquisition of information. Most of the “active trading” strategies rely on the premise of special skill or information of the managers. These strategies are bound to be very costly as they are characterized by notably high turnover, which increases transaction costs, and higher management fees, resulting from employment of larger amounts of human capital (Sharpe 1991, Carhart 1997). It follows, that for these strategies to be successful, their benefit needs to outweigh their cost. Conversely, passive management strategies such as indexing could be classified as a buy-and-hold

strategy with a predictably low portfolio turnover and low expenses. Thus observing a negative relationship between expenses (indicating information acquisition) and the excess returns would suggest that active management at its high levels does not add value to investors and passive investment strategies should be pursued.

Recognizing the existence of these dissimilar management styles, I make the simplification that the market can be divided into two types of traders: the “informed” and the “uninformed” ones. The “informed” managers believe in existence of some information or skill that can increase their performance even net of expenses. On the other hand, the “uninformed” managers believe that there is no such information that would be worth looking for. Said differently, these traders hold that active management with its extensive research creates more expenses than it can offset by potentially higher returns. Thus the “uninformed” traders generally follow a market index and focus on minimizing their expenses.

In this work I utilize the information equilibrium theory to address the claims of both types of traders. It is clear that acquisition of information and skills requires expenditure of time and other resources. Thus one would not engage in activity such as market research without expectations of appropriate rewards. Rational agents are on average able to learn from their experience. Therefore if the agents did not receive any rewards for their expenditure, they would no longer pursue the path that proved fruitless. Given these assumptions, in equilibrium, the marginal return to additional research or information will equal its marginal cost. In such equilibrium, all incentives to get more or less informed diminish.

Applying this framework to the EMH, it seems plausible that managers are able to outperform the market before expenses. Nevertheless, the risk-

adjusted excess returns disappear after the as the expenses are subtracted from the higher returns. If the managers that focus on costly research and trading were able to outperform the market net of expenses on continuous basis the relationship between their expenses and returns would be positive. Conversely, if the traditional form of EMH holds, trading on special skill or information is essentially a losing game as such practice can only increase expenses without enhancing the returns, which reflects a negative relationship between expenses and fund performance. Lastly, if the market is in informational equilibrium, there will be no relationship between expenses and returns net of fees as any excess returns created by extensive research will be exactly offset by higher cost.

In this work I test this relationship over a broad sample of 500 mutual funds. First, I employ the renowned Capital Asset Pricing Model (CAPM) as developed by Sharpe (1964) to analyze the actual performance of the funds. I estimate the following time series regression model for each fund in the dataset:

$$(1) R_{jt} - R_{ft} = \alpha + \beta (R_{mt} - R_{ft}) + \mu$$

Where  $R_{jt}$  is a return on a mutual fund net of fees in period  $t$ ,  $R_{ft}$  is a risk free interest rate at year  $t$  and  $R_{mt}$  is a return on a broad market portfolio such as the S&P 500. This model is widely accepted in the financial industry and allows me evaluate the relative risk-adjusted performance of the mutual funds. According to Sharpe (1964), the return on a security or a fund less the risk free rate is directly proportionate to the amount of risk that the fund takes on. This relies on the observation that investors need to be rewarded for taking on extra risk. Such reward is known as the CAPM risk premium. Thus risk, measured by the coefficient  $\beta$ , is the major determinant of returns. In general,  $\beta$  represents the sensitivity of expected excess returns on a fund or an asset  $j$  to the expected market returns, which is expressed by the following relationship:

The unexplained portion of the regression, reflected in the intercept  $\alpha$ , is then attributed to management skill and expertise. Positive alphas indicate that a fund was able to outperform the market on risk adjusted basis. Nevertheless, the EMH clearly suggests that the expected value of alpha is zero because, on average, funds cannot outperform the market as there is no information or practice that would enable them to continuously do so.

If active management and research do not add value, funds engaging in such practices will systematically underperform the index funds and will likely display significantly negative alphas as a result of high expenditures. However, in informational equilibrium, both actively managed and index funds will perform comparably, resulting in alphas that are mostly indistinguishable from zero. Furthermore, the average coefficient of  $\beta$  across the funds should be equal to unity as a random broad sample of widely diversified funds should in essence mimic the market, possessing on average as much risk as the market itself.

Thus I hypothesize the following:

$$H1: E(\alpha) = 0$$

$$H2: E(\beta) = 1$$

In the second part of the paper I use the respective alpha and beta estimates from (1) to examine the role of expenses and turnover in determination of fund returns. Inspired by Jensen and Ippolito, I expand the CAPM model by including the turnover and expense ratios of funds as well as the variable BMktRF ( $=\beta_j*(Rm_t - Rf_t)$ ), which is a multiple of estimated beta of a fund and the market return in excess of the risk free rate<sup>2</sup>. I construct a pooled dataset of the sample and estimate the following OLS panel regression model:

$$(2) R_{jt} - Rf_t = b \beta_j*(Rm_t - Rf_t) + \tau \text{turnover}_{jt} + e \text{expense}_{jt} + y \text{Year} + f \text{Fund} + \mu$$

---

2 For detailed discussion, see the data section.

Where *Year* and *Fund* are dummy variables unique to each fund and time period. These variables address the problem of correlation of the residuals. The coefficient  $\beta_j$  estimated in regression (1), becomes a part of an explanatory variable<sup>3</sup>. The way  $\beta_j$  is estimated in (1) results in the fact that the coefficient  $b$  on the variable  $\beta_j \cdot (R_{m_t} - R_{f_t})$  should be statistically insignificant in difference from unity and so this coefficient is not key for the inference about my hypotheses.

On the other hand the variables *turnover* and *expenses* play an essential role as they can explain some of the fund performance that was previously captured by alphas. I hypothesize that funds generally spend their resources efficiently. Therefore the coefficient on expenses should be statistically insignificant in difference from zero, supporting the irrelevance of expenses hypothesis. A positive coefficient would suggest that managers are not only able to offset the higher expenses created by research and trading, but that the extra returns of such strategy outweigh the extra costs.

Thus, unlike most studies that simply assume a negative relationship between expenses and returns net of expenses, I test the relationship and hypothesize a neutral impact of expenses on returns:

$$H3: e = 0$$

Lastly, to understand the connection between equations (1) and (2), one should take into account that the first model simply states that returns are determined by the movement of the market and an unexplained cluster of management skill and information. The second regression is then used to analyze this cluster and examine whether some of this unexplained portion of returns is attributable to expenses or fund turnover.

---

<sup>3</sup> Usage of an estimated coefficient as a part of an explanatory variable inevitably injects extra variation in the regression.

## Data

In my research it would be optimal to work with monthly Center for Research in Security Prices (CRSP) data that are used in most of the academic studies. However, my data selection is restricted by the research budget as the financial data of the mutual fund industry are generally very costly. I limited the range of my data to eight annual observations in years 2002-2010 for a universe of small-cap growth, large-cap growth, small-cap value and large-cap value equity mutual funds, as supplied by Lipper – the Thompson Reuters Company. Ideally all observation would be included for all time periods to make my panel data balanced. Nevertheless, as I point out in the limitations section, this is not the case and my dataset misses about 7% of its observations.

To construct a sample from this universe of 2191 funds I randomly select 500 mutual funds and categorize them according to their asset classes. The basic version of my dataset includes: fund returns net of expenses, turnover and expense ratios. Furthermore I add the excess return on the market ( $R_{m_t} - R_{f_t}$ ) denoted as  $MktRf$ . This variable was obtained from online “French and Fama Library” and is constructed as follows (Kenneth R. French - Data Library):

The excess return on the market is computed by subtracting the Treasury bill rate (obtained from Ibbotson Associates) from the value-weighted return on all stocks traded in the United States (obtained from CRSP). This variable is likely to move closely with the excess returns of any particular mutual fund and so I expect it to hold a significant explanatory power.

Furthermore, to answer my research question I construct the dependent variable for excess returns of the funds (*exreturn*) by subtracting the risk free interest rate on ten-year Treasury bills from the percentage return on the fund net



of fees. This provides me with a measure of returns in excess of the risk free rate that the fund was able to capture.

Lastly, it appears that the random sample is fairly representative of the market during the period. The distribution of the sample funds across categories is depicted in Table 1, and Table 2 summarizes my data. For detailed data summary see Appendix 1.

**Table 1**  
**Category Frequency in Percentages**

	Growth	Value
Large Cap	41%	21%
Small Cap	25%	12%

**Table 2**  
**Summary of Data**

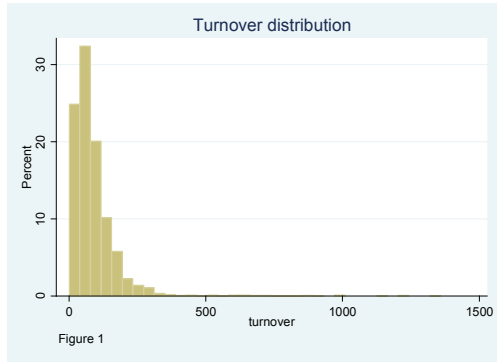
Variable	# Observations	Mean	St. dev	Minimum	Maximum
Exreturn	3701	3.44	25.9	-60.40	164.26
MktRf	4000	3.93	23.18	-39.94	32.12
Turnover	3658	90.82	97.99	1	1359
Expense	3613	1.46	.64	.07	12.42

Table 2 correctly points out that my panel is not balanced as the variables *exreturn*, *turnover* and *expense* are missing 299, 342 and 387 observations respectively. Although some observations are absent, the dataset does not suffer from substantial survivorship bias<sup>4</sup> as the vast majority of the sample funds survived throughout the examined period. In fact the mean number of periods observed per fund is 7.402 with minimum of 3 and maximum of 8 periods available per funds. Most of the unobserved periods seem to be a result of the simple fact that the fund were not yet in existence in the earlier years of the examined period. I further

<sup>4</sup> Survivorship Bias refers to a tendency to omit failed mutual funds from performance evaluation. If only funds that were successful enough to survive were included in the sample, the performance results could be skewed upwards as the sample would not reflect the inferior returns of funds that have gone out of business.

examine this potential problem of “creation bias” in the limitations section of this work.

