New York State Economics Association

Proceedings of the New York State Economics Association

Volume 3

63rd Annual Meeting
Rochester, New York
September 24-25, 2010

Editor: Richard Vogel

New York State Economics Association
Founded 1948

2009-2010 Officers

President:
Jeffrey Wagner, Rochester Institute of Technology

Vice-President:
Richard Vogel, SUNY-Farmingdale

Secretary:
Michael McAvoy, SUNY-Oneonta

Treasurer:
David Ring, SUNY-Oneonta

Editor (New York State Economic Review):
William P. O'Dea, SUNY-Oneonta

Managing Editor (New York State Economic Review):
John W. Piccione, Siena College

Editor (New York State Economic Proceedings):
Richard Vogel, SUNY-Farmingdale

Web Coordinator:
Wade L. Thomas, SUNY-Oneonta

Board of Directors:
Cynthia Bansak, Saint Lawrence University
Darius Conger, Ithaca College
Robert P. Culp, Dalton State College
Richard Dietz, Federal Reserve Bank of New York – Buffalo Branch
Arindam Mandal, Siena College
Patrick Meister, Ithaca College
F. Pan Shu, SUNY-Potsdam
Della Lee Sue, Marist College
Xu Zhang, SUNY-Farmingdale

Published, June 2011.

Table of Contents

Determining the Causes of the Recent U.S. Recessions and the Economic Slowdown in China
Suprabha Baniya and Florence P. Shu 4

Racial Differences in the Intensity of Breast Cancer Treatment
Emma Bojinova 10

The Effect of Strict Alcohol Policies on America's Sexually Transmitted Disease Rates
Karolyn Caprara *19*

Is Small Beautiful? Size-Productivity Relationship in Indian Manufacturing
Prabal K. De and Priya Nagaraj 28

Capital Flows to Russia, Ukraine, and Belarus: Does "Hot" Money Respond Differently to Macroeconomic Shocks?
Scott W. Hegerty 38

The Declining Exchange Rate: Impact on the U.S. Economy 2000-2009
John J. Heim 49

How much does "Crowd Out" Reduce the Effects of Keynesian Stimulus
John J. Heim 58

Macroeconomic Policy and Institutional Change in the Age of Limits to Growth
Kent Klitgaard 66

Job and Worker Flows: Evidence from New York State Counties
Arindam Mandal 78

The Political Economy of Education Reform: The Case of A Nation at Risk
Johanna Mitchell 89

United States Monetary Policy during the Financial Crisis, Credit Facilities Development and Potential
Marwan El Nasser and Richard Robinson 105

A Chair's Guide to Student Evaluations of Teaching
William P. O'Dea 115

Determining Optimal Firm and Consumer Research and Development Spending in the Medical Technology Sector
Sameer J. Shah …… 126

The Far Reaching Effect of Ground Water Reductions
Anthony Signorelli …… 136

Bankruptcy Prediction Using Neuro Fuzzy: An Application in Turkish Banks
Birol Yildiz and Soner Akkoc …… 142

The Effects of Structural Deregulation on State Banks in New York during the 1970s
Michelle Zagardo …… 154

Determining the Causes of the Recent U.S. Recessions and the Economic Slowdown in China

Suprabha Baniya* **and Florence P. Shu†**

ABSTRACT

Variables suspected as having possibly triggered the recent U.S. recession that begins from December, 2007 to have reached its trough in June, 2010 are mostly financial (NBER, 2010). These variables are mostly suspects by the short-term account. For instance, it is reported that the breakdowns in key credit markets -- such as bond and equity finance markets (Duca & Luttrell, 2010a: Duca, DiMarino, and Renier, 2010b) first occurred. In this thinking, heavy homebuilding finance and housing-backed consumer borrowing defaults trailed the recovery from the 2001 recession (Duca and Luttrell, 2010a; Duca, DiMarino, and Renier, 2010b) were cited by the most as dual major drivers behind the recent U.S. recession. This paper examines the causes of the recent U.S. recession by investigating the relationships between U.S. public debt-to-GDP, U.S. real estate mortgage delinquency, U.S. income inequality, and U.S. GDP level. Our findings based on annual data between 1987 to 2009 shows that to U.S. GDP performance, U.S. public debt-to-GDP shows statistically insignificant indirect association whereas U.S. income inequality and real estate mortgage delinquency proxies show statistically significant direct association in a multivariate linear regression analysis. Co-integration tests upset these finding results. To U.S. GDP in difference at the first order, all the three cause variables, in respective difference at the first order, shows indirect association. The difference of the U.S. public debt-to-GDP at the first order shows statistically significant indirect association, whereas the difference of the annual income inequality and the annual real estate mortgage delinquency proxies at the first order show statistically insignificant association. Despite short-term perspectives make blame on financial accounts, annual data between 1987 and 2009, suggests that U.S. public debt-to-GDP growth is on the other hand the major statistically significant factor of the recent U.S. recession.

INTRODUCTION

In this study, U.S. debt-to-GDP ratio, U.S. income inequality, and U.S. real estate mortgage delinquency are examined as causes to the recent U.S. recession which concurrently led to slight economic slowdown in China. Our study is aimed at finding out whether each of the three selected cause variables plays a role in the recent U.S. GDP downturn. U.S. GDP is often used as an indicator for the level of the U.S. economic activities. The three variables we examine as candidates for the causes to the recent U.S. recession provide information about how each is associated with the recent U.S. economic slowdown, taking a longer-term perspective using the annual data between 1987 and 2009.

* Departments of Economics and Department of Mathematics (BA/MA degree program), class of 2012, SUNY Potsdam, Potsdam, NY 13676.

† Associate Professor of Economics, Department of Economics, SUNY Potsdam, Potsdam, NY 13676 (Correspondence via e-mail:shufp@potsdam.edu)

The discussion given to this paper is carried out in the following order. First, we discuss theories. Then, we discuss the data, its sources, and two models to be used to help us decide the causaul relations between each of the identified variables and the recent U.S. GDP level. Third, we report our regression findings. This is followed by a section of the brief discussion about the recent U.S. recession and its impact on economic slowdown in China. The last is a section of conclusion. No sophisticated econometric methods are applied in this study to examine the cause-effect relation between the recent U.S. recession and the recent economic slowdown in China.

THEORIES

Below, some theories are introduced for explaining relations between the GDP of a nation and its debt-to-GDP ratio, its income inequality proxy, and/or its real estate mortgage delinquency rate.

Recent studies on the sustainability of fiscal policy to selected Euro-area countries, including Italy, France, Germany, and Portugal, shows sustainability evidenced for fiscal policy, though the Maastricht treaty three percent restriction on debt to GDP ratio may be temporarily violated (Greiner, Köller, and Semmler, 2007; Greiner, Semmler, and Gong, 2005). The studies suggest that the primary surplus to GDP ratio if evidenced as positive function of the debt to GDP ratio, then sustainability of the fiscal policy can be expected. The primary surplus to GDP ratio evidenced as positive function of the debt to GDP ratio reflects tax-base GDP growth leading to the sustainability of fiscal policy. Tax-base GDP growth when occurred with higher debt-to-GDP ratio implies debt produces positive economic stimulation effect leading therefore not to recession..

In an endogenous growth model with elastic labor supply, where agents differ in initial endowments of physical capital, García-Peñalosa and Turnovsky (2007) theorizes that the growth rate and the distribution of income are jointly determined. García-Peñalosa and Turnovsky (2007) also show that policies, especially those ones used in financing an investment subsidy, aiming at increasing the growth rate may result in a more unequal distribution of pre-tax income. Higher income inequality reported by the recent empirical literature may be by theory unfavorable to the growth rate of a consumption-driven economy. Despite families that first encountering liquidity constraints may have shortened multi-period planning horizons, their consumptions are not necessarily restrained by disposable income [Mariger (1987)]. The factor is important to be examined, for that a consumption-driven growth may occur as income inequality worsens and liquidity constrained consumption become increasingly financed by debt.

Sub-prime mortgage default and foreclosures triggered widening of the Libor-IOS spread is a feature unique of the recent U.S. recession (Duca and Luttrell, 2010a; Duca, DiMartino, and Renier, 2010b). Importance of the mortgage default to the pre-U.S. recession financial crisis has been widely commented by government research staff (Economic Report of the President, 2010; Duca and Luttrell, 2010a; Duca, DiMartino, and Renier, 2010b); and scholars Caballero, Farhi and Gourinchas 2008).

Based on aforementioned reasoning, we formalize a simple multivariate regression model aiming at examining whether our regression outcomes based on this simple model may shed lights to help determine causes of the recent U.S. recession. Below is the organized linear regression model.

$US\text{-}GDP_t = a + b1^* DGDPR_t + b2^* (I^*I)_t + b3^* MDR_t + u_t$ (Model I)

In the model, US-GDP stands for either U.S. annual or quarterly GDP in current or constant dollars. (I*I) stands for an U.S. yearly income inequality index, DGDPR stands for the U.S. annual debt-GDP ratio, MDR is U.S. annual mortgage delinquency rate, a stands for intercept of the regression, b1, b2, and b3 are regression coefficients with respect to variables DGDPR, I*I, and MDR, u is a stochastic disturbance term in linear regression Model I, and t is a time descriptor.

We begin with this simple linear regression assumption in Model I, then change this model accordingly when regression results suggest to do is necessary and beneficial. Otherwise, Model is proposed as behaving in agreement with the normality assumed in the classical linear regression model.

DATA AND ECONOMETRIC MODELS

Gathered from the BEA (Bureau of Economic Analysis) are the following data: The time-series U.S. annual GDP in current and constant dollars; and the time-series U.S. quarterly GDP in current and constant dollars. Yearly U.S. GINI index is used as U.S. income inequality proxy. Gathered from the Economic Report of the President is the annual U.S. Federal debt-to-GDP ratio. Annual U.S. real estate mortgage delinquency rates are also gathered to be used as the pre-financial crisis indicator proxy.

In this paper, we first use the ordinary least squares method to obtain regression estimation for Model I. The model later is refined to Model II, with details given below for points to be refined.

REGRESSION FINDINGS

The estimated linear regression results of Model I is illustrated in Table I, where the time-serial U.S. annual GDP in million of 2005 dollars is used as a regressor, annual U.S. debt-to-GDP ratio, U.S. yearly GINI index, and U.S. annual real estate mortgage delinquency rate are used as regressands. The estimated regression results, reported in the second column, Table I, show statistical significance for intercept and for the regression coefficient of variable I*I at the, two-tailed, level of the significance at 1%. However, the regression coefficients of both the variables DGDPR and MDR do not show such a statistical significance. U.S. annual GDP in 2005 dollars is positively correlated to U.S. yearly GINI index and U.S. I*I. Using the estimated results, we cannot generate a prediction for the recent U.S. recession. Our prediction merely on the base of magnitude of I*I from 2007 to 2009 shows wider income inequality should have brought increases to U.S. GDP during these years. This direction of the prediction is reverted because the negative estimate of the regression coefficient for DGDPR (or, U.S. debt-to-GDP ratio). From 2007 to 2009, high U.S. debt-to-GDP ratio reduces the level of U.S. annual GDP in 2005 dollars, based on the regression outcomes of Model I. Positive estimate obtained for the regression coefficient of variable MDR (or, U.S. annual real estate mortgage delinquency rate) suggests that U.S. annual GDP

should have been rising, which actually dropped, between 2007 to 2009. We suspect that multicollinearity may have caused the estimated regression coefficients to be significant for variable I*I and insignificant for both the regressands DGDPR and MDR.

Tabel 1: Regression Results

Variables/Test Statistics	**Estimated Regression Coefficients (Linear Regression Model I)**	**Estimated Regression Coefficients (Model II)**
Constant Term	-52,555	330.4646
	(-5.86476**)	(1783454**)
DGDPR (Model I);	-50.1163	-41.4793
DGDPRD (Model II)	(-1.06555)	(-6.96606**)
I*I (Model I)	142,712.4	-2436.72
or I*ID (Model II)	(6.130013**)	(-.93262)
MDR (Model I);	296.9651	-22.9315
MDRD (Model II)	(1.960401)	(-1.30447)
R-Square	0.808027	0.847676
Adjusted R-Square	0.777715	0.822289
F-statistic Significance	5.1E-07	1.45 E-07
t-critical value at $\alpha/2 = 0.025$	2.093	--
Degree of freedom	19	--

Note: t-statistics of the regression coefficients are reported in the parenthesis under each regression coefficient estimate . t-statistics of the regression coefficients that are denoted with ** are significant at $\alpha/2 = 0.005$. t-statistics with * are tested statistically significant at $\alpha/2 = 0.025$.