Additionally, several other important observations arise from Table 2. The mean excess return (3.44%) is surprisingly high over a period of several economic slumps. Average turnover of a fund is



90.8%, suggesting that an average fund in the sample turned over about 91% of its holdings during a year. Turnover rate is calculated by dividing the fund’s total sales or purchases (whichever is less) by its average monthly assets. The rate then represents the percentage of the fund holdings that change over the course of the year. Taking this knowledge into account I notice an outlier in the turnover data: 1359. It seems unlikely that a fund would turn all its assets over more than 13 times in a year. In fact, Figure 1 illustrates that only a negligible percentage of the turnover observations are greater than 600. Since the large outlier might impact coefficient estimation I drop the outliers beyond five standard deviations from the mean (turnover of 580). This seems theoretically justifiable as it is improbable that even very active funds would turn their assets over more than six times a year (Wermers 2002). By omitting the potential outliers I drop 23 observations<sup>5</sup>, but the estimates of the model change only very marginally.

To account for the unique characteristics of each fund and each year I construct dummy variables that also enable me to address the problem of correlation between residuals. It is reasonable to assume that the funds are unique as different fund managers arguably possess different skills and employ dissimilar

<sup>5</sup> Only 0.63% of turnover data is omitted under this restriction.

investment strategies. For the sake of simplicity, I assume that fund managers did not change during the examined period or that if a manager left, the fund carried on the investment strategy, style and know-how of the original manager. This also implies that the targeted level of risk of the fund, reflected by the fund's Beta, has not changed during the period, which is an essential assumption of the CAPM analysis. Using the methodology of Ippolito I test this assumption against my data by running a regression for each fund with a dummy variable  $D$  and its interaction term with the  $MktRf$  variable. The regression equation takes on the following form:

$$(3) \quad R_{jt} - R_{ft} = \alpha + \beta (R_{mt} - R_{ft}) + c D + d (R_{mt} - R_{ft}) * D + \mu$$

where  $D$  is a dummy variable for years 2006-2009. If the coefficient  $d$  were statistically significant in difference from zero, the assumption of constant beta would be highly questionable. I find that 82 funds or about 16% of my sample display betas that are not stable at the 95% confidence interval. Exclusion of these funds from my analysis however does not change the results substantially.

## Limitations

This study faces several key limitations that need to be addressed. First, due to the nature of my data, I am unable to separate trading costs and management expenses that are both reflected in the total expenses. Such division would enable me to make a stronger argument about the impact of management fees on the risk-adjusted returns. Nonetheless, the correlation<sup>6</sup> between expenses and turnover, which is directly related to trading costs, is relatively low. Therefore most of the expenses seem attributable to management fees. Consequently my results are mostly indicative of the role of management expenditure in determination of returns.

---

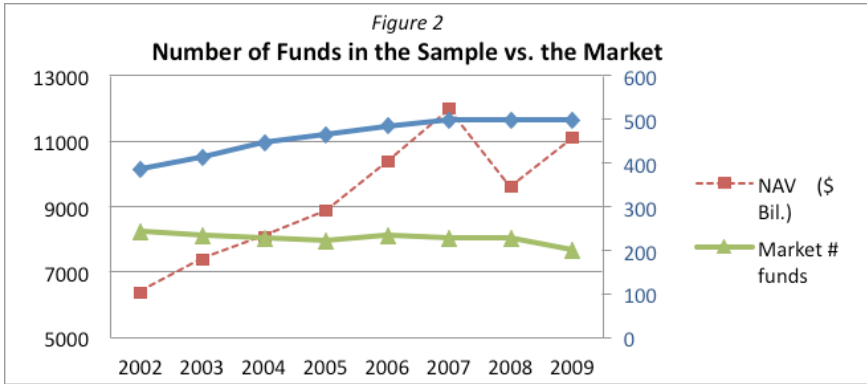
<sup>6</sup> The correlation coefficient is 0.17

Second, my research faces the “black box” problem for I am unable to ascertain changes in portfolio holdings or drifts in the beta of the funds. In particular, I assume that targeted level of risk (beta) and core investment strategy of the fund remained constant during the period. These assumptions are theoretically sound as most funds position themselves as pursuing a certain investment style and strive to retain this image in the eyes of investors. However, if this were not the case, my estimates of the panel data regression would not accurately reflect the true relationships between the dependent and explanatory variables.

Third, my panel dataset is not balanced and suffers from a survivorship bias. In fact, because of the constrained data selection process I can examine only those funds that were still operating in 2011. Thus no funds in my dataset cease their existence during the analyzed period and 114 funds were not yet in business at the beginning of 2002. As Figure 2 indicates the number of funds in the market declined by approximately 6% over the period 2002-2009, while the number of operating funds in my sample actually increased by 28%. This “creation bias” may skew my results. Nevertheless, it seems to have a relatively minor impact on my estimates and so it does not substantially threaten the credibility of my conclusions. In addition, it is interesting to note that the net asset value (NAV) of the mutual fund industry had been increasing at an unprecedented rate until the financial crisis in 2008.

Fourth, this study can be subject to the criticism that the very limited number of time series observations used for estimation of alphas and betas may cause such estimates to be seriously inaccurate. Although a greater number of observations in the regression analysis would certainly be very beneficial, the utilized dataset should provide a good general sense of the size of the true coefficients. Future research should employ quarterly observations for the given period to produce more accurate estimates of the alphas and betas.

Lastly, my analysis faces several econometric problems. Researchers have abundantly documented that using pooled data of this kind is likely to result in positive correlation among residuals. The often complex correlation between the residuals across time and across the industry would have a diminishing impact on the size of the standard errors of the estimates and could occur through two avenues: The error terms may be correlated for a group of funds during a given year (due to the value and small stock effects etc.) or may be serially correlated for a specific fund (high performers may have generally positive residuals) (Ippolito 1989). I address this issue by including dummy variables for years and funds as well as using HAC standard errors.



## Results

First, let us focus on the relative performance of funds across the time period to understand the patterns in their returns. Table 3 provides an overview of the coefficients alpha and beta estimated for each fund using the CAPM model (1). As I expected, the mean beta is close to the beta of market which is a unity. This finding supports my hypothesis H2 that on average the widely diversified funds hold as much risk as the market itself does (H2:  $E(\beta) = 1$ ).

Furthermore, the average alpha of the sample is negative, suggesting that the funds on average slightly underperformed the market on risk adjusted

basis. Nonetheless, a more useful approach to evaluate the validity of this claim is to test whether the individual alphas of the particular funds are lower than zero. The results of this analysis are summarized in Table 4. This table also includes a comparison with the results of Ippolito (1989) and Jensen (1964) who employed a similar methodology in their prominent studies.

**Table3**  
**Overview for the Sample**

	Obs	Mean	St. Dev.	Min	Max
<b>Alpha</b>	500	-0.62	2.56	-11.47	8.51
<b>Beta</b>	500	1.05	0.18	0.49	1.71

**Table4**  
**Analysis of Estimated Alphas:\***

	Zero	Negative	Positive	Total	Mean Alpha	Mean Beta
My results (2002-09)	485	10	5	500	-0.62	1.05
Ippolito (1965-84)	127	4	12	143	0.81	0.88
Jensen (1945-1964)	98	14	3	115	-1.1	0.84

*\*calculated at 95% confidence interval*

**Proportions of Samples:**

	Zero	Negative	Positive
My results (2002-09)	97%	2%	1%
Ippolito (1965-84)	89%	3%	8%
Jensen (1945-1964)	85%	12%	3%

Table 4 clearly indicates that the vast majority, 97% of the sample, neither outperformed nor underperformed the market at a 95% confidence interval. This observation supports my hypothesis H1 that on average mutual funds neither outperform nor underperform the market after expenses (H1:  $E(\alpha) = 0$ ). It is worth noting that two percent of the sample funds underperformed the market significantly while one percent of the funds substantially outperformed the benchmarks. These findings are generally in accord with the results of Ippolito and Jensen, although their proportions of the samples that displayed alphas

indistinguishable from zero are smaller. This fact may be attributable to numerous factors ranging from very dissimilar market and economic conditions involving financial uncertainty to shortcomings of my data.

Additionally, I estimate the mean alphas and turnover ratios by different fund categories to examine potential patterns in the industry. The results of this estimation are reported in table 5.

**Table 5**

<b>Mean Estimated Alphas by Categories:</b>			<b>Mean Turnover by Categories:</b>		
	Growth	Value		Growth	Value
Large Cap	-1.25	-1.01	Large Cap	94.04	58.26
Small Cap	-0.15	1.27	Small Cap	122.23	71.92

The table provides several appealing observations. The largest negative alphas on average were documented in the large cap growth category while small cap value category displayed on average the greatest positive alphas. Since most of the security research and information is readily available in the large cap growth category, it is reasonable to assume that the marginal return on research and information is the lowest in the category. This would be especially true if mutual fund managers who actually carry out the research would be slower to act on certain information than public traders. Conversely, information is generally scarce among small cap stocks and particularly in the small cap value category. Therefore marginal return to research could be the highest in this stock class, enabling mutual fund managers to truly benefit from their security analysis and trading expertise.

Such hypotheses are generally supported by my findings. The large mean alpha in small cap value category indicates that the fund managers in this equity class were able to beat the “uninformed” market after expenses. Generally one can notice that the mean alphas, as indicators of management skill, are on

average lower in the large cap segments than in the small cap categories. This observation might be partially caused by so called “size effect” which states that small cap stocks generally outperform large cap stocks. This anomaly to the CAPM model has been widely documented. French and Fama for instance argue that on average holding small stock enables an investor to capture greater excess returns than holding other asset classes for any given level of risk (1992). The size effect represents a premium that is not associated with management skills. Therefore, the alphas of funds focused on small cap stocks may be overstated. To test this claim I estimate the CAPM model (1) with *SMB* as an explanatory variable. The variable *SMB* is constant for every year and reflects by how much the small cap market portfolios, constructed by French and Fama, outperformed the large cap market portfolios (Kenneth R. French - Data Library). Under such estimation the disparity among the alphas largely disappears and the alphas become generally more negative, except for the category large cap growth, where the mean alpha slightly increases. The results of this estimation are presented in Appendix 3.

Turnover correctly reflects the phenomenon that value investors tend to wait more and trade less than growth investors. The growth investors believe that they can frequently trade on certain information even if it is not fully supported by the fundamentals of the stock (Strong 2004). For these reasons the turnover of growth funds is generally higher than turnover of value funds. Additionally, I would expect the mean turnover on the large cap growth category to be the highest because abundant information that one can trade on is available and because most day trading strategies focus on this asset class. Nevertheless, this is not the case and small cap value segment actually displays the highest mean turnover. This seemingly puzzling fact is not extremely surprising as even the most prominent researchers fail to reach a consensus regarding the relationship between returns, asset classes and fund turnover.



Considering these observations about the actual performance of the funds over the examined period I further analyze the role that expenses and turnover play in determination of these returns. I estimate the panel data regression (2) for each of the fund classes large cap growth (LG), large cap value (LV), small cap growth (SG) and small cap value (SV) as well as for the entire sample. In the estimation for particular fund classes I omit the fund and year dummy variables as most of the dummy variables would be dropped due to collinearity. Nevertheless, for the overall sample I run regressions both with (Overall 1) and without (Overall 2) the dummy variables. Additionally, I employ HAC standard errors in all my estimations because my sample suffers from severe heteroscedasticity. The results are presented in Table 6.

**Table 6**  
**Estimated Coefficients by Categories:**

<i>Dependent variable: Exreturn = <math>R_{jt} - R_{ft}</math></i>						
	LG	LV	SG	SV	Overall 1	Overall 2
BMktRf	1.00** (0.008)	1.00** (0.007)	1.00** (0.012)	1.00** (0.022)	1.00** (0.054)	1.00** (0.006)
Expenses	-1.14** (0.391)	-0.50* (0.291)	0.74 (1.456)	0.00 (1.257)	3.89 (2.819)	0.12 (0.701)
Turnover	0.00 (0.002)	0.00 (0.002)	0.00 (0.002)	0.00 (0.005)	0.00 (0.003)	0.00 (0.001)

\*\* Statistically significant at 95% confidence level

\* Statistically significant at 90% confidence level

In parenthesis: Standard Errors

Overall 1: Dummy variables for years and funds were employed

Overall 2: No dummy variables were employed

Several important observations arise from the results. The coefficient on BMktRf, which represents the CAPM market premium ( $R_m - R_f$ ) multiplied by the previously estimated beta of the fund, is not statistically significant in difference from one. This is exactly what I expected because beta is originally estimated as a coefficient on MktRf. Thus there will likely be a one to one relationship between BMktRf and excess returns.

In the regressions for the entire sample, the coefficients on expenses are statistically insignificant in difference from zero which supports my hypothesis that the impact of expenses is in essence neutral ( $H_3: e = 0$ ). The coefficients obtained from the sub-samples of asset classes are in size similar to the coefficient on estimation with no dummy variables. Therefore, it appears that the relatively high coefficient on expenses in estimation of Overall 1 is purely due to the inclusion of dummy variables for years and funds.

More interestingly, the coefficient  $e$  is negative and statistically significant in difference from zero in the large cap categories while remaining insignificant in the small cap categories. This finding supports the proposed theory that the lack of readily available information in the small cap categories increases the marginal returns on information and consequently enables managers to offset their research and management expenses by resulting higher returns. In contrary, it seems that, in the large cap segment, research and active trading that drive expenses are likely to lower the net returns of the funds, which supports the EMH. These findings are not impacted by the size effect and hold among several model specifications.

Turnover seems to be almost perfectly neutral across the entire sample and the sub-categories because the estimated coefficients are indistinguishable from zero. This would imply that the amount of trading itself does not have a substantial impact on the returns. Additionally, turnover is a proxy for trading expenses that are already incorporated in the expense ratios. Therefore, as I have noted before, one may expect high correlation between the variables turnover and expenses. Nevertheless, the relatively low correlation coefficient of these variables (0.17) indicates that most of the expenses are due to research and management fees rather than trading expenses. In future, it would be beneficial to obtain data for management fees and expenses separately as this would strengthen my inference from the empirical results.

## Conclusion

Combining the analysis of individual fund and market performance, it is clear that most of the funds across different asset classes neither substantially outperformed nor underperformed the market on risk-adjusted basis during the period 2002-2009. I document that although this was true for a majority of the funds, some asset classes yielded greater risk adjusted excess returns than others. In fact, the estimation of individual fund's alphas indicates that funds in the small cap categories outperformed those in the large cap segment of the market on risk adjusted basis. This could be caused by the fact that small stock tends to generate greater returns than other asset classes, for any given level of risk. This so called "size effect" is confirmed by my empirical analysis and artificially inflates the alphas of managers who focus on small cap stocks. Using the French and Fama methodology, I find that the differences in alphas diminish after I account for the size effect. This suggests that, abstracting from the size effect, the managers on average performed comparatively well in all of the categories.

Nevertheless, even after I account for the size effect, the key result indicated by my analysis remains unchanged: Contrary to a popular public view, there does not seem to be a negative relationship between expenses and returns net of fees. The estimated coefficient on expenses that is indistinguishable from zero suggests that the mutual funds on average spend their resources efficiently. In other words fund expenses that generally increase due to research and active management are at least offset by resulting higher returns. This observation holds across all estimations presented in this study and is theoretically justifiable assuming existence of costly information in financial markets.

However, several interesting exceptions arise from the estimations of particular sub-classes. I document a significant negative impact of expenses on

excess returns in the large cap categories, while observing a non-negative effect in the small cap categories. This finding remains unaffected even after accounting for the size effect, suggesting its relative strength.

Therefore, I conclude that, although the semi-strong EMH holds in most cases, the managers focusing on asset classes with low availability of information may experience high returns to information and consequently outperform the market before expenses. In addition, my empirical analysis indicates expenditure efficiency, suggesting that the overall mutual fund market as well as its small cap segment is in a relative informational equilibrium. In such equilibrium the traders carry out just enough research that its marginal cost equals marginal benefit of the information gained.

On the other hand, the large cap class of the mutual fund market does not appear to be in such equilibrium. I find that in this category the increased expenses negatively contribute to the fund returns. Therefore, in large cap, more research and management is unlikely to increase returns. In fact, greater active management is likely to be counterproductive and so I hold that the large cap funds are not spending their resources efficiently. To bring this market segment into equilibrium, rational agents would cut their expenses, decrease the amount of research and human capital they employ or would focus on more profitable market segments. It seems reasonable to believe that the reasons why this has not been the case lie in the problems of imperfect information and bounded rationality. Future research should focus on such differences between the two markets segments and should identify any conditions specific to the large cap funds that could reconcile this disparity.

## Bibliography

- Carhart, Mark. "On Persistence in Mutual Fund Performance ." *The Journal of Finance* 52, no. 1 (1997): 57-82.
- Fama, Eugen , and Kenneth French. "The Capital Asset Pricing Model: Theory and Evidence." *Journal of Economic Perspectives* 18, no. 3 (2004): 25-46.
- Fama, Eugen, and Kenneth French. "The Cross-Section of Expected Stock Return." *The Journal of Finance* XLVII, no. 2 (1992): 247-265.
- French, Kenneth. "Kenneth R. French - Data Library." Tuck - MBA program web server. [http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html) (accessed October 20, 2010).
- Gray, Wesley, and Andrew Kern. "Do Hedge Fund Managers Have Stock Picking Skills?." Social Science Research Network. <http://papers.ssrn.com/abstract=1477586> (accessed September 25, 2010).
- Grinblatt, Mark , and Sheridan Titman. "The Persistence of Mutual Fund Performance." *The Journal of Finance* XLVII, no. 5 (1992): 1977-1984.
- Grossman, Sanford J. and Joseph E. Stiglitz. "On the Impossibility of Informationally Efficient Markets." *American Economic Review*. no 3 (1980): 393-408
- Hendricks, Darryll, Jayendu Patel, and Richard Zeckhauser, "Hot hands in mutual funds: Short-run persistence of performance, 1974-88" Working paper, Harvard University. (1991)
- Henriksson, Roy, D. "Market Timing and Mutual Fund Performance: An Empirical Investigation." *Journal of Business*, 57 (1984), pp. 73-96.
- «Investment Company Institute.» 2010 Factbook. <http://www.icifactbook.org> (accessed September 19, 2010).
- Jensen 1967, Michael. «The Performance of Mutual Funds in the Period 1945-1964 .» *The Journal of Finance* 23, no. 2 (1967): 389-416.
- Karz, Gary. "The Efficient Market Hypothesis." Investor Home . <http://www.investorhome.com/emh.htm> (accessed October 10, 2010).
- Lehmann, Bruce, and David Modest. "Mutual Fund Performance Evaluation: A Comparison of Benchmarks and Benchmark Comparisons." *Journal of Finance* 42 (1987): 233-265.
- Malkiel, Burton. "The Efficient Market Hypothesis and Its Critics." *Journal of Economic Perspectives* 17, no. 1 (2003): 59-82.
- Malkiel, Burton. "Returns from Investing in Equity Mutual Funds 1971 to 1991." *The Journal of Finance* 2 (1995): 549-572.
- Peterson, James, Paul Pietanico, Mark Riepe, and Fran Xu. "Explaining After-Tax Mutual Fund Performance." *The Financial Analysts Journal* 58, no. 1 (2002): 75-86.
- Sharpe, William . "Capital Asset Prices: A theory of Market Equilibrium under Conditions of Risk." *Journal of Finance* 19 (1964): 425-42.
- Sharpe, William. "The Arithmetic of Active Management." *The Financial Analysts' Journal* Vol. 47, (1991): 7-9.
- Strong, Robert A.. *Practical investment management* . 3. ed. Mason, Ohio: Thomson/South-Western, 2004.
- Wermers, Russ. "Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses." *The Journal of Finance* 55, no. 4 (2002): 1655-1695.