Recall that U.S. annual GDP in 2005 dollars as well as all the regressands used in Model I are time series data. Whether these variables each contains unit root is then considered. Our findings point out that variable I*I shows as a case with unit root. The other regressands do not present to be as cases with unit root. Considering this finding, regression Model II is organized after the several unit root tests.

$US\text{-}GDPD_t = a^d + b1^{d*} DGDPRD_t + b2^{d*} (I*ID)_t + b3^{d*} MDRD_t + u^d_t$ (Model II)

For Model II, the regression results are reported in column three, Table I. These results, still in the process of being further examined for mathematical and statistical plausibility, already suggest that DGDPRD provides explanation to a very large part of GDPD in the recent U.S. recession.

When we use the causal analysis suggested by Granger (Studenmund, 2006) to identify if that DGDPRD is a cause to how U.S. GDPD performed, or, that U.S. GDPD is a cause to how DGDPRD performed, our findings suggest that U.S. GDPD leads how DGDPRD performed the next period of the time. From Model II regression results, it is clear that when DGDPRD is higher, U.S. GDPD decreases the next period of the time. And, higher U.S. GDPD is associated with much more rapid decrease in DGDPRD than the decrease in U.S. GDPD the next period of the time. This finding helps us to conclude that U.S. GDPD leads DGDPRD in direction of moves. Strong indirect DGDPD correlation with U.S. GDPD well explains how U.S. GDPD varies as a result of the changes in DGDPD in the recent years. Recent high DGDPD, led by increases in U.S. GDPD between 2000 and 2007 well explains occurrences in the recent U.S. recession beginning November, 2007 lasting until the present. Since U.S. GDPD led

changes in DGDPD, we suggest that US GDP growth by fiscal or monetary expansionary stimulation are keys to the rebound of the near future U.S. GDP performance.

We now turn to the next section, where a brief discussion is given about the recent U.S. recession and economic slowdown in China. We now turn to the next section, where the recent U.S. recession and economic slowdown in China is the subject of the study.

RECENT U.S RECESSION AND ECONOMIC SLOWDOWN IN CHINA

Literature in both politics (such as Drezner, 2009) and economics (such as Caballero, Farhi, and Gourichas, 2008) have led readers to dismiss the articulated growing concerns about U.S. dependence on China and other authoritarian capitalist states as a source of credit to fund the United States' trade and budget deficits. Furthermore, our study finds that the recent U.S. recession indeed turned to become a damper on the rapid economic growth in China. However, regression results, ideally analyzed using monthly if not daily data, are not possibly carried out due to limited accessibility to the Chinese data. Our findings based on quarterly data between 2008 and 2009 are in need of being revised using monthly data when it becomes accessible to us.

CONCLUSIONS

In their article published in the *Brookings Papers on Economic Activity*, Caballero, Farhi, and Gourichas (2008) considers that global asset scarcity led large capital flows toward the United States as origins of asset bubbles that eventually burst into the recent financial crisis. Shortage of assets in the world economy triggered a partial re-creation of the bubble in commodities market, and oil markets in particular. This first phase of the financial bubble is represented by an increase in petrodollars seeking financial assets in the United States. As U.S. real economy and global growth slowed down, the slowdown worked to reverse the tight commodity market conditions then destroyed the commodity bubble. After flowing into U.S. economy, a housing bubble not later sustained by real economic growth led to the recent U.S. recession. However, this recession is not as bad as it appears in duration, because the past record shows that U.S. GDP performance when excellent always led to reduction in U.S. public debt to GDP ratio. U.S. public debt to GDP ratio has not dropped, as a result of the recent U.S. economic slowdown. Literature indicating that the current recession suffering from the mortgage delinquency finds no parallel evidence to support the assertion that mortgage delinquency always led to U.S. GDP slowdown. We strongly believe that a real economic growth in response to macroeconomic stimulation remains to be the key to the recovery from this recent recession. U.S. GDP growth in the past often leads to lower U.S. public debt to GDP ratio, lower mortgage delinquency ratio, and even a better GINI index performance.

REFERENCES

Caballero, R. J., E. Farhi and P.-O. Gourinchas. 2008. Financial Crash, Commodity Prices, and Global Imbalances. *Brookings Papers on Economic Activity*, 2008: 1-55.

Drezner, D. W. 2009. “Bad Debts: Assessing China's Financial Influence in Great Power Politics.” *International Security*, 34: 7-45.

Duca, J. V. and D. Luttrell. 2010a. “Recovering from Housing and Financial Crisis.” *Economic Letter: Insights from the Federal Reserve Bank of Dallas*, 5(7): 1-4.

Duca, J. V., D. DiMartino, and J. J. Renier. 2010b. “Fed Confronts Financial Crisis by Expanding Its Role as Lender of Last Resort.” *Economic Letter: Insights from the Federal Reserve Bank of Dallas*, 4(2): 1-11.

Economic Report of the President. 2010. U.S. Publishing Co.

García-Peñalosa, C. and S. J. Turnovsky. 2007. “Growth, Income Inequality, and Fiscal Policy: What Are the Relevant Trade-Offs?” *Journal of Money, Credit and Banking*, 39: 369-394.

Greiner, A., U. Köller, and W. Semmler. 2007. “Debt Sustainability in the European Monetary Union: Theory and Empirical Evidence for Selected Countries.” *Oxford Economic Papers*, 59: 194-218.

Greiner, A., Semmler, W., and Gong, G. 2005. *The Forces of Economic Growth: A Time Series Perspective.* Princeton Unversity Press.

Mariger, R. P. 1987. “A Life-Cycle Consumption Model with Liquidity Constraints: Theory and Empirical Results.” *Econometrica*, 55: 533-557.

NBER's Business Cycle Dating Committee. September 20th, 2010 Report. http://www.nber.org/.

Studenmund, A. H. 2006. *Using Econometrics: A Practical Guide*. (Pearson/Addison-Wesley)/Prentice Hall: 5th edition.

Racial Differences in the Intensity of Breast Cancer Treatment

Emma Bojinova*

ABSTRACT

Breast cancer has the highest mortality rate among females diagnosed with cancer in the United States. Racial discrimination along with the stigma of inferiority of the minority groups can adversely affect health. The goal of this paper is to investigate if there are racial differences in the intensity of treatment of breast cancer patients. Based on cross-sectional data from the Healthcare Cost and Utilization Project, different specifications are used to estimate if African-Americans, Hispanics, and other racial groups are treated less intensively in comparison with white Americans when they are admitted to hospitals in 35 states.

INTRODUCTION

Cancer is the second leading cause of death for Americans and breast cancer has the highest mortality rate among females diagnosed with cancer. Racial discrimination, both individual and institutional, along with the stigma of inferiority of the minority groups, can adversely affect health. Hence, it is important to understand whether there is racial discrimination in the intensity of treatment, which can explain the higher mortality rate for African-Americans, or if there are other factors that account for this disparity. Thus, the purpose of this paper is to investigate whether there are racial differences in the intensity of treatment for breast cancer, which is proxied by hospital expenditures recorded for an inpatient claim.

Based on cross-sectional data from the Healthcare Cost and Utilization Project (HCUP) for year 2002, different specifications are used to estimate if black and other racial groups such as Hispanics, Asians, Native Americans, and others are treated less intensively in comparison with white Americans when they are admitted to hospitals in 35 states. The results suggest that there is some evidence for the presence of racial discrimination in this particular year.

BACKGROUND AND PREVIOUS LITERATURE

According to the Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute the age-adjusted breast cancer incidence rate for the period 2002-2006 was 123.8 per 100,000 women per year. The breakdown by race is as follows: the incidence rate is highest for white females (127.8 per 100,000 women), followed by black females (117.7 per 100,000), Asian/Pacific Islanders (89.5 per 100,000), and Hispanics (88.3 per 100,000 women). The lowest incidence rate is recorded for American Indian/Alaska Native females (74.4 per 100,000). In contrast, black females have

* Department of Economics and Finance, Canisius College, 2001 Main St., Buffalo, NY 14208.
Phone: (716) 888-2721. E-mail: bojinove@canisius.edu

the highest age-adjusted mortality rate (33 per 100,000) as compared to other races for the same period. Breast cancer mortality rate is approximately 38.1% lower for white women (23.9 per 100,000) and between 87.5% and 264% lower for the other 3 racial groups mentioned above.

The early detection of breast cancer is very important for a patient's survival. Mammography is especially useful for identifying breast cancer at an early stage even before physical symptoms develop. Many studies have shown that early detection increases treatment options and decreases mortality. However, the decision to undergo breast cancer screening depends on whether the person has health insurance, as well as on the educational level and awareness of breast cancer symptoms. For instance, women who lack health insurance, who are poor, less educated, or live without a husband tend to have the lowest prevalence of mammography use due to their limited access to health care. According to the National Center for Health statistics, African-American, Hispanic, and American Indian women are more likely to be diagnosed with breast cancer at a later stage of disease development, which will affect their hospital expenditures.

Furthermore, these expenditures depend not only on the clinical status of patients, but also on the duration of stay, reason for admission, and whether this is first admission or one of the followed re-admissions to a hospital. For example, a woman in an advanced stage of breast cancer during her first admission to a hospital will undergo more diagnostic and therapeutic procedures and as a result her expenditures will be high as compared to say her third re-admission when she can have chemotherapy or radiation therapy only. A number of studies have shown that there is a direct relationship between length of stay and hospital charges but as the length of stay decreases expenditures decrease less than proportionately because the latter is associated with higher intensity of treatment during early stay in hospital. Also, for cancer patients it was found that the "cost of treatment may decrease with severity because of the futility of any further active intervention, while at the same time mortality rate goes up for each stage and substage" (Medstat Disease Staging Software Reference Guide, HCUP 2002).

A lot of research has been done in explaining the differences in cancer survival rates among different socioeconomic groups. Some of the studies considered several cancer sites, whereas others concentrated on a single cancer site. A study by Kravdal (2000) found that the excess mortality was about 15 percent lower for patients who had a post-secondary education as compared to those with compulsory schooling after controlling for age, stage at the time of diagnosis, and registered differences in tumor characteristics. However, Kravdal did not find clear indication that host factors such as comorbidities and immune functions, as well as treatment and care differences matter for the differential in survival rates. Majority of breast cancer survival studies found some evidence suggesting that socially advantaged have better survival rates after controlling for possibly earlier detection of the disease among people corresponding to higher social classes (see for example LeMarchand et al., 1984; Bassett and Krieger, 1986; Karjalainen and Pukkala, 1990; Gordon et al., 1992; Ansell et al., 1993; Schrijvers et al., 1995). Figueroa and Breen (1995) analyzed cases of breast and cervical cancer diagnosed in the period 1989-1990 in San Francisco, Detroit, and Atlanta. They found that 87% of the breast cancer cases were

diagnosed late, when already tumor was malignant. They found that a significant part of the variation in diagnostic stage was explained by the residence in underclass area. The likelihood of late-stage diagnosis also increased with age and was higher for females living without a spouse. Katz and Hoffer (1994) found similar results for breast cancer patients living in Ontario, Canada.

Several studies found evidence that health insurance matters with regards to breast cancer screening, surgical procedures or other treatment procedures. Mitchell and Hadley (1997) considered nonelderly women diagnosed with breast cancer using hospital inpatient discharges for 1988 and 1991 in five states (CA, MD, MA, NJ, and NY). The authors found that the probability of breast-conserving surgery is 2.7% lower for females enrolled in HMOs, 4.8% lower for Medicaid and 6.6% lower for self-pay patients as compared to females having private insurance plans. A study by Thorpe and Howard (2003) found substantial differences in cancer spending by insurance status based on the Medical Expenditure Panel Survey for 1996-1999. They considered 5 big cancer types, among which is breast cancer. Their results showed that nonelderly cancer patients without health insurance have higher risk of being inadequately treated especially if they are of Hispanic origin.

In addition, there are a number of studies based on HCUP data, some of them were concentrated on the relationship between hospital volume and survival, and others looked at racial discrimination. Bach et al. (2001) studied patients 65 years old or older diagnosed with lung cancer between 1985 and 1996 and had surgery at a hospital that participated in the Nationwide Inpatient Sample (76 hospitals). Their results suggest that “patients who undergo resection for lung cancer at hospitals that perform large numbers of such procedures are likely to survive longer than patients who have such surgery at hospitals with a low volume of lung-resection procedures.” Dimick et al. (2003) found similar results for patients undergoing surgery for colorectal cancer based on 1997 HCUP data using logistic regressions. Andrews and Elixhauser (2000) examined whether there is difference in the rate of receiving therapeutic procedures between Hispanic and white non-Hispanic patients based on 1993 discharge data for California, Florida and New York (states that account for half of the Hispanic population in the USA). Their findings provide evidence that Hispanics are undertreated in a sense that they are less likely to receive major therapeutic procedures for 38% of the 63 conditions they examined and more likely for 6% of the conditions as compared to non-Hispanic whites. There are a number of other HCUP studies looking at racial or sex disparities for patients diagnosed with various diseases including other types of cancer (Ball and Elixhauser, 1996; Harris et al, 1997; Andrews and Elixhauser, 2000; Shenn, 2002). However, according to my knowledge there are no studies looking at racial differences in the intensity of treatment for breast cancer patients based on HCUP data. Therefore, my paper can be considered as a contribution in this less researched area.

DATA DESCRIPTION

The source of data is the Healthcare Cost and Utilization Project (HCUP) for year 2002. The nationwide inpatient sample (NIS) consists of approximately 7.85 million hospital stays from about 1,000

hospitals in the United States. It covers 35 states and is designed to approximate a twenty-percent sample of the U.S. community hospitals, which allows for making inferences for the country as a whole. The advantages of using HCUP data is the availability of large number of inpatient records, the good data on health insurance, hospital characteristics, and different disease diagnoses. It should be noted that the unit of observation in this data set is an inpatient claims record not the patient itself. Thus, there is a possibility that the same individual went to a hospital for treatment multiple times in a given year which can affect the estimates. The data is also censored because we do not observe the whole universe (population) - just the individuals that go to a hospital and are treated file a claim.

I concentrate my research on breast cancer inpatient stays and restrict the HCUP sample to discharges with principal diagnosis “breast cancer” (based on ICD-9-CM codes), which reduced the sample to 22,678 observations. Observations with missing values for the variables of interests, i.e. race and total charges, are deleted. There are 156 observations for breast cancer patients that are male, which are dropped from the sample to avoid potential unobserved gender differences with regards to treatment and disease development. Furthermore, Georgia does not report race due to confidentiality of reports. There is also missing race data for some of the other states. As a result, after deleting these observations the sample size decreases by 6,084 inpatient records.

The following variables from the NIS are used in the study: total charges (*totchg*), median household income category for patient's zip code (*zipinc*), length of stay (*los*), number of procedures on this record (*npr*), died during hospitalization (*died* – a dummy variable equal to one if died and zero otherwise), expected primary payer (*pay1*), age in years at admission (*age*), and whether the admission was elective (*elective*). The dependent variable is *lntotchg*, which is the natural logarithm of total charges for an inpatient stay. I use the logarithmic transformation to account for possible skewness of the expenditure distribution (to rule out big outliers in the sample). I create two dummy variables for race – black and other race (each variable is equal to one if the patient is black or other race respectively, and zero otherwise). The indicated category is white. I collapse Hispanics, Asians, Native Americans, and others into one dummy variable called *otherrace* because I am generally interested in the potential differential between African-Americans and non-Hispanic whites. I also create categorical variables for health insurance status and median household income for patient's zip code. The expected primary payer variables are Medicaid, Medicare, and private insurance (incl. HMOs and PPOs). The indicated category is other expected primary payer, which includes self-pay, charity, and the like. The median household income for patient’s zip code is not a continuous variable but instead income is reported in ranges or categories, i.e. from $1 to $24,999, from $25,000 to $34,999, from $35,000 to $44,999, and from $45,000 or more. Thus, the following dummy variables are created respectively: low income, below median income, and above median income. The indicated category is high income ($45,000 or more).

As mentioned in a previous section, hospital expenditures depend on disease severity. I control for this relationship by including variables such as disease staging and AHRQ comorbidity measures. I merge national inpatient sample with the HCUP severity data by a common code that is uniquely defined

in both datasets. The AHRQ comorbidity measures define thirty different coexisting medical conditions, which are likely to be present prior to the hospital stay and are not directly related to the reason for hospital admission or principal diagnosis. All these comorbidity measures are defined as categorical variables (equal to one if the patient has the disease and zero otherwise). The presence of comorbidities can increase the cost of treatment, so it is important to rule out their impact on total expenditures for breast cancer patients. Disease staging criteria, developed by Medstat, define the severity for different medical diseases. It is measured on a scale from 0 to 4 with stage one being a disease with no complications; stage 2 is a disease with local complication; stage 3 corresponds to an increased disease complexity – it involves multiple sites or has systemic complications; and stage 4 is death. The staging variable (*ds_stage*) is measured on a continuous scale (has substages) to better represent the severity of a particular disease.

The summary statistics by race are presented in Table 1. Whites seem to be a little older compared to breast cancer patients that are African-Americans (blacks) or belong to other race. The length of stay in a hospital is somewhat longer for African-Americans than for other racial groups on average (3.41 vs. 2.62 and 2.41). In addition, in this data set African-Americans have relatively higher hospital expenditures, lower rate for elective admission, higher mortality rate, more advanced stage of the disease at admission, and are poorer on average in comparison with whites and patients of other races.

Table 1: Summary Statistics by Race

	Black			***Other Race***			***White***		
Variable	**Obs**	**Mean**	**Std. Dev.**	**Obs**	**Mean**	**Std. Dev.**	**Obs**	**Mean**	**Std. Dev.**
age	1672	57.31	14.42	1634	56.75	14.15	12053	62.79	14.45
los	1672	3.41	4.73	1634	2.62	3.8	12053	2.41	2.77
npr	1672	1.85	1.22	1634	1.91	1.26	12053	2.03	1.24
totchg	1672	18162.04	22945.78	1634	17930.95	18109.24	12053	15409.9	14147.13
lntotchg	1672	9.47	0.78	1634	9.51	0.73	12053	9.39	0.69
died	1670	0.03	0.17	1633	0.02	0.14	12039	0.01	0.12
elective	1664	0.72	0.45	1628	0.79	0.41	12025	0.84	0.37
lowincome	1672	0.12	0.33	1634	0.07	0.26	12053	0.02	0.13
belowmedian	1672	0.28	0.45	1634	0.17	0.37	12053	0.18	0.38
abovemedian	1672	0.26	0.44	1634	0.21	0.41	12053	0.25	0.43
ds_stage	1580	1.32	0.71	1614	1.2	0.58	10936	1.16	0.51

Hospital expenditures may be correlated with hospital characteristics. Therefore, I control for this possible relationship using dummy variables for hospital ownership, location, region, and size. Such data is provided in the supplemental HCUP hospital data set, which I merge with the NIS and severity data set. It is important to mention that the perception of a patient about the disease and the social support she receives from her family and friends can influence the timing of hospitalization and length of stay, which indirectly affects hospital expenditures. However, they cannot be easily measured and are not available in the HCUP data.

ECONOMETRIC MODEL AND ESTIMATION TECHNIQUES

The incurred hospital expenditures serve as a proxy of the intensity of breast cancer treatment. Using expenditures, however, should be done with caution because there may be issues with co-insurance and health insurance reimbursement to doctors for government provided insurance such as Medicaid and Medicare. In addition, there is a possibility that a part of these expenditures may be due to defensive medicine. Therefore, I employ various specifications to estimate the possibility of racial discrimination in terms of intensity of treatment for breast cancer patients controlling for health insurance status. To avoid the possibility for highly skewed expenditures or having big outliers that change dramatically the mean, I use the logarithm of expenditures as the dependent variable and estimate several model specifications via ordinary least squares (OLS). Since there can be omitted variable bias that causes heteroskedasticity I employ the Huber-White correction to the OLS regressions. Thus, standard errors will be consistent and inferences can be made. The basic model can be specified as:

$$\text{lntotchg} = \alpha + \beta X + \gamma Race + \theta HI + \varepsilon$$

where *X* is a vector of inpatient claims' characteristics such as age in years at admission, length of stay in the hospital, number of procedures on record, dummy variables for patient's median income, whether the person died in the hospital and whether the admission was elective. The coefficients on the race dummy variables (*black* and *otherrace*) show the difference between the respective base category and the indicated category, white female patients with breast cancer, in terms of log of total expenditures. I control for differences based on the health insurance status using three dummy variables (*Medicaid, Medicare, and Private Insurance*).

In the next specifications, I include controls for hospital characteristics such as location (urban or rural), ownership/control (public, private for profit, and private not-for-profit), size approximated by the number of beds (small, medium, and large), and region (Northeast, Midwest, South, and West). Furthermore, to take into account that expenditures depend on the disease severity, I include dummy variables for disease staging and comorbidities. In addition, I expand the model by adding an interaction term between age and race to account for possible differences in hospital expenditures for women of different ethnicities at different ages. Thus, the expanded model is as follows:

$$\text{lntotchg} = \alpha + \beta X + \gamma Race + \theta HI + \varphi Hospital + \delta Severity + \lambda(XRace) + \varepsilon$$

I also estimate the expanded model via OLS for Medicaid and Medicare claims separately, which solves the problem with having various out-of-pocket expenditures and prices for given procedures provided to patients with different types of health insurance. Finally, I estimate quantile regression models introduced by Koenker and Bassett (1978) that look at the median or different percentiles of expenditures (25% or 75%). As a result, I can make inferences for a possible racial differential at more similar expenditure levels.

ESTIMATION RESULTS

The signs of the OLS estimated coefficients make sense except for the ones on race, private insurance, and elective admission in the specifications where these variables are statistically insignificant. As expected the coefficients on *npr* and *los* are positive and statistically significant at 1% level of significance showing that the log of hospital expenditures goes up as the number of procedures and length of stay increase. The coefficient on *age* is negative, which is consistent with the theory that older patients are treated less intensively on average (makes sense for elderly patients only). The results show that patients with low income, below and above median income have lower hospital expenditures as compared to those coming from high-income zip codes (all coefficients are significant at the 1% or 5% level in all specifications but the last two). The coefficients on *died* and *disease staging* are negative implying that at more advanced stage of the disease, incl. dying in the hospital, patients will have not too many opportunities for treatment and as a result their expenditures will be lower. *Died* is statistically significant in all specifications and *ds_stage* is significant in specifications 1 through 6. Regression results also show that elective admissions lead to higher hospital expenditures on average. This variable is only insignificant in specification 7 with controls for hospital ownership. According to the estimates from the regressions, government (Medicaid or Medicare) or firm provided insurance (private insurance) are associated with higher expenditures per inpatient record (excl. specification 7 for all three types of insurance and specification 6 for Medicaid) as compared to self-pay, charity, or other types of insurance. This result seems plausible considering the possibility that uninsured people will tend to spend less on treatment procedures. The coefficients on health insurance variables are significant at 1% level of significance for specifications 1 to 5 (with controls for hospital size and location). When I add controls for hospital region and ownership some of the signs change and some of the coefficients become insignificant. Specifications 3 and 7 have different comorbidity measures as controls. It turns out that only five comorbidities are significantly affecting expenditures in specification 3 (deficiency anemias, uncomplicated diabetes, metastatic cancer, obesity, and peptic ulcer disease) and only three in specification 7 (metastatic cancer, obesity, and peptic ulcer disease).

The main variables of interest in the regressions are the race variables and the interaction terms of the race variables with age: *black, otherrace, ageblack, and ageotherrace*. The interaction terms are positive and statistically significant in specifications 1 through 6. The coefficients on the two race variables are negative and statistically significant in all but specification 7 which provides some evidence for the presence of racial discrimination in the intensity of breast cancer treatment of African-Americans and other races compared to white Americans.

When I estimate the model for Medicare and Medicaid claims separately with all necessary controls I ignore issues of having difference in prices for various procedures and variation in out-of-pocket expenditures that complicate the analysis. The results from the regression based on Medicare claims illustrate that age, number of procedures, length of stay, disease staging, hospital size, location and ownership are significant predictors of hospital expenditures. Some of the income variables and

comorbidities are also statistically significant. The coefficients on black and other race, as well as on the interaction terms between age and race, are statistically insignificant from zero. The same can be said for the race variables estimated via OLS using Medicaid claims from HCUP. A possible explanation for these results is that patients of different races with such insurance plans are more alike/have similar characteristics and are less likely to be treated differently.