**Friedrich von Hayek:**  
**The Socialist-Calculation Debate, Knowledge Arguments,  
And Modern Economic Development**  
*Cara Elliott*

**Introduction**

At the close of the nineteenth and the commencement of the twentieth century, socialism began to gain momentum as a large-scale movement in Europe and the United States. This popularity was supported by an increased influence of the working class in society, which put pressure for representation upon European parliaments and began to secure concrete improvements in labor protection laws (Backhouse, 2002: 269). Moreover, socialist proponents looked hopefully towards the living example of the Soviet Union, which began its socialist experiment in 1917 following the success of the Bolshevik Revolution. Socialism, which found its economic grounding in the legacies of such men as David Ricardo and Karl Marx, tended to encourage a more central and vital role for government intervention in the economy. Thus economists who favored a socialist-oriented change in contemporary societies began to develop theories intended to address such issues as “where, when and how the state should intervene in economic life” (Backhouse, 2002: 269) and how societies might be successfully reorganized so as to be based upon these new precepts.

These developing theories contrasted with those of the opposition contingent of both past and contemporary economists. As a result, a rich discourse of opposing ideologies appeared in the early decades of the twentieth century, coming from such men as Otto Neurath, Henry D. Dickinson, Maurice Dobb, and Oskar Lange on the socialist side, and Ludwig von Mises and Friedrich von Hayek on the opposing. The debates, which focused on such subjects as the role

of the competitive market, central planning boards, prices, and entrepreneurial knowledge, were valuable to the participants not only in terms of their motives, which were often political, but also in that the very nature of the debates continuously challenged the economists and their theories. This encouraged deeper analysis, understanding, and innovation in the theories developed by economists on both sides of the dispute, allowing for stronger and more refined arguments. Oskar Lange himself, when speaking about a challenge made to socialist theories by Ludwig von Mises, wrote, “Socialists have certainly good reason to be grateful to Professor Mises, the great *advocatus diaboli* of their cause. For it was his powerful challenge that forced the socialists to recognize the . . . very existence of . . . a problem [in the system]” (Lange, 1938: 57). The theories that appeared at this time, especially with regards to Friedrich von Hayek’s “Knowledge Arguments,” caused significant intellectual reverberations that continue to have implications in the socialist debate and economic discipline as a whole in recent economic dialogues.

Hayek’s arguments are especially significant in their challenge to the traditional neoclassical conception of static equilibrium and of perfectly informed, uniform actors in a free-market economy that has “reached” this state of equilibrium. Hayek’s contribution is in conceptualizing a dynamic, consistently changing equilibrium that responds to and thus account more effectively the actions of individuals in those economies. This conceptualization allows economists to visualize and analyze market economies in a much more dynamic fashion and is particularly vital with regards to the current state of affairs in free-market societies worldwide. The world is in the midst of a charged atmosphere still experiencing the shocks of the 2008 financial crisis, the essential collapse of countries such as Greece and Ireland last year, and a generally pessimistic attitude about the ability for Western, free-market economies, especially the United States, to continue to

compete with countries experiencing growth at exponential rates, particularly China, a country that continues to follow a socialist model – albeit an evolved, complex one. The significance of Hayek’s arguments will not be in their ability to be any sort of band-aid to solve the Western world’s problems. However, they are also challenging, thought-provoking, and *non-traditional* arguments which could exist in a larger forum of debate and exchange that focuses upon forward-thinking economic theories that could help to revitalize stagnant or struggling free-market economies in the modern world.

This paper will explore, first, the “socialist-calculation debate” of the early decades of the twentieth century between such scholars as Ludwig von Mises, Oskar Lange, and Friedrich von Hayek. One product of that debate was Hayek’s “Knowledge Arguments,” which will be the second topic of focus in the paper. Finally, the third section of the paper will demonstrate different examples of theories developed in various economics circles in recent years, with special attention to those who claim foundation in Hayek’s work. These papers studied in this section include Israel M. Kirzner’s “Entrepreneurial Discovery and the Competitive Market Process: An Austrian Approach” (1997) and Fikret Adaman and Pat Devine’s “On the Economic Theory of Socialism” (1997).

### **The Socialist-Calculation Debate**

The Socialist-Calculation Debate, consisting of a series of arguments on the subject of the “possibility of a rational economic calculation in a socialist economic system” (Adaman and Devine, 1997: 55), opened with Enrico Barone’s 1908 paper “The Ministry of Production in the Collectivist State,” which outlined a mathematical model for a collectivist state, and continued into the 1920s and 1930s with the opposing literature of the Austrian and Socialist schools (Adaman and Devine, 1997: 55).



## German Literature, 1920s

In 1920, Ludwig von Mises began the “German-language” segment of the debate with his “Economic Calculation in the Socialist Commonwealth,” originally published in *Archiv für Sozialwissenschaften 1920*, which challenged a number of proposed socialist models (Mises, 1972: 75-91). Mises sought to respond especially to Otto Neurath, who had written a paper in 1919 outlining a calculation theory based upon the example given by “war economy.” Neurath argued that during wartime, government assumes responsibility for the planning of material distribution, suppressing the market price system characteristic of peacetime. Moreover, production in wartime is not profit-seeking – which “leads to recurrent periods of over-production and unemployment” (Caldwell, 1997: 1859) – but rather works to achieve maximum productive capacity. Neurath asserted that the “central planning” that occurred during war should continue during peace, with the government acting as a central “giant enterprise” (Caldwell, 1997: 1859). He then went so far as to argue that money would be unnecessary in this new society in which “production would be driven by objectively determined needs rather than the search for profits . . . [and] all calculation regarding the appropriate levels of inputs and output could be handled in ‘natural’ physical terms” (Caldwell, 1997: 1859).

Mises, a monetary theorist, especially disagreed with Neurath in terms of his plans for the dissolution of the monetary system, mentioning that, as money serves as a uniform means of exchange across different factors of production, “for the practical purposes of life monetary calculation always suffices. Were we to dispense with it, any economic system of calculation would become absolutely impossible” (Mises, 1972: 79). However, he also hoped to outline an obstacle for the conceptualization of the socialist order with or without the existence of money, so his response to Neurath went on to provide a model in which socialist states’

ownership of capital, or factors of production, was assumed, therefore negating any competitive market for these goods. Mises saw this inclination as making it difficult, or perhaps impossible, for socialist states to assign value to inputs, and, thus, outputs (Caldwell, 1997: 1859). Essentially, without a competitive market, which, through the processes of supply and demand naturally and continuously determines prices and acts as an economic calculator of sorts, Mises argued, “in the socialist commonwealth every economic change becomes an undertaking whose success can be neither appraised in advance nor later retrospectively determined. There is only groping in the dark. Socialism is the abolition of rational economy” (Mises, 1972: 80).

### **English Literature, 1930s**

In the following decade, partly in response to a popular shift towards socialist thought that occurred in Britain as demonstrated by a variety of groups, including political and labor parties as well as academics (Caldwell, 1997: 1859-60), English literature took up the debate. Friedrich von Hayek, who was at the forefront of this movement, built and elaborated upon arguments begun by Ludwig von Mises during the 1920s and formulated a series of critiques of socialism, addressing in his works the arguments of such men as Henry D. Dickinson and Maurice Dobb, and later, Oskar Lange. One of Hayek’s first obstacles was to tackle Dickinson’s proposal of the possibility of mathematical calculation to determine values in a socialist society through the employment of Léon Walras’ system of equations and the utilization of the “auctioneer.” In Dickinson’s plan, a central planning board using this “Walrasian set of equations” would take on the function of the market as the determinant of prices (Caldwell, 1997: 1860). Hayek, in response, outlined a variety of issues with Dickinson’s system: first, the collection and processing of large amounts of necessary information; second,

the difficulty associated with formulating and solving the required equations; and, third, the static nature of such a system and its inability to compete with the natural adaptations of a free market (Caldwell, 1997: 1860).

Hayek continued this trend of underlining the difficulties associated with suggested socialist solutions in his critique of Maurice Dobb's contention that "if consumption decisions were subjected to central control, most of the problems associated with central planning would be alleviated" (Caldwell, 1997: 1861). Hayek rejected this claim, arguing that this change would likely be unacceptable to a society such as Great Britain, or others like it, which was accustomed to a system of consumer freedom. Also in this critique, Hayek addressed "market socialism," which had not been concretely proposed as yet but was in theory an alternative to what its adherents saw as a distorted capitalist system, in which the market was no longer truly free or competitive, but rather dominated by corporations and monopolies (Caldwell, 1997: 1861). By imagining a system in which "managers of monopolized industries [would be] directed to produce so that prices covered marginal costs . . . duplicating the results of competitive equilibrium," (Caldwell, 1997: 1861) Hayek addressed a few issues which, at that point, were rather underdeveloped, but which he would go on to better conceptualize in his later works. These obstacles were the difficulty for a socialist order to replace entrepreneurs acting in a free market, and the issue of managerial incentives in a society that was not profit-oriented (Caldwell, 1997: 1861).