With regards to the quantile regressions, the results seem to be mixed. The coefficients on the health insurance variables and comorbidities, which are statistically significant in the previous specifications estimated by OLS, in all quantile regressions are insignificant. Race variables are also insignificant for the median and 25th percentile. For the 75th percentile, the regression coefficient estimates for *black, otherrace, ageotherrace* have the expected signs and are significant at 1% level of significance. The results suggest that racial discrimination is present only for the inpatient claims in the 75th percentile (claims with higher expenditures) and not for the 50th and 25th percentiles.

CONCLUSION

This paper tries to investigate whether there are disparities in the intensity of breast cancer treatment among different racial groups such as whites, blacks, and others. Based on cross-sectional data from the Healthcare Cost and Utilization Project for year 2002, I estimate different specifications using ordinary least squares and quantile regressions. The coefficients on race variables are negative and statistically significant in most of the specifications providing some evidence for racial discrimination. This result is also confirmed by the quantile regression for inpatient claims in the 75th percentile of hospital expenditures. However, further evidence is necessary to prove that African-Americans and other races are treated less intensively when admitted to hospitals for breast cancer procedures. A possibility for future research is to investigate this question using time-series data, as well as analysis by regions or states.

ENDNOTES

1. The regression results are available from the author upon request.

REFERENCES

Altekruse S.F., C.L.Kosary, M. Krapcho, N. Neyman, R. Aminou, W. Waldron, J. Ruhl, N. Howlader, Z. Tatalovich, H. Cho, A. Mariotto, M.P. Eisner, D.R. Lewis, K. Cronin, H.S. Chen, E.J. Feuer, D.J. Stinchcomb, B.K. Edwards (eds). SEER Cancer Statistics Review, 1975-2007, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2007/, based on November 2009 SEER data submission, posted to the SEER web site, 2010.

Andrews, R.M. and A. Elixhauser. 2000 “Use of major therapeutic procedures: are Hispanics treated differently than non-Hispanic Whites” *Ethnicity & Disease*, 10(3): 384-394.

Ansell, D, S. Whitrnan, R. Lipton and R. Cooper. 1993. “Race, Income, and Survival from Breast Cancer at Two Public Hospitals” *Cancer*, 72(10): 2974-2978.

Bach, P.B. , L.D. Cramer, D. Schrag, R.J. Downey, S.E. Gelfand, and C.B. Begg. 2001. “The influence of hospital volume on survival after resection for lung cancer” *New England Journal of Medicine*, 345(3): 181-188.

Ball, J.K. and A. Elixhauser. 1996. “Treatment differences between blacks and whites with colorectal cancer” *Medical Care*, 34(9): 970-984.

Bassett, M.T. and N. Krieger. 1986. “Social Class and Black-White Differences in Breast Cancer Survival” *American Journal of Public Health*, 76(12): 1400-1403.

Dimick, J.B., J.A. Cowan Jr., G.R. Upchurch Jr., and L.M. Colletti. 2003. “Hospital volume and surgical outcomes for elderly patients with colorectal cancer in the United States” *Journal of Surgical Research*, 114(1): 50-56.

Dowell, M.A., B. Rozell, D. Roth, H. Delugach, P. Chaloux, and J. Dowell. 2004. “Economic and clinical disparities in hospitalized patients with type 2 diabetes” *Journal of Nursing Scholarship*, 36(1): 66-72.

Figueroa, Janis B. and Nancy Breen (1995) “Significance of Underclass Residence on the Stage of Breast or Cervical Cancer Diagnosis” *The American Economic Review*, 85(2): 112-116.

Gordon, N.H., J.P. Crowe, J. Brumberg, and N.A. Berger. 1992. “Socioeconomic Factors and Race in Breast Cancer Recurrence and Survival” *American Journal of Epidemiolog,* 135(6): 609-618.

Harris, D.R., R. Andrews, and A. Elixhauser. 1997. “Racial and gender differences in use of procedures for black and white hospitalized adults” *Ethnicity & Disease*, 7(1): 91-105.

Karjalainen *S.* and E. Pukkala. 1990. “Social class as a prognostic factor in breast cancer survival” *Cancer*, 66(4): 819–826.

Katz, S.J. and T.P. Hofer. 1994. “Socioeconomic Disparities in Preventive Care Persist Despite Universal Coverage” *The Journal of the American Medical Association*, 272(7): 530-534.

Kravdal, Oystein. 2000. “Social Inequalities in Cancer Survival” *Population Studies,* 54(1), 1-18.

LeMarchand L., L.N. Kolonel, A. Nomura. 1984. “Relationship of ethnicity and other prognostic factors to breast cancer survival patterns in Hawaii” *Journal of the National Cancer Institute*, 73:1259-65.

Mitchell, Jean M. and Jack Hadley. 1997. “The Effect of Insurance coverage on Breast Cancer Patients’ Treatment and Hospital Choices” *The American Economic Review*, 87(2): 448-453.

Shen J.J. 2002. “Severity of illness, treatment environments and outcomes of treating acute myocardial infarction for Hispanic Americans” *Ethnicity & Disease*, 12(4):488-98

Thorpe, Kenneth E. and David Howard. 2003. “Health Insurance and Spending among Cancer Patients” *Health Affairs*, 189-198.

Schrijvers C.T., J.W. Coebergh, L.H. Van der Heijden, and J.P. Mackenbach. 1995. “Socioeconomic status and breast cancer survival in the southeastern Netherlands, 1980-1989” *European Journal of Cancer*, 31A(10):1660–1664.

The Effect of Strict Alcohol Policies on America's Sexually Transmitted Disease Rates

Karolyn Caprara*

ABSTRACT

This paper was created to investigate the role that alcohol policies play on America's sexually transmitted disease (STD) rates. Alcohol consumption has been linked with risky sexual behavior and poor decision making. Risky sexual behavior includes multiple sex partners and unprotected sex; both enhancing the probability of contracting an STD. Enforcing strict alcohol policies may reduce or help control the chances of this event from occurring. This paper attempts to locate the relationship between alcohol policies (alcohol tax, legal drinking age and blood alcohol concentration level) with STD rates, specifically Gonorrhea and Chlamydia, of 15-24 year olds in America for the year 2008.

INTRODUCTION

Sexually transmitted diseases have been uncovered for many years, yet the exchange from one person to another still exists. According to the Center of Disease Control and Prevention, in 2008, just in America alone, there are over 300,000 reported cases of gonorrhea and over 1.2 million reported cases of Chlamydia. In both instances, age groups 15-19 and 20-24 were held accountable for the majority of these cases, by overwhelming percentages. The concentration of this study is placed upon youths and young adults, because they face this higher risk then any other age group. Incidence rates of Chlamydia and gonorrhea – the two most common reportable STDs—are highest among American teenagers and young adults (Grossman, Kaestner and Markowitz, 2004). This is quite alarming since antics such as sexual education, television/radio commercials promoting safe sex with condoms, awareness and unlimited amount of information is set out to educate the youth of America; yet, decades after STDs were first discovered, individuals are still practicing unsafe sex and endangering themselves and others.

One major cause could be due to alcohol consumption and the risks that come along with drinking alcohol. Consuming alcohol increases the chances of experiencing unprotected sex and multiple sexual partners. The amount one consumes at one time and how frequently the alcohol is being consumed, can either increase or decrease the probability of having risky behavior even occur. Specifically, having 10 drinks a day on average results in roughly a 60% increase in the probability of possessing multiple partners within the past year (Luong). Consequently, the relationship between the amount of alcohol consumed and the frequency of engaging in sexual relations will have a positive relationship.

Behavioral economics explains such irrational actions of unprotected sex or having multiple partners, as a result of temptation. Something is considered 'tempting' when more of 'it' (whatever 'it' may be, in

* Department of History, Economics and Politics, Farmingdale State College, Farmingdale, NY 11735. Paper from the 4th Annual NYSEA Undergraduate Paper/Presentation Competition.

this instance, sexual activity) is consumed in a state of arousal (hot state) then in a rational state (cold state). For most of us, however, self control issues arise because we underestimate the effect of arousal (Thaler and Sunstein, 2009).Behavioral economist George Loewenstein refers to this as “hot- cold empathy gap”. Hot and cold empathy gap is when an individual in a cold state downplays the significant changes in characteristic traits, like desires and behaviors, once they find themselves in a ‘hot’ state. Temptation for sexual activity can also be stimulated or enhanced with the consumption of alcohol; which faces its own temptation. When alcohol is present, temptation for sexual activity and risky behavior prevails.

Measures can be taken to help reduce the risk of alcohol consumption and the risky behavior associated with it. One way is to try to limit the amount of alcohol one is willing consume at one time. This could be done by raising the taxes, which will raise the total price. In theory, as prices increase and disposable income is held constant, consumption should decrease. This tends to happen, because the new increased priced item will take more out of the disposable income, leaving less available funds for other goods and services. Then the individual has to assess where their greater personal utility lies; between the alcohol and the other item/s. A plausible relationship between higher alcohol taxes and the incidence of sexually transmitted diseases is predicted on a positive correlation between STDs, risky behavior, and alcohol consumption (Luong). Therefore, if overall prices are increased, consumption will fall, resulting in a lower probability that the individual will result to risky sexual behavior. Lowering the chances of risky behavior from occurring will ultimately lower the likelihood of contracting or spreading a sexually transmitted disease.

Another action that can potentially lower the rate of STD’s in America due to alcohol consumption is increasing the legal age limit on purchasing alcohol. This could decrease the availability to minors by decreasing the chances of obtaining and consuming the alcohol. If less people have the ability and the right to purchase alcohol, then theoretically, there should be less alcohol being purchased, distributed and consumed. If there is less alcohol in circulation, then the probability of minors engaging in unprotected sex with multiple partners *should* decrease. The problem here, however, is it’s difficult to determine if alcohol is the reason for casual intercourse among the youth or if these acts are occurring just as frequently in a sober manner as well.

Other economic factors that can influence alcohol consumption, for better or for worse, are household incomes, unemployment, and the gross domestic product (GDP) of the state. If these issues are playing a negative role (having little or no income due to lack of productivity) in a majority of people’s lives, this could cause people to consume alcohol. According to PeaceHealth, a medical group of the northwest region, low income populations are at a higher risk of abusing alcohol more frequently. Individuals with more stress and lower incomes tend to consume alcohol, smoke cigarettes and result to substance abuse over people who are middle class or above.

LITERATURE REVIEW

SEXUAL ACTIVITY DUE TO ALCOHOL CONSUMPTION

There have been numerous studies done on the relationship between teen-young adults binge drinking and sexually transmitted diseases, due to unsafe sexual practices. Researchers on this subject matter have mostly found a direct, positive correlation between consumption and sexual activity that has the possibility of leading to sexually transmitted diseases. For example, Graves and Leigh (1995) show that young adults who drink heavily are more likely to be sexually active and to have multiple partners, and those who are heavy drinkers are also less likely to use condoms (Grossman, Kaestner, Markowitz, 2004). This research suggests that in a [hot] state of drunkenness, peoples actions are less thought out and often more impulsive (then in a cold state), leading to irresponsible actions.

An experiment was conducted by Leif Crowe and William George to observe the effects of alcohol consumption on human responses to sexual stimuli. The results showed that alcohol and sexual inhibition was both psychological and pharmacological; meaning that alcohol alone and learned expectancies about alcohol made up a major role in the sexual activity under the influence. Thus, there is scientific basis for the belief that alcohol consumption might increase the likelihood of participating in a risky sexual encounter (Chesson, Harrison, Kassler, 2000).

ALCOHOL TAXES: THE EFFECT ON CONSUMPTION AND SEXUAL ACTIVITY

Taxes on alcohol, including beer, liquor, and wine influence prices and peoples decisions on their purchases. This is based on the theory that higher prices will result in lower consumption and vise versa. In order to enforce a stricter alcohol policy, raising the alcohol tax can have a positive influence on lowering the probability of individuals acting on risky sexual behaviors, which could lead to STDs.

According to research done in Canada by May Luong on sex, teen pregnancies, STDs and beer prices, increasing beer taxes by $1 has made a major impact in the reduction or gonorrhea and syphilis rates. Specifically, employing cross-state data from 1981-1995, they find a $1 increase in the per gallon beer tax to be correlated with 25.4% and 93.3% reduction in all-age gonorrhea and syphilis rates, respectively (Luong). This shows how a small marginal increase in the total price of alcohol, can have a significant impact on STDs by showing a decline in the consumption of alcohol first.

A similar research that was done in America by Harrell Chesson, Paul Harrison and William J. Kassler, had very comparable outcomes to the research that was done by Luong in Canada. However, instead of using such a large time frame, the American study was done on an annual basis and used both liquor and beer tax in the analysis. They estimated that a $1 increase in the per-gallon liquor tax reduces gonorrhea rates by 2.1 percent and a beer tax increase of $ .20 per six-pack reduces gonorrhea rates by 8.9 percent, with similar though more pronounced effects on syphilis rates (Chesson, Harrison, Kassler, 2000).

LEGAL AGE LIMIT

There have been many disputes and much controversy over what the legal drinking age should be in America. A few decades ago, the drinking age was 18 years old; then jumped right up to 21 years of age. Even though this might have upset a good amount of the population, the 3 year change has proven to be a good decision; at least in the terms of gonorrhea rates. A report by Chesson (2000) reports that raising the age limit lowered gonorrhea rates for ages 15 to 19. In addition, the authors find that a one year increase in the minimum legal age is associated with between a 6.5 to 8% fall in gonorrhea rates among 15 to 29 year olds (Luong). It's assumed that reducing the amount of alcohol that's allowed to be purchased would result in less alcohol winding up in the hands of minors. Other studies, focusing on America's youth, show responsiveness to both the age limit and the zero tolerance laws.

DATA

For this regression model, a cross sectional analysis of the United States was applied for the year 2008. The independent variables consisted of beer tax, liquor tax, wine tax, and the percent of the population that's legally allowed to purchase and consume alcohol. Also, I factored in the economic stress (that could lead to alcohol consumption) of unemployment rate and earnings per capita, of each state. The dependent variables are gonorrhea and Chlamydia rates per 100,000. Therefore:

$F(G,C)= \alpha + \alpha 1B + \alpha 2L + \alpha 3W + \alpha 4P + \alpha 5U + \alpha 6\ EPC.$

G = gonorrhea **W =** wine tax

C = Chlamydia **P =** percent of legal consumption age

B = beer tax **U =** unemployment

L = liquor tax **EPC =** earnings per capita

The rates were based on ages 15 to 24, because they are the highest rate of new reported cases in both categories. Gonorrhea and Chlamydia were both used because they are the highest reported as well. The taxes on beer, wine, and liquor are all based on a per gallon addition. Below, you will find the data on 45 out of the 50 states. The missing states include Nevada, Oregon, Texas, Utah, and Wisconsin. These states were left out due to missing data reports in the independent variables.

The descriptive statistics show that the mean for each variable and the standard deviation (SD), which is a measure of how spread out the numbers are. Gonorrhea and Chlamydia have means of 96.6 and 388.8, and standard deviations of 64.5 and 118, respectively. As far as independent variables are concerned liquor, wine, and beer taxes have means of $5.82, $0.77, and $0.27 and standard deviations of $4.51, $0.51, and $0.25, respectively. The mean of the percent of legal age- consumers is 73.15%, unemployment has 5.29% and earnings per capita are $39,378. Their standard deviations are 1.47%, 1.23% and $5,952, respectively.

REGRESSION MODELS

Gonorrhea Regression Results

Regression Statistics	
Multiple R	0.649643
R Square	0.422036
Adjusted R Square	0.330779
Standard Error	52.78839
Observations	45

	df	*SS*	*MS*	*F*	*Significance F*
Regression	6	77323.2	12887.2	4.624681	0.001283
Residual	38	105891	2786.614		
Total	44	183214			

	Coef.	*Std Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	1201.14	411.50	2.9189	0.005874	368.101	2034.1	368.10	2034.18
Spirits Tax (Per Gallon)	3.57872	2.0813	1.7194	0.0936	-0.6347	7.7922	-0.63477	7.79223
Table Wine Tax (Per Gallon)	-9.3478	21.950	-0.425	0.6726	-53.784	35.088	-53.7844	35.0887
Beer Tax (Per Gallon)	-5.02812	46.626	-0.107	0.9146	-99.417	89.361	-99.4176	89.3613
% allowed to consume alcohol	-15.4291	5.6893	-2.711	0.0099	-26.946	-3.9116	-26.9466	-3.91162
Unemployment Rate	20.9808	7.4306	2.8235	0.0075	5.93832	36.023	5.93832	36.0234
Earnings per Capita	-0.00252	0.0014	-1.785	0.0821	-0.0053	0.0003	-0.00538	0.00033

Based on these results the equation for gonorrhea would be:

$$G = 1201.145 + 3.58L - 9.35W – 5.03B -15.43P + 20.98U – 0.002EPC$$

This indicates that wine tax, beer tax and earnings per capita have a negative correlation; which was predicted to happen. This implies that as these taxes and earnings per capita grow, the rate of gonorrhea will decrease. Unemployment had a positive correlation(as expected) indicting that as the unemployment rate increases, so does the gonorrhea rate. Liquor tax and the percent of legal aged consumers had unanticipated results. Liquor tax implies that as the taxes increase the rate of gonorrhea will rise as well. This could be indicating that individuals are purchasing a higher proof liquor to make up for the extra addition on price. The percent of legal aged consumers suggests that as the percentage of legal drinkers grows the lower the rate of gonorrhea will be.

The astonishing part is with 91% confidence that the liquor tax and 99% confident the percent of legal aged consumers, has great significance on the rate or gonorrhea. In addition, wine tax and beer tax hold barely any significance at all. The most significant (and without unusual results) is the unemployment rate with 99% confidence that it is relevant to the overall rate. According to the adjusted R-Squared, the independent variables presented explain 33% based on the dependant variable (gonorrhea). *Since wine tax and beer tax obtain low significance the new equation would be:

G = 1201.145 + 3.58L -15.43P + 20.98U – 0.002EPC

Next we will analyze how the same independent variables affect the rate of Chlamydia in the United States for the year 2008.

Chlamydia Regression Results

Regression Statistics	
Multiple R	0.611802
R Square	0.374302
Adj R Square	0.275508
Stand. Error	100.9222
Observation	45

	df	*SS*	*MS*	*F*	*Significance F*
Regression	6	231533.8	38588	3.788698	0.004695
Residual	38	387040.7	10185		
Total	44	618574.5			

	Coef.	*Std Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	2837.3	786.72	3.60	0.001	1244.729	4429.99	1244.729	4429.995
Spirits Tax (Per Gallon)	-0.185	3.979	-0.04	0.963	-8.240	7.871	-8.240	7.871
Table Wine Tax (Per Gallon)	9.909	41.966	0.23	0.815	-75.046	94.863	-75.046	94.863
Beer Tax (Per Gallon)	61.149	89.141	0.68	0.497	-119.307	241.606	-119.307	241.606
% Legal Age	-34.34	10.877	-3.15	0.003	-56.361	-12.322	-56.361	-12.322
Unemployment Rate	23.202	14.206	1.63	0.111	-5.557	51.960	-5.557	51.960
Earnings per Capita	-0.002	0.003	-0.7	0.444	-0.008	0.003	-0.008	0.003

Based upon these results the equation for Chlamydia rates would be:

C = 2837.36 – 0.185 L + 9.91W + 61.149B -34.34P +23.2U -0.002EPC

In this situation, liquor tax, unemployment rate and earnings per capita had expected results. Liquor tax and earnings per share have a negative relationship with Chlamydia rates (as they raise, Chlamydia rates decrease). Unemployment had a positive relationship where Chlamydia rates increase when unemployment increases. Wine tax, beer tax, and percent of legal aged consumers had unexpected results. Wine and beer tax both shared a positive relationship; implying that they both increase the rate instead of lowering it as predicted. However, none of the explained variable have any true significance, except for the percent of legal aged consumers. With 99% confidence, the percentage has a significant impact on the Chlamydia rates; as the pool of legal aged drinkers increase, the number of Chlamydia cases decrease. The R-Square implies that 37% of the dependent variable is being explained.

Since all of the dependent variables, besides the percent of legal aged consumers, are under 90% confidence levels, they get dropped from the equation; making the new Chlamydia equation:

C = 2837.36 -34.34P

CONCLUSION

This study was done to test the effectiveness of stricter alcohol policies on the United States sexually transmitted disease rates; specifically gonorrhea and Chlamydia. What we found was unpredicted and, for the most part, somewhat surprising. Both gonorrhea and Chlamydia used the same independent variables for their separate analysis (beet tax, liquor tax, wine tax, percent of legal aged consumers, and stress variables of unemployment and earnings per capita), but received very different results. Gonorrhea was affected significantly by liquor tax, unemployment rates and earnings per capita. This indicates that both alcohol policies and economic stress pay a part in determining the gonorrhea rate. Chlamydia was only correlated with age limit; however, in a different manner then predicted. This indicates that the higher the percent of legal drinkers, the lower the rate of chlamydia will be. This could be implying that the legal age limit should either stay where it is or even become lower. If it's lowered, the percent of legal aged consumers would increase, which should result in lower rates. This could also suggest that raising the age limit has led to underage binge drinking. However, these assumptions would need further investigation.

REFERENCES

Bureau of Labor Statistics: http://www.bls.gov

Center of Disease Control and Prevention: http://www.cdc.gov

Chesson, Harrell and Harrison, Paul and Kassler, William J. 2000."Sex under the Influence: The Effect of Alcohol Policy on Sexually Transmitted Disease Rates in the United States." *Journal of Law and Economics*. 43(1): 215-38.l

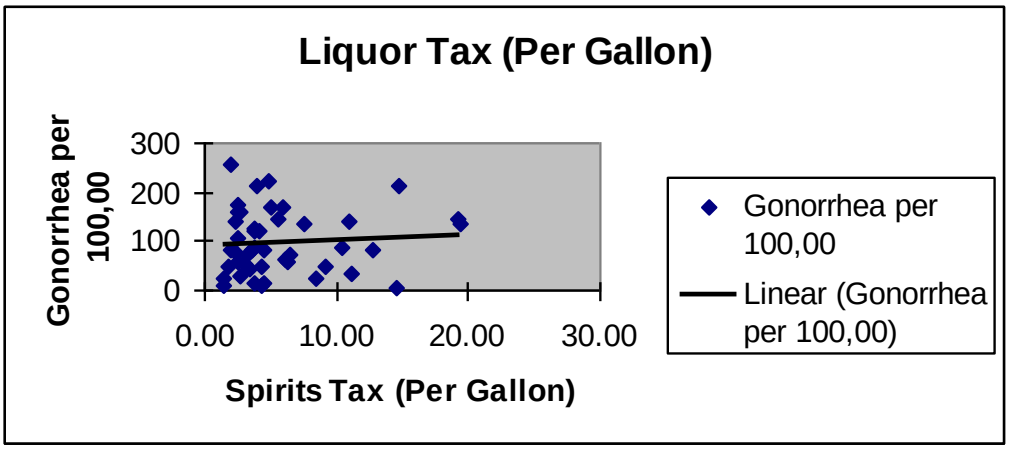
Grossman, Michael, Kaestner, Robert and Markowitz, Sara. 2004. "An Investigation of the Effects of Alcohol Policies on Youth STDs." Working Paper 10949. *National Bureau of Economic Research.*

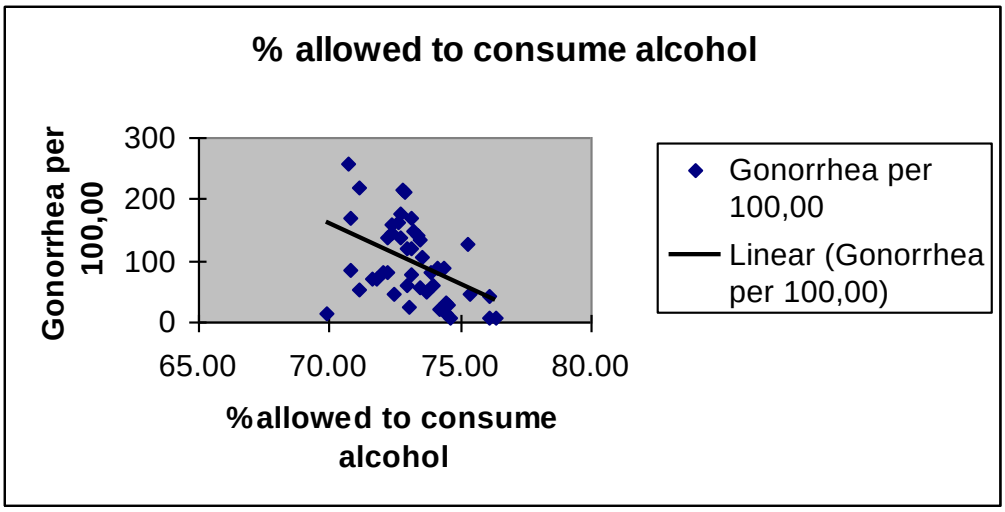
Luong, May and Sen, Anindya. 2006. "Sex, Teen Pregnancies, STDs, and Beer Prices: Empirical Evidence from Canada." Working Paper 06001, University of Waterloo, Department of Economics.