### **The Lange Debates On Market Socialism**

Oskar Lange, a Polish immigrant based in the United States, also wrote on the subject of market socialism; however, he acted as its ardent advocate, entering the discourse during the latter half of the 1930s to respond first to Mises and later to Hayek. In his paper "On the Economic Theory of Socialism," first

published in *Review of Economic Studies* in 1936-7, and again as a book in 1938, Lange's framework proposed a free market for both consumer goods and labor in conjunction with public ownership of capital and other means of production, which he argued would reduce – although not eliminate – social gaps, given the lack of income disparity that originates in private capital ownership (Lange, 1972: 92-3). Lange's model required given prices determined by a Central Planning Board so as to allow for a "subjective equilibrium condition" in which there is a "combination of factors which minimizes average cost, [a level of] output which equalizes marginal cost and the price of the product, and the best allocation of the ultimate productive resources" (Lange, 1972: 97). Lange's model then necessarily rejected Mises' argument that the market must be the determinant of prices, arguing instead that the "parametric function of prices" which occurs in competitive markets is retained in a system of fixed prices (Lange, 1972: 98-9).

Lange's paper also addressed and refuted Hayek's computation argument (in which Hayek outlined the difficulty of computation given the breadth of required information for equations and subsequent issue of *solving* those equations). Lange claimed that "the only equations which would have to be 'solved' would be those of the consumers and the managers of production" whereby, to solve these equations, consumers need only to spend their income and managers of production need only to produce at those levels determined by the equilibrium requirements as previously defined in this section (Lange, 1972: 103). Moreover, those prices required by the managers of production to determine their production levels should be subject to a "trial and error" method of adjustment for price finding, in which the prices are raised or lowered according to whether there is a surplus or shortage of their respective goods. Given their broader knowledge base and attention to the market, Lange believed the Central Planning Board setting the prices and making the necessary adjustments would actually be better suited for

this role than private entrepreneurs, allowing for a shorter process of movement towards equilibrium than occurred the competitive market (Lange, 1972: 103-104).

While Mises never explicitly responded to Lange's challenge, he continued to reject socialist proposals in his writings, focusing on refuting the socialist proposition of a centrally planned mechanism for determining prices. He saw prices as having a very specific determinant; that is, "[prices are] brought about by the interplay of the valuations of all individuals participating in the operation of the market" (Mises, 1966: 331, cf. Caldwell, 1997: 1863). This entrepreneurially-based scope allowed for an adaptive market that socialism simply could not reproduce as "the ever-changing structure of prices that exists within a market system, the messy groping that appears so archaic, ends up being a passably efficient system for revealing relative scarcities" (Caldwell, 1997: 1863). In this system, the entrepreneur is "the essential actor of the piece" (Caldwell, 1997: 1863) reacting to and, in fact, causing changes in the market; for example, "where shortages have existed . . . the resulting price increases [are] driven by entrepreneurs recognizing, in the face of the uncertainty of the real world, the profit opportunities available" (Kirzner, 1997: 70). According to Mises, the competitive market system, which tends to include explicit private property laws and a propensity towards profit-maximizing actions, is a necessary prerequisite to the role of Mises' entrepreneur, therefore, in socialist economies, the equilibrating actions of the entrepreneur would be null and void.

Hayek, like Mises, continued to write arguments against socialist systems, and, in his own response to Lange, not only made use of his earlier addressed objections to the mathematical calculation system proposed by Dickinson, but also outlined new issues, such as the nature of equilibrium and the role of knowledge in the market.

## Hayek's Knowledge Arguments

Hayek's "Knowledge Arguments," a series of papers published in the late 1930s and throughout the 1940s, explored the nature of equilibrium and especially the process by which that state is reached, and, in so doing, issued a challenge to both traditional assumptions about static equilibrium and socialist theorists. His argument focused on the role of knowledge in competitive markets, responding directly to Lange's contention that a Central Planning Board would better determine prices than individuals participating in a naturally fluctuating free market. Many scholars have asserted that the arguments presented in these pieces were "seminal . . . both in development of Hayek's ideas and in [their] implications for the calculation debate" (Caldwell, 1997: 1865). His ideas were first presented in his 1937 "Economics and Knowledge," and were further refined and developed in later works such as his 1945 "The Use of Knowledge in Society."

### "Economics and Knowledge," 1937

"Economics and Knowledge," which was featured in a 1937 issue of *Economica*, sought to outline a more definite and clear concept of equilibrium. Essential to this concept was Hayek's departure from a "pure theory of stationary equilibrium," which neither accounts for change nor for time and also assumes uniform, perfect knowledge among individuals, making the models based upon these premises inapplicable to real world situations. Hayek's study of knowledge, denoted as "data," begins with a distinction between "objective real facts, as the observing economist is supposed to know them," and "subjective [data] as things known to the persons whose behavior we try to explain" (Hayek, 1937: 39). Thus, a state of equilibrium only lasts so long as the "external data correspond to the common experiences of all the members of the society" (Hayek, 1937: 41). Understanding equilibrium in this manner breaks the restraints of stationary models

and allows for equilibrium analysis to be applied to more realistic, “progressive” society. Also important to this line of argument, however, is Hayek’s careful mention that his distinction does not intend to imply that there is not a relationship between subjective data, or individual plans, and objective data, or external facts. Rather, “subjective data of different people [would never] correspond unless they were due to the experience of the same objective facts” (Hayek, 1937: 43).

Furthermore, this correspondence of knowledge and intentions by entrepreneurs in the market as Hayek understands it is an evolutionary process with a consistent tendency towards equilibrium as economic actors “come more and more into agreement . . . [or] become more and more correct” (Hayek, 1937: 44). This tendency, (which does not necessarily ever lead to an absolute state of equilibrium) or the process “by which individual knowledge is changed,” is Hayek’s next subject. He asserts that economists should remember “how little we actually know about the conditions under which an equilibrium will ever be reached” (Hayek, 1937: 48). This limitation derives from a “constancy of the data” as a condition of equilibrium. This constancy does not exist in the real world, as individuals consistently change their expectations and subsequent actions “as they gain experience about the external facts and other people’s actions,” leading to a continuous and seemingly infinite process of changes (Hayek, 1937: 47-8).

While this process is of interest to Hayek, it is in the ensuing analysis of the nature of knowledge – “how much and what sort different individuals possess” – that his most interesting insights come to light (Hayek, 1937: 48). He begins by assuming individually “relevant knowledge,” which, when taken together, allow for a “spontaneous interaction of a number of people, each possessing only bits of knowledge, [to] bring about a state of affairs in which prices correspond to costs” – in other words, a “*Division of Knowledge*” similar in function to the much-studied division of labor (Hayek, 1937: 49, emphasis in the original). This

knowledge, is, again, consistently changing as individuals become aware of new facts. It is a learning process that may occur either completely by accident or through the unexpected results of any executed action. Therefore, “it is only relative to the knowledge which a person is bound to acquire in the course of the carrying out of his original plan and its successive alteration that an equilibrium is likely to be reached” (Hayek, 1937: 51). In other words, Hayek argued that a greater understanding and, moreover, refinement of this “division of knowledge” theory would allow for economists to better comprehend the equilibrating nature of the market. Hayek saw this issue as the

central question of all social sciences, how the combination of fragments of knowledge existing in different minds can bring about results which, if they were to be brought about deliberately, would require a knowledge on the part of the directing mind which no single person can possess. To show that in this sense the spontaneous action of individuals will under conditions which we can define bring about a distribution of resources which can be understood as if it were made according to a single plan, although nobody has planned it, seems to me indeed an answer to the problem which has sometimes been metaphorically described as that of the ‘social mind’ (Hayek, 1937: 52).

Hayek concludes that, economists, then, should attempt to find a way to deliberately put to use all of this knowledge so as to better fathom the evolutionary process of equilibrium and to formulate a model which would, in application, allow for a closer state of absolute equilibrium than present models and states. Hayek continued to refine these theories in the next few years of his career, and thus Hayek’s later work, “The Use of Knowledge in Society,” delves deeper into, and therefore more clearly expresses, the arguments formed in the 1937 paper.

### **“The Use of Knowledge In Society,” 1945**

Published in the fall 1945 issue of *The American Economic Review*, “The Use of Knowledge in Society” begins by asking the question, “What is the problem we wish to solve when we try to construct a rational economic



order?” (Hayek, 1945: 519). Hayek, in contrast to the traditional practice of seeking to solve this economic calculation problem by employing a system of given preferences and perfect knowledge, underscores that a proposed calculation which at least attempts to be applicable to society must make use of the “dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess” (Hayek, 1945: 519). This problem naturally deals with the concept of planning – in the traditional definition of the word, taken to mean the decisions pertaining to the allocation of resources within a society – and, more specifically, the “planners” making those decisions. The debate, then, over the advantages of either central planning, in which there is one plan determined from above, or competition, in which there is decentralized planning by many different individuals acting in the market, should be resolved by answering the question of which system makes better use of this dispersed knowledge. In order to make this determination, a better definition of the different kinds knowledge and their relative significances is required.

Hayek differentiates between scientific knowledge and tacit knowledge, defining the former as “those which we should with greater confidence expect to find in the possession of an authority made up of suitably chosen experts” and the latter as “more likely to be at the disposal of particular individuals . . . the knowledge of the particular circumstances of time and place” (Hayek, 1945: 521). While greater emphasis is generally placed upon scientific knowledge, Hayek argues that tacit knowledge should hold equal import given that “practically every individual has some advantage over all others in that he possesses unique information of which beneficial use might be made” (Hayek, 1945: 522). This information is formed by way of a wide array of sources, such as popular or localized publications or advertisements; isolated interactions between entrepreneurs; or individual desires – all “special knowledge of circumstances of

the fleeting moment not known to others” (Hayek, 1945: 522). Moreover, through the process of acting in the competitive market, knowledge is consistently and continuously changing, constituting an ongoing discovery process. The nature of tacit knowledge, then, is such that it cannot be expressed numerically, making it *impossible* to integrate in a calculation-based socialist order, while at the same time acting as a natural component of a responsive capitalist system (Hayek, 1945: 524).

The second half of Hayek’s 1945 paper focuses upon arguments in support of the price mechanism. Hayek argues that prices function not only a medium of exchange but also as a conveyor of information. They are a “kind of machinery for registering change, or a system of telecommunications which enables individual producers to watch merely the movement of a few pointers . . . in order to adjust their activities to changes of which they may never know more than is reflected in the price movement” (Hayek, 1945: 527). The nature of the price mechanism is one the primary tenets of Hayek’s argument against those who would advocate for “conscious direction,” (central planners or socialist economists) because one of the most miraculous features of the price mechanism is that it has “evolved without design” (Hayek, 1945: 527). In this sense, it is one of many social institutions which could not be instantaneously replicated, or worse, replaced with different systems, because they exist as part of an evolutionary process consisting of many generations of improvements and development in which success is achieved “by building upon habits and institutions which have proved successful in their own sphere and which have in turn become the foundation of the civilization we have built up . . . [having originally] stumbled upon [them] without understanding” (Hayek, 1945: 528).

From Hayek’s perspective, then, the socialist proposal was doomed from the start, because it intended to transform society at its very roots, therefore

undermining and, in fact, nullifying those evolutionary processes Hayek outlined. The contemporary implications of Hayek's "Knowledge Arguments" were therefore not small; he underscored an issue for both mainstream and alternative school proponents by challenging traditional equilibrium assumptions as well as outlining a system in which the applicability and efficiency of socialist models could not compare to competitive market processes which encouraged the best utilization of dispersed knowledge. The exact nature and significance of these implications continue to be debated and built upon in modern economic circles.

### **Modern Implications of the Knowledge Arguments**

The concepts expressed in Hayek's Knowledge Arguments – dispersed knowledge among individuals; the discovery process surrounding that knowledge in a competitive market setting; the nature and role of tacit knowledge; and the function of prices as conveyors of information about the market – have been significant in different ways for different schools of modern economists.

### **The Austrian School**

The modern Austrian school's conceptualization of the market is influenced by both Mises' lesson of the market being driven by profit-oriented entrepreneurs as well as Hayek's contribution of an understanding of the role, nature and continuous augmentation of knowledge in the market (Kirzner, 1997: 67). Additionally, Austrians have learned to appreciate the function of competition with regards Hayek's discovery procedure in that "for the modern Austrian approach, this perception of competition as the dynamic, driving force for discovery in the market process has become central" (Kirzner, 1997: 69). The Austrian School, then, represents a break with standard neoclassical economics in that the Austrians do not see the static equilibrium model that focuses its interest on the *endpoint* as being sufficient to explain or understand what happens

in actual market economies. Moreover, while neoclassical economics no longer tends to adhere to the time-honored assumption of perfect information, it has done comparatively little to address the role of the discovery process as outlined in Hayek's Knowledge Arguments.

In this sense, the Austrian school's work exists as the most orthodox modern application of Hayek's theories, as Hayek's "idea that 'the pure theory of stationary equilibrium' is inadequate as a tool for understanding the workings of a market economy, and that it should be replaced by a view of the market as a competitive-entrepreneurial process for the discovery and coordination of knowledge, has become a central tenet of Austrian thought" (Caldwell, 1997: 1866). The Austrian school is the best representation of Hayek's legacy within modern economic circles, and it is the only school that employs Hayek's arguments together in understanding the market, equilibrium, and the processes behind each of these phenomena.

### **The Socialist School**

The Knowledge Arguments have been significant for modern socialist economists as well, especially for those that have made attempts to reconcile Hayek's arguments with socialist systems. For example, in Fikret Adaman and Pat Devine's "On the Economic Theory of Socialism," printed in a 1997 issue of *New Left Review*, the authors advocate for a system of participatory planning, in which "the values of individuals and collectives interact and shape one another through a process of cooperation and negotiation . . . [enabling] tacit knowledge to be articulated and economic life to be consciously controlled and coordinated" (Adaman and Devine, 1997: 75). Their proposed model differentiates between "market exchange" and "market forces," in which market exchange is taken to mean "transactions between buyers and sellers" and market forces refers to

“the process whereby changes are brought about in the underlying allocation of resources, the relative size of different industries, the geographical distribution of economic activity” (Devine, 1992: 79-80, cf. Adaman and Devine, 1997: 76).

Market exchange, then, encourages information to be generated through the participation of individuals acting in the market, which then makes the best use of existing productive capacity while also determining any necessary changes that will need to be made in the structure of that capacity for better future usage. Market forces, on the other hand, will not be incorporated into the participatory setting, planned instead from above. Through this combination of free participation in the market with planning from above, the authors attempt to express a model in which Hayek’s concept of tacit knowledge can, in fact, be incorporated into a socialist economy. The issue with this model is that Hayek’s concept of an individual acting in the market resembles Mises’ profit-motivated entrepreneur, and so does not tend, as in Adaman and Devine’s model, to “promote cooperation and recognition of interdependent common interest” (Adaman and Devine, 1997: 78), but rather makes choices that are motivated by profit (Kirzner, 1997: 78). This orientation towards profit is the driving force behind Hayek’s discovery process, thus, while Adaman and Devine’s system accounts for tacit knowledge, it still misses some of Hayek’s main points.

Adaman and Devine, however, differ from many other socialist economists in at least seeking to integrate Hayek’s Knowledge Arguments into their models. There have been a variety of other socialist models recently proposed which do not take into account the insights of Hayek’s work, especially ignoring the role of tacit knowledge. One of these models is outlined in Pranab Bardhan and John Roemer’s work, expressed in such papers as their 1992 “Market Socialism: A Case for Rejuvenation,” which focuses on the issues of calculation and motivation instead of knowledge, and, through a “bank-centric

system of insider monitoring” seeks to solve the managerial motivation problem (Bardhan and Roemer, 1992: 105). Echoing Oskar Lange’s claims, their system’s “main bank and the group partners . . . have more “inside” information . . . [and] are likely to be capable of detecting . . . trouble more easily than a diffuse body of stockholders” (Bardhan and Roemer, 1992: 109). In this sense, Bardhan and Roemer outline a model which may seem applicable to real world situations, but, like Lange before them, do not account for the role of knowledge as expressed by Hayek in their work. The differences between Adaman and Devine and Bardhan and Roemer’s theories, which both exist in the realm of market socialist theories, illustrate the nature of the debate that surrounds socialism today, even from within, as socialist economists continue to search for a working model which accounts for their opponents’ challenges.

## **Conclusion**

Friedrich von Hayek’s Knowledge Arguments stand as both some of his most insightful and significant work as well as noteworthy developments in the greater scope of economic thought as a whole. They had implications both within Hayek’s own contemporary economic circles, especially with regards to the socialist-calculation debate, and continue to influence economic theorists today. The propagation of Hayek’s concepts is particularly evident in the “alternative” Austrian school of thought, which, like Hayek, challenges neoclassical standards to move closer to real world situations so as to create applicable, working models for market economies. Also notable, modern socialist economists such as Fikret Adaman and Pat Devine have worked to address Hayek’s theories, attempting to integrate at least some of the ideas into a workable socialist model.

This continuing debate mirrors that of the development of the Knowledge Arguments themselves. Hayek was a key economist among those participating

in the socialist-calculation debate, and, therefore, it follows that the Arguments developed as a product of that debate, in one of his many attempts to challenge his opponents. As Israel Kirzner maintains, the socialist-calculation debate was a “catalyst in the development and articulation of the modern Austrian view as a competitive-entrepreneurial process of discovery . . . it was through the give-and-take of this debate that the Austrians gradually refined their own position” (Kirzner, 1988: 1, cf. Lavoie, 1985, cf. Caldwell, 1997: 1861) In this sense, the evolution of the Knowledge Arguments demonstrates a broader tendency of economists to rely upon debates within and between different circles of thought to better develop their theories and, ultimately, to come to a more complete understanding of the world around them, and, more importantly, apply that understanding with the hopes of improving that world.