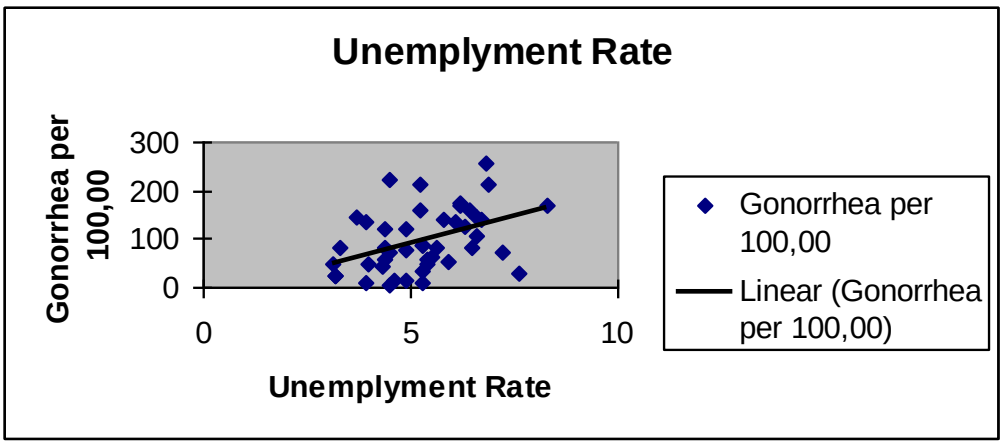
Thaler, Richard H. and Sunstein, Cass R. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, Connecticut and London, England: Yale University Press.

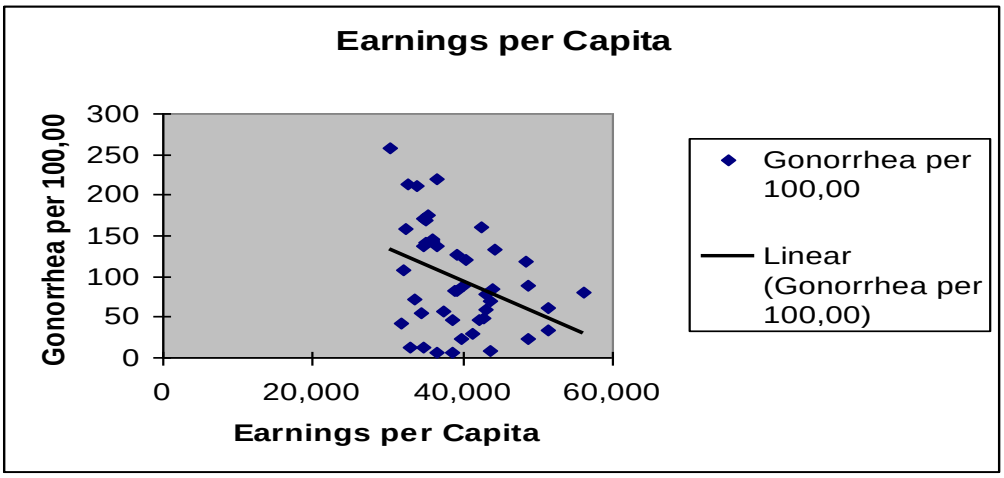
APPENDIX

GONORRHEA CHARTS

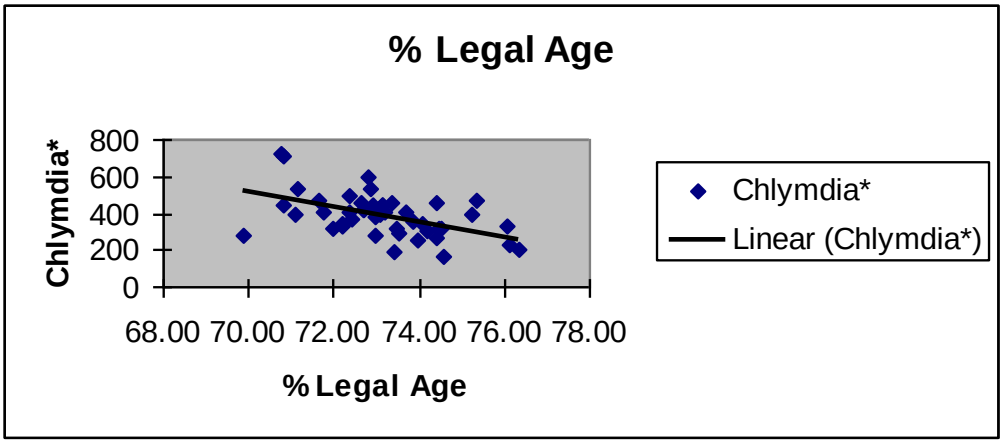








CHLAMYDIA CHARTS



Is Small Beautiful? Size-Productivity Relationship in Indian Manufacturing

Prabal K. De[*] and Priya Nagaraj[†]

ABSTRACT

We study the size-productivity relationship of manufacturing firms in India. Like other developing countries, India has been characterized by large size heterogeneity of firms within the same industry. In the context of growing liberalization of trade, licensing and monopoly policies since 1991 and growing incentives to small and medium scale firms, productivity implications of size heterogeneity has become more important. We find that smaller firms that are in the lowest quintile of the asset distribution are more productive than their larger counterparts. For identification, we exploit the panel nature of the data and use production-function estimation technique as developed by Levinsohn and Petrin (2003) to estimate the total factor productivity and Arellano and Bond (1991) type of dynamic estimation to control for potential endogeneity arising out of the persistence of productivity.

INTRODUCTION

Growth literature has long documented the importance of productivity enhancement in raising income, particularly in developing economies. Historically, manufacturing sector has been the principal component of income growth and has been subject to intense research in the past decade in search of explanations of productivity variation. One among many peculiar characteristics of manufacturing sectors in developing countries is the size-heterogeneity of firms. Industries within the manufacturing sectors in developing countries have been characterized by coexistence of firms of such heterogeneous sizes that Tybout (2000) calls this phenomenon a form of 'dualism'. However, productivity implications of such heterogeneity are not clear. While big firms have scale advantage and often have better access to capital and other resources through political and economic channels, small firms are likely to have leaner management and more flexibility in resource allocation. This polemic has become particularly important in the last decade or so for at least two reasons. A number of economies have liberalized their trade and licensing regimes to allow easy entry of firms and to reduce monopolies of either big state enterprises, or big private firms. Second, much richer micro-datasets are increasingly being available to empirically assess the dynamic changes in productivity.

In this paper we investigate whether small firms are likely to be more productive. Size-productivity relationship in the manufacturing sectors of developing countries is an important question in its own right for the following reasons: First, theoretical literature has not been conclusive. Therefore, there is value in estimating a reduced form relationship rather than estimating some specific structural model. Second,

[*] City College, City University of New York. pde@ccny.cuny.edu

[†] The Graduate Center, City University of New York. PNagaraj@gc.cuny.edu We are grateful to the seminar participants at CCNY and the New York State Economic Association Annual Meeting for helpful comments.

empirical evidence of size-productivity relationship is scant, particularly for developing countries. Finally, as discussed in the next section, India has been pursuing a policy of providing incentives to small firms over the last decade or so. Similar trends in policy can also be found in the other parts of the developing world also. Without empirical evidence, we cannot evaluate such a policy.

SMALL SCALE SECTOR IN INDIA

Being labor intensive, small enterprises promote employment, especially of unskilled labor. They play an important role in economic growth by contributing to the GDP (Beck et al, 2005). Among the Asian countries, India has been unique in terms of its focus on the development of small and medium enterprises since Independence. Over the decades, the policies governing the small scale industries changed from promotion to protection. Financial incentives and reservations for the small scale sector emerged as the policy instruments in the 1960's. By 2003-2004, the Small Scale Industry units accounted for more than 40 percent of gross value of output in manufacturing sector and about 34 percent of total exports of the country. It is the highest employment-providing sector after agricultural sector. Though 87 percent of the SSI units are unregistered, the registered units account for 72 percent of the total SSI production and 87 percent of total SSI exports. (Source: Development Commission (SSI), Third Census).

There are a number of ways in which the ministry helps the small firms - from directing credit, easing collateral requirement for loans, providing easy credit for technology improvement, providing training for entrepreneurs and providing technical training to tax concessions. Though the Government has now undertaken a number of policy initiatives in this direction, promoting SMEs might not be in the best interest of the economy if the SMEs are not more productive than their larger counterparts. Therefore, empirical findings on size-productivity relation will have bearing on the impact of policies promoting small and medium scale enterprises. This is true for not only India, but other countries too.

THEORETICAL PRIORS AND RELATED LITERATURE

While one would expect big firms to reap the traditional benefits of economies of scale, and in case of developing countries, where institutions are typically weak, other benefits such as access to license, finance and government contracts, we have not found a theoretical approach that succinctly models it. Idson and Oi (1999) argue that workers in large firms reap the benefits of increasing returns (brought in by big volumes) by having less idled time and producing more. This provides an explanation for higher productivity in larger firms. Tybout (2000) also attributes the (potential) higher productivity of large firms in developing countries to variables like increasing returns and lobbying power. One of the earliest of such attempts was made by Williamson (1967). In a model of hierarchical control, he showed that benefits of increasing returns from growing in size are countervailed by the increasing cost of managerial complexity. The models that followed, built on this advantage of small firms – leaner and more flexible management. Further, small firms can be more receptive and adaptive to the new technology. One advantage of having a smaller scale is that the production process is less deeply entrenched in existing technology. Finally,

analyzing manufacturing firms in nine African countries, Van Biesebroeck (2005) found that large firms are more productive and have higher growth rates. However, Bigsten and Gebreeyesus (2007) find in Ethiopia that small firms actually grow faster than the larger ones. In the United States, Dhawan (2001) finds that smaller firms are more productive.

DATA AND SIZE DISTRIBUTION OF FIRMS IN INDIAN MANUFACTURING

The data used is taken from Centre for Monitoring Indian Economy (henceforth, CMIIE)'s PROWESS database. CMIE compiles this data from the audited financial results of listed and unlisted firms. The firms in Prowess accounts for almost 75 per cent of all corporate taxes and over 95 per cent of excise duty collected by the government of India.

We use an unbalanced panel data of manufacturing firms for the period 1994 to 2008. The choice of time period has been dictated by the intention to keep the most recent data and avoid having a lopsided sample. Number of firms covered in the PROWESS database has jumped steeply in 1994. The effects of the economic reform of 1991 were evinced gradually by 1993. The data has detailed information on financial variables including break up of assets, liabilities, income and expenses. All variables have been deflated by the wholesale price index. Since the database contains information on only the wage bill and not on employment, we infer the amount of labor by deflating the wage bill by average industry wages. The data on wage per worker has been taken from Central Statistical Organization's (CSO)'s Annual Survey of Industries (ASI) database. The PROWESS data has firms from 22 NIC-2 digit industries.

Table 1 shows the distribution of assets in terms of mean and median over five quintiles. We can make several key observations from this table. First, there is substantial heterogeneity in terms of firm size. The mean size in the highest quintile is more than 100 times bigger than the mean size in the lowest quintile. Second, the average size in the highest quintile has been steadily rising over the years, signifying an increasing concentration at the top. Third, as the higher up we go in quintiles, more skewed the size distribution becomes, as reflected by the wedge between mean and median.[1]

Table 1: Distribution of Assets among Indian Manufacturing Firms (in INR Million)

year	Quintile1		Quintile2		Quintile3		Quintile4		Quintile5	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
1994	34	34	93	92	179	175	404	381	4210	1440
1995	34	35	90	89	179	176	403	378	4090	1430
1996	40	42	102	100	205	199	480	456	4790	1770
1997	43	45	108	106	217	211	519	489	5320	1970
1998	41	44	106	104	220	218	525	498	5650	2010
1999	37	39	98	94	207	203	496	465	5320	1900
2000	34	35	95	94	199	192	478	448	5020	1810
2001	34	35	94	91	202	194	493	459	4970	1830
2002	32	33	91	90	196	191	485	450	5200	1870
2003	27	28	85	85	179	172	429	406	4950	1700
2004	23	24	79	77	177	170	446	424	5170	1770
2005	21	21	77	76	186	178	488	454	5840	1980
2006	23	23	89	86	223	215	590	556	7040	2270
2007	28	27	108	105	293	288	812	769	9380	3230
2008	40	38	185	171	519	510	1360	1270	14700	5530

Note: Within every industry and year, firms have been divided into five quintiles according to the size of its assets

EMPIRICAL STRATEGY

In order to estimate the effects of a firm's size on its productivity, we start with the following reduced form baseline specification:

$$TFP_{ijt} = \beta_0 + \beta_s SMALL_{ijt} + \beta_m imports_{ijt} + \beta_x exports_{ijt} + \beta_r rnd_{ijt} + \beta_o Ownership_{ijt} + \beta_h BusinessHouse_{ijt} + age_{it} + agesq._{it} + \lambda_i + \lambda_j + \lambda_t + \eta_{ijt} \ldots\ldots\ldots\ldots (1)$$

The dependent variable is the Total Factor Productivity of a firm i, in industry j and period t. SMALL is an indicator of the size of the firm by assets within the industry it belongs to. For each industry and year, firms have been categorized into five quintiles according to their asset sizes. The variable is therefore comparable across industries and across years. The indicator small here refers to the lowest quintile, that is, the lowest 20 percent of the firms by asset size. β_s is our coefficient of interest. The rest of the control variables in equation (1) have been chosen to ameliorate the omitted variable bias. The choice of variables is driven mostly by either previous literature or theoretical prediction.

In what follows, we will estimate a variety of this baseline specification. We will start with the random effects model where the firm-specific time-invariant unobserved heterogeneity λ_i is assumed to be uncorrelated with η_{ijt}. The second set of estimates is calculated under the alternative assumptions of fixed effects model. Finally, we dynamize this model by including the lagged dependent variable in the right hand side and by using the lagged values of exports and R&D as instruments of themselves as

proposed in Arellano and Bond (1991), and Arellano and Bover (1995) and applied in TFP estimation by Fernandez (2007) and Khandelwal and Topalova (2010) among others.

CALCULATION OF TFP

The estimation of total factor productivity using the traditional method of Ordinary Least Square leads to simultaneity issues. With a positive shock to productivity, the use of inputs also increases. The residual will therefore be a biased estimate of productivity. To overcome this simultaneity issue, we use the method pioneered by Levinsohn and Petrin (2003) by using intermediate inputs (such as material and fuel) as a proxy in place of investment. Conditional on capital, profit-maximizing behavior leads more productive firms to use more intermediate inputs.

IDENTIFICATION AND ENDOGENEITY CONCERNS

Our estimation captures the relationship between firm size and productivity. There are two threats to the claim that any correlation is causal. First, there may be omitted firm characteristics (either unobserved, or lacking in our data) that may affect both firm size (and growth) and productivity. The second concern is that the causality may run in the reverse direction to our claim – firms experiencing higher productivity may acquire new assets and grow in size.

We take a variety of measures to deal with these potential endogeneity concerns. First, we include firm fixed effects. This controls for the unobserved time-invariant firm characteristics such as managerial ability and spatial characteristics. Second, we also include year fixed effects that control for across-the-economy policy changes in a particular year and industry fixed effects that control for industry specific heterogeneities. Together, they control for industry specific changes in regulations over time.

However, while these measures ameliorate the endogeneity concerns in a static model, the endogeneity might arise due to the time-dependency of TFP.(Syverson,2010). Hence we also consider dynamic models developed by Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998). These models further improve upon our specification in the following ways. First, in the absence of exogenous instruments, ABB estimators use lagged values of the control variables as instruments. This makes the endogenous variables pre-determined and, therefore, not correlated with the error term in equation (1). Secondly, the first-differenced GMM estimator naturally removes time-invariant firm characteristics, thus preserving the benefits of using a fixed effects model. Moreover, if TFP is indeed correlated with its past values, then the static model produces inconsistent estimates because the regressor(s) will be mechanically correlated with the error term. ABB-style system GMM controls for this by using past values as instruments. Finally, as Roodman (2006) notes, for a short time (T), long units (N) panel, there is less likelihood that correlation of the lagged dependent variable with the error term will decline over time to be rendered eventually insignificant. ABB estimators are particularly relevant for these cases.

MAIN RESULTS

Table 2 summarizes the main results of the paper. Columns of Table 2 present results from estimation of six variations of equation 1. We start from estimating the simplest relationship between size and productivity with no controls (column (1)). We see that there is a strong positive correlation between belonging to the lowest asset quintile and productivity. Column (2) introduces year and industry fixed effects to control for unobserved industry and time variability. Then column (3) and column (4) present results from the random effects (generalized least square) models. For all four columns, we see that being small means more productive and significantly so. Coefficient estimates go down slightly when we introduce controls to our estimation. Next, we include firm fixed effects in equation (1). Being small is still significant with slightly lower coefficient estimate (column 5, row 1). Finally, we present estimates from the dynamic model *a la* Arellano and Bond (2001). In this case, two lags of the dependent variable TFP are included as regressors and the estimation is performed as a system generalized method of moments procedure.

As column (6) shows, the dynamic model gives us the best fit. All regressors are significant. The first and second lags of the dependent variable TFP are significant and less than unity.[2] Even in this case, small firms turn out to be significantly more productive, though both coefficient estimate and standard errors are less in the case of dynamic model. Therefore, the overwhelming evidence from Table 2 tells us that small firms are more productive than their larger counterparts, after controlling for the unobserved variables at the level of the firm, industry and time and for the dynamism in TFP.

ROBUSTNESS CHECKS WITH BALANCED PANEL

Table 2 reports estimation results of equation (1) with an unbalanced panel of firms. Firms enter and leave industries all the time. However, as discussed in the earlier section, some exit behavior may bias the sample towards firms with higher productivity. Therefore, we restrict the sample to firms who stayed on throughout the sample period of 1994-2007 to form a balanced panel of firms. The results from estimating the same set of models (except for single regressor regressions to avoid clutter) are presented in Table 3.

Except for the first two columns, Table 3 is organized in the same way as Table 2. Evidence in this table also confirms the earlier evidence from unbalanced panel of firms that small firms are significantly more productive than their larger counterparts. The sign and significance pattern of the other control variables confirm that entry and exit of firms do not seem to affect the results in any particular way.

Table 2: Size-Productivity Relationship – Regression Results for UNBALANCED PANEL (Dependent Variable: Total Factor Productivity at firm-level)

	(1) RE	(2) RE	(3) RE	(4) RE	(5) FE	(6) AB
SMALL (Firms in the lowest 20% in asset size)	0.0748***	0.0745***	0.0695***	0.0694***	0.0524***	0.0422***
	(0.0075)	(0.0071)	(0.0072)	(0.0071)	(0.0089)	(0.0064)
L.TFP						0.1168***
						(0.0036)
L2.TFP						0.2320***
						(0.0031)
Industrial Group Dummy			-0.0861	-0.0843	0.0000	0.5426***
			(0.1132)	(0.1145)	(0.0000)	(0.1051)
Ownership Indian Private			-0.0204	-0.0136	0.0000	0.7221***
			(0.1127)	(0.1142)	(0.0000)	(0.1012)
Ownership Foreign Private			-0.0975	-0.1064	0.0000	0.5232***
			(0.1130)	(0.1134)	(0.0000)	(0.1124)
Observations	39886	39886	39822	39822	39822	26424
Year Fixed Effects		yes	yes	yes	yes	
Industry Fixed Effects		yes		yes		
Firm Fixed Effects					yes	yes

Robust standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Notes:

I. All regressions include firm demographics such as age and age squared and exports, imports and R&D as percentage of sales.

II. Ownership is a categorical variable with categories “Indian private ownership”, “foreign private ownership” “Indian government ownership” respectively. The third one is the base category in regression.

Table 3: Size-Productivity Relationship – Regression Results for BALANCED PANEL (Dependent Variable: Total Factor Productivity at firm-level)

	(1) RE	(2) RE	(3) FE	(4) AB
SMALL (Firms in the lowest 20% in asset size)	0.0875***	0.0876***	0.0497***	0.0236**
	(0.0101)	(0.0098)	(0.0126)	(0.0100)
L.TFP				0.3987***
				(0.0092)
L2.TFP				0.2221***
				(0.0071)
Industrial Group Dummy	-0.0669***	-0.0727***	0.0000	0.1356***
	(0.0154)	(0.0266)	(0.0000)	(0.0235)
Ownership Indian Private	0.1036***	0.1257***	0.0000	0.3205***
	(0.0132)	(0.0165)	(0.0000)	(0.1157)
Ownership Foreign Private	0.0657***	0.0655***	0.0000	0.5393***
	(0.0123)	(0.0131)	(0.0000)	(0.1155)
Observations	13496	13496	13496	11466
Year Fixed Effects	yes	yes	yes	
Industry Fixed Effects		yes		
Firm Fixed Effects			yes	yes
Robust standard errors in brackets				
*** p<0.01, ** p<0.05, * p<0.1				

Notes:

I. All regressions include firm demographics such as age and age squared and exports, imports and R&D as percentage of sales.

II. Ownership is a categorical variable with categories “Indian private ownership”, “foreign private ownership” “Indian government ownership” respectively. The third one is the base category in regression.

SUMMARY AND POLICY IMPLICATIONS

In this paper, we have used firm level panel data to estimate the differences in the productivity of large and small firms in the manufacturing sector in India. Such exercise has been motivated by several stylized, theoretical, empirical and policy observations. Firm size heterogeneity is widespread among developing countries. Mammoth firms coexist with smaller firms and continue to produce similar products. However, theoretically both small and large firms have productivity advantages and disadvantages such as scale economies vs. smaller and more flexible management structure. Empirical evidence from the US and the rest of the world has been piecemeal and mixed. Finally, several countries including India have been pursuing policies to promote small and medium scale enterprises. With availability of new firm-level micro data, new evaluation of size-productivity relationship contribute to both researchers’ and

policymakers' understanding of the implications of firm size heterogeneity for productivity and growth in the economy.

We combine several strategies for identification. We calculate total factor productivity by using Levinsohn-Petrin method to control for simultaneity between input choice and productivity shocks. In estimating our main specification, we exploit the panel nature of the data to remove unobserved time-invariant firm heterogeneities. We also use dynamic panel data technique to control for -persistence of total factor productivity.

As far as policy implications are concerned, our findings buttress the case for encouraging small firms. However, this does not mean that large firms need to be broken up or firm growth should be stifled or firms should be stopped from being merged or acquired. A deeper understating of size-productivity relationship is required, particularly of the channels through which such productivity benefits accrue to smaller firms.

ENDNOTES

1. Note that these are sector-wide numbers that aggregate *all* industries. It may be improper to compare firms in, say, textile industry to heavy industries like steel and automobile. However, listing all industries will lead to unmanageable clutter. Such results, along with a longer version of the paper are available upon request.
2. When regress TFP on its first lag, we get a coefficient of 0.58 – similar to the range of 0.6 - 0.8 found by, for e.g., Ábrahám and White (2007) and Foster, Haltiwanger, and Syverson (2008).

REFERENCES

Ábrahám, Árpád and White, Kirk T. 2007. *The dynamics of plant-level productivity in U.S. manufacturing.* Rochester, NY: University of Rochester Working Paper.

Arellano, M., & Bond, S. 1991. "Some tests of specification for panel data: Monte carlo evidence and an application to employment equations." *The Review of Economic Studies, 58*(2), 277-297.

Arellano, M., & Bover, O. 1995. "Another look at the instrumental variable estimation of error-components models." *Journal of Econometrics, 68*(1), 29-51.