## Bibliography

- Adaman, Fikret and Pat Devine, 1997. "On the Economic Theory of Socialism." *New Left Review* 221 (January): 54-80.
- Backhouse, Roger E., 2002. *The Ordinary Business of Life: A History of Economics from the Ancient World to the Twenty-First Century*. Princeton and Oxford: Princeton University Press.
- Bardhan, Pranab and John E. Roemer, 1992. "Market Socialism: A Case for Rejuvenation." *The Journal of Economic Perspectives* 6 (Summer): 101-116.
- Caldwell, Bruce, 1997. "Hayek and Socialism." *Journal of Economic Literature* 35 (December): 1856-1890.
- Hayek, F. A., 1937. "Economics and Knowledge." *Economica* 4 (February): 33-54.
- Hayek, F. A., 1945. "The Use of Knowledge in Society." *The American Economic Review* 35 (September): 519-530.
- Kirzner, Israel M., 1997. "Entrepreneurial Discovery and the Competitive Market Process: An Austrian Approach." *Journal of Economic Literature* 35 (March): 60-85.
- Lange, O., 1938. "On the Economic Theory of Socialism." in Alec Nove and D. M. Nuti (eds.), 1972. *Socialist Economics*. Middlesex: Penguin Books Ltd., 92-110.
- Lange, Oskar and Fred M. Taylor, 1938. *On the Economic Theory of Socialism*. Benjamin E. Lippincott (ed.) London: Oxford University Press.
- Mises, L., 1920. "Economic Calculation in the Socialist Commonwealth." in Alec Nove and D. M. Nuti (eds.), 1972. *Socialist Economics*. Middlesex: Penguin Books Ltd., 75-91.



# **History of Usury**

## **The Transition of Usury Through Ancient Greece, The Rise of Christianity and Islam, And the Expansion of Long-Distance Trade and Capitalism**

*Cheryl Olechnowicz*

### **Introduction to Usury**

Society and its ideas, markets, and institutions are in the constant process of change. These transforming factors contribute to the evolution of economics. Usury is one prominent economic issue that demonstrates this evolution. As it has developed, usury, the lending of money at interest or excessive interest, has been debated for almost two millennia (Visser, 1998, Usury).

During the lifetime of Aristotle, 384-322 B.C., the lending of money for profit was believed to be unnatural and dishonorable (Madra, 2010, Ancient Greece). Aristotle and his beliefs of usury provided a foundation of ideas for future perspectives on the practice. This negative connotation associated with usury continued in history as is evident in the development and spread of Christianity and Islam during the Middle Ages. The Christian church drew on biblical passages and moral and religious reasons to define usury as a sin. The Church placed a ban on the practice of usury to prevent this “evil”. In Islam, the Quran and the teachings of the Prophet Muhammad led Muslims to also view usury as a crime.

As the world has developed, usury has lost its negative connotation in the West and has become a social norm. The Christian church has lifted its ban on usury while a gradual decrease of the importance of religion is seen. Long-distance trade has developed which also contributes to the increasing emergence of usury. The expanse of trade has led to more people being involved in the market and the augmentation of new ideas on usury.

The rise of capitalism too has affected societies' views of usury. Capitalism does incur the making of self-profit as well as rates of interest. Both of these would have been looked down upon in Aristotle's time and the Middle Ages. In the present however, the West has grown accustomed to capitalism. Interest rates do not carry any negative connotation and usury is no longer considered a sin.

The Christian church and Islam both drew on Aristotle's beliefs on usury to help develop their own disapproving views of the practice. However, as time passed, society developed economically and socially and the Church lifted its ban on usury. Islam has developed as well, yet it still continues to view usury as detrimental to society. The debate on usury has witnessed countless arguments over the past two millenniums, and it will continue to perceive them due to varying opinions and the religious passages in the Bible and the Quran.

### **History of Usury**

Aristotle was revered for his contributions to philosophy and economics. His writings and ideas on usury were significant in Ancient Greece and his influence continues to be seen today. Aristotle distinguished between natural and unnatural exchange to define his view on usury. Natural and unnatural exchange is also known as *arête* (the art of being a good citizen) versus *chrematistike* (the acquisition of wealth) (Madra, 2010, Ancient Greece).

The discrepancy between the two types of exchange heavily influenced people's thought in Ancient Greece and the Middle Ages. As time progressed, this difference became less important, and it ultimately contributed to a less critical view of usury. On this distinction of exchange he says the following:

“There are two sorts of wealth-getting, as I have said; one is a part of household management, the other is retail trade: the former necessary and honorable, while that which consists in exchange is justly censured; for it is unnatural, and a mode by which men gain from one another. The most hated sort, and with the

greatest reason, is usury, which makes a gain out of money itself, and not from the natural object of it. For money was intended to be used in exchange, but not to increase at interest. And this term interest, which means the birth of money from money, is applied to the breeding of money because the offspring resembles the parent. Wherefore of a modes of getting wealth this is the most unnatural” (Aristotle, mid 300 B.C., cited from Medema, 2003, Excerpts from Politics).

To understand the quote one must comprehend the Greek *polis*; or an independent city state where the citizens have a large role in their public life (Backhouse, 2002, p. 23). In order to survive and carry out their civic role, the citizens required materials to continue living on their estate. This was termed “household management” which was considered to be perfectly natural exchange. The people did take part in trade, however only for necessary items they could not produce themselves. This is the reason Aristotle terms natural trade “necessary and honorable” (Aristotle, mid 300 B.C., cited from Medema, 2003, Excerpts from Politics).

People involved in this household management therefore had a limit on the natural amount of wealth they could accumulate. The ultimate goal for the Ancient Greeks was to obtain the “good life” which entailed being a citizen of the *polis* (which as explained above entailed household management). Aristotle believed in this good life and urged people to acquire it.

Unnatural trade involved one person benefiting from another, an action viewed as usury. The excessive accumulation of wealth solely for profit was considered abnormal and ethically wrong. One would be acting rationally for their “self interest”. If in doing so, one disregards others, then acting in “self interest” is viewed as wrong.

The Greeks viewed usury as the “most hated sort” of trade (Aristotle, mid 300 B.C., cited from Medema, 2003, Excerpts from Politics). Lending money at a high interest rate was using money to make a profit. This was frowned upon

because money was meant to be used for exchange, not for making more money. Aristotle viewed unnatural exchange as a producer of avarice which led to social problems (Kozel, 2006, p.20) He believed that people obsessed with attaining wealth, would be too preoccupied to participate in the *polis* and fail to perform their civic duties (Kozel, 2006, p. 25).

Aristotelian thought is continued and reflected in the Christian church during the Middle Ages. Religion is combined with Aristotle's ideas to influence economic thought on usury during this time period. Citizens involved in trade questioned whether profit was considered moral. They turned toward the Church to address this problem. They looked at Jesus who had his followers give up all their possessions (Backhouse, 2002, p.33). Saints were respected and followed, yet not as extreme. Some Saints did not believe in owning property, because they did not want people to become obsessed with the accumulation of it (Backhouse, 2002, p.34). This fixation with acquiring wealth has always been one main argument against usury.

The Saints in the Christian church reflect Aristotle's negative views on wealth. St. Paul urged people to give up their worldly possessions (Madra, 2010, Middle Ages). They would not have wealth and they would not be distracted with the goal of accumulating money. St. Augustine argues that "wealth should be a means not an end" (Madra, 2010, Middle Ages). St. Augustine is agreeing with Aristotle that the natural exchange of money is deemed appropriate. People need enough money as a "means" to survive. It should not be an "end" and the only goal in one's life.

The Church and the Saints supported a ban on usury by drawing from Aristotle as well as from the Bible. The following biblical passage swayed many Christians that usury was a sin.

“But love ye your enemies, and do good, and lend, hoping for nothing again; and your reward shall be great, and ye shall be the children of the Highest: He is kind unto the unthankful and to the evil” (Luke vi: 35, cited from Nelson, 1949, p.8).

When Christians heard the word of a disciple speak of lending but “hoping for nothing”, they followed their example and adhered to it. The Church thus placed a ban on usury for these religious and moral reasons.

The Crusades also posed an economic reason for the prohibition of lending with interest. Usurers were seen as taking advantage of profits in “commodity corners and loans on the security of lands” that had been placed on the market by nobles gone to fight (Nelson, 1949, pg.7). With the elimination of these usurers, these profits would then be directed toward the promotion of the crusades instead. For efficient and influential promotion of the Crusades, several Popes required the inhibition of usurers (Nelson, 1949, pg.7). Leading up to the Crusades, discussion on the definition of a usurer had been unclear. The Popes then curtailed all usury to solidify the distinction. These actions let it be clear that one of the motives for the Crusades was the elimination of usury. The Crusades also experienced much land being placed on the market by men fighting in the Holy Wars (Nelson, 1949, pg.7). Usurers then exploited the market, seeking profit from the absence of these warriors. While Christianity opposed usury, Islam was seen holding similar views.

The decline of the Roman Empire was followed by the growth of Islam. The golden age of Islam continued to see religious and Aristotelian influences. Muslims drew on these influences to develop their argument that usury was morally and ethically wrong. The Prophet Muhammad argued that no interest should be required in transactions (Madra, 2010, Middle Ages). Muhammad was as admired and esteemed as Aristotle; and the ideas of the two men were revered.

The Qur'an, similar to the Bible, contains passages advising against the practice of usury. The Muslim word for excess when speaking of usury is *riba*. *Riba* is defined as lending money for interest without any risk to the lender (Jones, 1989, *Islam and Usury*). The following Qur'anic passage addresses the issue of *riba*.

“O you who believe! Eat no *Riba* (usury)” (Jones, 1989, cited from the Qur'an, Al Imran 3:130).

The ethics explained in the Qur'an were not to be reconciled with. In addition to the Qur'an, several Hadith were composed urging against the practice. The Hadith are narrations written describing the words and actions of Muhammad to provide them as a guide on how Muslims should lead their lives (Brown, 2009, p.89).

Muslim scholars emphasize the “consumable nature” of money and how it can lead to the distraction of an individual (Jones, 1989, *Islam and Usury*). The Qur'anic passage supports this view and helps Muslims understand the divine adverseness to usury. Muslims believe that God “permits trade yet forbids usury” (Visser, 1998, *Usury*). In the market individuals can make a profit through determination and efficiency in which a value-creating process occurs. While interest is set, profit is susceptible to change. One must work to guarantee that they receive profit, while with interest one knows the amount that they will receive (Visser, 1998, *Usury*).

Many Muslims view usury as the exploitation of the poor. Making money by abusing an economic relationship with the poor is strongly urged against. In Islamic society they have a Principle of Distributive Equity that its economy aims to maintain (Visser, 1998, *Usury*). Usury prevents this equity from being reached. Usury is viewed as making the wealthy more affluent, and the poor more deprived.

Usury is defined by many as a love for money. However, this passion for wealth is only one defining aspect of usury. Many, including the Church and philosophers, criticize usury not only because it is considered an act of greed, but also because it is delineated by the negative morals that one exhibits when one performs usury. Some philosophers and religious advisors disparage usury because it is the act of taking advantage of others. It is a disruption on individuals when they are consumed with the idea of making money, and they may neglect their other duties to society and their families. In addition to greed, the consequences on individuals due to partaking in usury are reasons in themselves to vilify the act.

### **Changing and Persisting Views on Usury**

History has seen a great deal of change in society and the economy. Throughout the world, expansion has occurred and markets have developed. Change is inevitable and is seen every day. A decrease in the importance of religion, the emergence of long-distance trade, and the development capitalism has had an effect on many people's views on usury.

In the West, an emanation of trade has posed new thoughts on individuals' actions and decisions in the market. The decreasing importance of religion along with decreasing government censorship contributed to more new economic ideas. These emerging ideas influenced peoples' shifting opinions on usury. In the West, usury no longer carries a negative connotation and it is no longer viewed as a sin.

The world has evolved and trade has become more complex and defined. In the current globalized world, making profit is present and abundant in the economy. Making profit off of others is seen in exchange between two individuals and exchange on a global scale. Profit drives the market and keeps society in motion. Although some disapprove of the practice, usury is now a widely accepted social behavior.

Long-distance trade resulted in the creation of industries and commercialism. These then led to towns to support these markets and merchant capitalists became prominent (Hunt, 2002, p12). The number of people involved in the market and in trade has increased. As this number became augmented, people became to accept usury more.

The Church lifted its ban on usury around the 18<sup>th</sup> century as the practice was widely debated. No clear answer on its permissance was found. One philosopher states that while religion has influenced most laws, such as the Christian church and usury, its effects have gradually been “purged away during the past two centuries so that today there is almost nothing is left of them” (Berman, 1974, p.26).

Many view the crisis of religion being present in law as a result from the decrease in self-identification with religion and the Church (Berman, 1974, p.95). A society whose political and religious aspects have no principles of change is believed by some to be a society in danger (Berman, 1974, p.139). As stated previously, change is inevitable and societies must adjust to the evolving times.

As history unfolded, the Church came to realize that usury was economically detrimental to itself. During the Middle Ages, monasteries that existed were capable of lending money (Noonan, 2005, p.131). However, usury was banned and therefore the monasteries did not participate in such practices. As the twelfth century passed economists saw urban churches develop. These churches were also available to lend money (Noonan, 2005, p.131). These economic reasons, combined with an increase in long-distance trade and changing ideas, contributed to the lifting of the usury ban.

The Enlightenment philosophers and the ideas of Adam Smith helped influence a lifting on the ban of usury. In the past acting in one’s own “self-



interest” helped define usury. However, ideas such as Smith’s thoughts on public good and on society shed new light on the debate. Self-interest is argued to be congruent with a flourishing society. Smith’s *Wealth of Nations* describes how a society can prosper even while individuals act in their own self interest (Backhouse, 2002, p.123). While this is possible, he emphasizes that justice must be present in order for society to function properly.

The emergence of capitalism and a free market changed the West’s view on usury. Capitalism led to the break-up of medieval feudalism which was an economic system that supported usury (Hunt, 2002, p.11). Privately owning inputs for production and making a profit became more common and accepted. Capitalism has many definitions, and many individuals began to define it as “honest trade and entrepreneurialism” (Visser, 1998, Usury).

The influence of capitalism was seen as early as the Middle Ages. A subtle shift towards accepting usury is present in loans during this time period. Individuals involved in the market agreed that if the “lender shared in the risk of the venture, the loan was legal” and it was not prohibited (Jones, 1989, Islam and Usury). Laws prohibiting usury rarely intervened with commercial capitalism. Merchants were able to receive a loan if their agreements made them susceptible to risk as well (Jones, 1989, Islam and Usury).

As commercialism became more prominent, a pro-capitalism movement developed in response to a pro-usury movement (Visser, 1998, Usury). The pro-capitalism movement certainly contained more momentum and support. Usury gradually was changing from being viewed as a morally wrong act against others, to being viewed as a more personal action which was not considered ethically wrong.

While the West experienced a major shift in its view on usury, Islam has held fast to its negative connotation on the practice. It has developed with the changing world, but ultimately Islam has not altered its view.

Many Muslims continue to view usury as detrimental to society. Religiously, the Muslims adhere strictly to their sacred texts and the teachings of their Prophets. They do not easily allow modern times to change their opinion on their traditions. Concerning capitalism, the Islamic perspective speaks that “the greatest problem in the capitalist economy is that of the crises and interest which plays a peculiar part in bringing about the crises” (Visser, 1998, Usury).

The evolution of the market economy is unavoidable and the Muslims are seen adjusting to it. To continue being an active participant in the developing world, Muslims have established a new system of Islamic banks. These banks do lend money, however they do not do so usuriously (Jones, 1989, Islam and Usury). The banks are expected to share the risk concerning money with the borrower. Agreements made between the borrower and lenders of the banks do not entail a “predetermined amount over and above principle” (Jones, 1989, Islam and Usury). Money must not be made from money.

In the 1960s, the first modern bank was created in Egypt, and the consecutive three decades have seen great expansion of the system (Visser, 1998, Usury). By doing so, Muslims are attempting to make their national banks function in accord with the teaching of Muhammad and the Qur’an. Muslims claim that their system of banks provides a stable, equitable, and more lucrative, system of lending (Visser, 1998, Usury).

## **Conclusion**

The debate over usury has been intriguing philosophers, economists, and society for the past two millennia. Countless religious, social, and economic reasons are utilized to argue for and against the practice.

Beginning with ancient Greece around 300 B.C., Aristotle is seen emerging with ideas opposing usury. His arguments of natural versus un-natural trade influence people of his time as well as provide the foundation for future arguments. An emphasis is placed on the importance of the Greek *polis* and the citizens' duty and involvement in it. He urges individuals not to be caught up in greed and money making so that they will be able to fulfill their civic duties.

Christianity and Islam emerge during the Middle Ages and draw on Aristotle's teachings. Each respective religion also is seen reviewing sacred texts to support a ban on usury. The Christian church and its Saints explain how usury is morally wrong. They argue how one can be corrupted for exploiting others and being consumed with making a profit. Islam draws on Muhammad's teachings and the Qur'an to prohibit usury. They argue for equity among all; and they frown upon usury as it can take advantage of the poor.

Long-distance trade, capitalism, and a decrease on the emphasis of religion appeared as time passed and the world evolved. More people became involved in the markets and trade became more prominent. New ideas and thoughts on usury emerged as it became a more common practice.

In the West, most of society accepts usury and no longer considers it a sin or a socially unmoral practice. The decline in the importance of religion has contributed to this. The sacred texts and the traditions of the Church have less influence on individuals and their actions in the economy.

Islam however has not experienced such a change of viewpoints on usury. Muslims continue to see the practice as ethically wrong. In response to the evolving world, modern Islamic banks have developed. These banks do not lend money usuriously; and they were created so that all national banks will adhere to Muslim religious law.

The teachings of Aristotle and the sacred texts of both Islam and Christianity will always be available for discussion. Philosophers and economists will draw from these to argue for and against the practice of lending money at interest. While it is important to understand that usury will always be disputed, it is also essential to comprehend that the world is susceptible to change and that adjustments can be made accordingly.

## **Bibliography**

Aristotle, mid 300 B.C. "Excerpts from *Politics*." In Steven G. Medema and Warren J. Samuels, eds., 2003. *The History of Economic Thought: A Reader*. London and New York: Routledge, pp. 5-13

Backhouse, Roger E., 2002. *The Ordinary Business of Life (OBL): A History of Economics from the Ancient world to the Twenty-First Century*. Princeton and Oxford: Princeton University Press.

Berman, Harold J., 1974. *The Interaction of Law and Religion*. Abingdon Press

Brown, Daniel W., 2009. *A New Introduction to Islam*, Second Edition. Wiley-Blackwell

Hunt, E.K. 2002. *History of Economic Thought, A Critical Perspective*. Updated Second Edition. M. E. Sharpe, Inc.

Jones, L. Norman, 1989. *God and the Moneylenders: Usury and Law in Early Modern England*. Oxford: Basil Blackwell.

Kozel, Philip, 2006. "Aristotle's Discourse on Commodity Exchange." In *Market Sense: Toward A new Economics of Markets and Society*. New York and London: Routledge, pp. 17-30

Madra, M. Yahya (2010) *The Ancient Greece and why Aristotle causes problems for Contemporary Economists*. Notes of History of Economic Thought and Analysis

Madra, M. Yahya (2010) *The Middle Ages: Economic thought in the Service of Religion*. Notes of History of Economic Thought and Analysis

Nelson, Benjamin, 1949. *The Idea of Usury, from Tribal Brotherhood to Universal Brotherhood*. Princeton, Princeton University Press.

Noonan, John, 2005. *A Church that can and cannot Change: the Development of Catholic Moral Teaching*. Notre Dame, Ind: University of Notre Dame Press

Visser, Wayne A.M.; McIntosh, Alastair, 1998. *History of Usury Prohibition*. Accounting, Business and Financial History, Routledge, London