Beck, T., Demirgüç-Kunt, A., & Maksimovic, V. 2005. "Financial and legal constraints to growth: Does firm size matter?" *The Journal of Finance, 60*(1), 137-177.

Bigsten, A., & Gebreeyesus, M. 2007. "The small, the young, and the productive: Determinants of manufacturing firm growth in Ethiopia." *Economic Development and Cultural Change, 55*, 813-840.

Blundell, R., & Bond, S. 1998. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics, 87*(1), 115-143.

Dhawan, R. 2001. "Firm size and productivity differential: Theory and evidence from a panel of US firms." *Journal of Economic Behavior & Organization, 44*(3), 269-293.

Fernandes, A. M. 2007. “Trade policy, trade volumes and plant-level productivity in Colombian manufacturing industries.” *Journal of International Economics, 71*(1), 52-71.

Foster, L., Haltiwanger, J., & Syverson, C. 2009. “The slow growth of new plants: Learning about demand?” Unpublished Manuscript.

Idson, T. L., & Oi, W. Y. 1999. “Workers are more productive in large firms.” *American Economic Review, 89*(2), 104-108.

Khandelwal, A., & Topalova, P. 2010. “Trade liberalization and firm productivity: The case of India.” *Forthcoming, Review of Economics and Statistics.*

Levinsohn, J., & Petrin, A. 2003. “Estimating production functions using inputs to control for unobservables.” *Review of Economic Studies, 70*(2), 317-341.

Roodman, D. 2006. “How to do xtabond2. Paper presented at the *North American Stata Users' Group Meetings 2006.”*

Syverson, C. 2010. “What determines productivity?” *NBER Working Paper.*

Tybout, J. R. 2000. “Manufacturing firms in developing countries: How well do they do, and why?” *Journal of Economic Literature, 38*(1), 11-44.

Van Biesebroeck, J. 2005. “Firm size matters: Growth and productivity growth in African manufacturing.” *Economic Development and Cultural Change, 53*, 545-583.

Williamson, O. E. 1967. “Hierarchical control and optimum firm size.” *The Journal of Political Economy, 75*(2), 123-138.

Capital Flows to Russia, Ukraine, and Belarus: Does "Hot" Money Respond Differently to Macroeconomic Shocks?

Scott W. Hegerty*

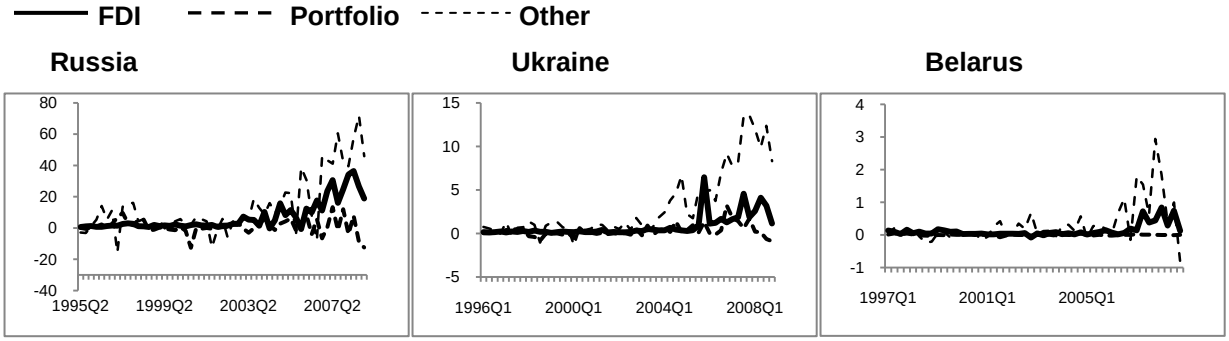
ABSTRACT

Capital flows into the former Soviet bloc have increased tremendously since the mid-1990s. This study applies the structural VAR model of Ying and Kim (2001) to investigate the macroeconomic "push" and "pull" factors behind net flows of FDI, portfolio, and other investment into Russia, Ukraine and Belarus. Impulse-response and variance decomposition analysis shows that domestic income and monetary shocks, as well as foreign income and interest-rate shocks, have effects that vary by flow and by country. Russian FDI and portfolio investment show significant, but different, responses to income and foreign interest-rate shocks. In addition, Belarus responds to improved macroeconomic fundamentals.

INTRODUCTION

With the ongoing transition to a market economy, foreign capital has poured into the former Soviet sphere. Some countries, particularly those that reformed rapidly enough to join the European Union in 2004, have been particularly attractive destinations for this investment. While others (such as the Balkan countries and most former Soviet republics) have not received as much in relative terms, they have still been able to attract a considerable amount of foreign capital. Russia, for example, attracted more than U.S. $16 billion in Foreign Direct Investment (FDI) and $8 billion in portfolio investment during the fourth quarter of 2007. Figure 1 shows the growth in FDI, portfolio, and Other investment flows into Russia, Ukraine, and Belarus since the mid-1990s. While FDI has been increasing, particularly in recent years, Other investment (classified by the IMF to include loans and trade credits) has grown even faster.

Figure 1. Net Capital Inflows (Real, Billions of U.S. dollars).



Source: International Financial Statistics of the IMF. Deflated with U.S. GDP Deflator, 2005 = 100.

* Department of Economics and Finance, Wehle School of Business, Canisius College, Buffalo, NY 14208

While these flows can be beneficial—FDI can facilitate technology transfers and other positive spillover effects, and portfolio investment can help deepen a country's capital markets—there are also risks. Capital can also be withdrawn, particularly "hot money" flows such as portfolio investment. These outflows can cause major problems—including, as Melecky (2005) noted, slower future growth. Russia, for example, experienced a capital outflow in late 2008 that could be simultaneously attributed to the world recession, domestic factors, or investor dissatisfaction after the Georgia invasion.

As a result, it is important to empirically model the macroeconomic determinants of capital flows, both for various countries as well as for different flows. This paper examines the effects of macroeconomic shocks for Russia, Ukraine, and Belarus in particular. At the same time, two sources of ambiguity in the literature—the effects of income shocks and the differences between the responses by different types of capital flow—are given considerable attention. The application of a structural VAR using quarterly data provides some surprising and useful results.

RELATIONSHIP TO THE LITERATURE

The so-called "push-" and "pull factors" that drive capital flows can include income and monetary shocks. Economic theory, however, suggests that the effects of these shocks are ambiguous; they can decrease capital flows as well as increase them. Glick and Rogoff (1995) show that current accounts are driven by savings and investment decisions that differ depending on the persistence of income shocks. Since extra income can be saved or invested abroad in varying proportions, the current account (and thus capital flows) can be pro- or countercyclical and must be modeled empirically.

Likewise, this paper addresses another source of ambiguity: Whether FDI and non-FDI flows really behave differently from one another. The literature is divided between two views. One, proposed by Claessens *et al.* (1995) suggests that these differences might be minimal. On the other hand, Chuhan *et al.* (1996) and Sarno and Taylor (1999) find that these flows might indeed exhibit unique properties.

The three countries in question are also dissimilar, since each is in a different stage in the transition process. Belarus is the smallest and has the weakest market economy; as a result, it has attracted the smallest net inflows. Russia's oil exports make it more dependent on global factors, but the country's political situation might make investors wary. Ukraine has also recently undergone a series of currency devaluations, in addition to having a rocky relationship with Russia. Thus, variation among the countries in general can also be expected.

This type of analysis has not been done for these countries, however. Previous research on capital flows to Central and Eastern Europe has neither focused on individual countries nor included foreign macroeconomic influences. Lipschitz *et al.* (2006) provide a theoretical framework behind the movement of capital (or lack thereof) to transition economies, emphasizing the importance of structural and institutional factors, without focusing on any specific economy. Working papers such as those by Ötker-Robe *et al.* (2007) and Von Hagen and Siedschlag (2008) describe recent developments regarding the growth of capital inflows, but omit Russia, Ukraine, and Belarus entirely. Other studies, such as Garibaldi

et al. (2001), include these three countries, but focus on more microeconomic variables. Likewise, Lane and Milesi-Ferretti (2007) omit these three economies and the role of foreign shocks in their analysis. Hegerty (2009) investigates the role of foreign shocks using a VAR methodology, but looks only at six recent EU entrants in Central Europe.

As a result, there is ample room in the literature for a study that looks at the macroeconomic determinants of these countries capital flows, while also focusing on the theoretical ambiguities regarding income shocks and "hot money." Using a structural VAR method based on Ying and Kim (2001) to model the effects of macroeconomic variables for Russia, Ukraine and Belarus, this study finds that FDI, portfolio, and Other flows do indeed respond differently from one another, and that the effects of income shocks vary from country to country as well.

METHODOLOGY

The effects of a set of domestic and foreign macroeconomic shocks on capital flows will be assessed with a structural Vector Autoregressive (SVAR) model that will allow for impulse-response and variance decomposition analysis for Russia, Ukraine, and Belarus. Quarterly time series beginning in the mid-1990s are used to study the impacts of these macroeconomic shocks to net FDI, portfolio, and Other investment for each country. Per Ying and Kim (2001) (or a later extension of the paper by De Vita and Kyaw, 2007), each flow is modeled as a function of a set of (unobservable) shocks to Foreign income, the Foreign interest rate, Domestic productivity, Domestic money, and the capital flows themselves:

$$flow_t = f_1\left(u_t^{FY}, u_t^{FR}, u_t^{DS}, u_t^{DM}, u_t^{FLOWS}\right) \quad (1)$$

Because the shocks are unobservable, the model uncovers the underlying structural model through the following VAR model:

$$Y_t = \sum_{i=0}^{\infty} A_i U_{t-i} = A(L)U_t \quad (2).$$

In this extension of the model, which includes disaggregated capital flows in place of the capital account, A_i is matrix of impulse responses to endogenous variables to exogenous shocks and

$$Y_t = \left(\Delta y_t^*, \Delta r_t^*, \Delta y_t, \Delta m_t, fdi_t, port_t, other_t\right)' \quad (3a) ;$$

$$U_t = \left(u_t^{FY}, u_t^{FR}, u_t^{DS}, u_t^{DM}, u_t^{fdi}, u_t^{port}, u_t^{other}\right)' \quad (3b) ;$$

$$A(L) = \sum_{i=0}^{\infty} A_i L^i \quad (3c) .$$

$$\begin{bmatrix} \Delta y_t^* \\ \Delta r_t^* \\ \Delta y_t \\ \Delta m_t \\ fdi_t \\ port_t \\ other_t \end{bmatrix} = \begin{bmatrix} * & 0 & 0 & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 & 0 \\ * & * & * & * & 0 & 0 & 0 \\ * & * & * & * & * & 0 & 0 \\ * & * & * & * & * & * & 0 \\ * & * & * & * & * & * & * \end{bmatrix} \begin{bmatrix} u_t^{FY} \\ u_t^{FR} \\ u_t^{DS} \\ u_t^{DM} \\ u_t^{fdi} \\ u_t^{port} \\ u_t^{other} \end{bmatrix} \quad (4) .$$

The structural model makes use of a set of long-run restrictions (following the Blanchard-Quah decomposition) that stipulates the following: Foreign variables can only be affected by foreign shocks, and monetary shocks do not have any long-run impact on domestic output. In addition, capital-flow shocks do not affect any other variable. The three capital flows are ordered as in (3a) or (3b) because FDI is thought to be more stable than portfolio investment, and Other investment is shown (using the sample standard deviations of the time series) to be more variable than portfolio investment. Thus, the *A(L)* or *A(1)* matrix is as above. Except for the flows, all variables are in logs, as well as first differences. The flow series are in levels, but deflated by each country's nominal GDP for the sake of stationarity (see below). Using the software *JMulTi* (see Lütkepohl and Krätzig, 2004), impulse-response functions and forecast error variance decompositions are obtained for capital flows for each country.

RESULTS

Quarterly data from the International Financial Statistics of the IMF are used in this study. The data and exact time spans are explained further in the Appendix. The Phillips-Perron stationarity test is first performed on each variable; the results are provided in Table 1. Capital flows are tested using two deflators that are standard in the literature: each country's GDP deflator (to capture real flows in domestic currency) or GDP itself. While the macroeconomic variables are first-difference stationary, or I(1), the test suggests that the flows as shares of GDP show more evidence of stationarity in levels. A SVAR(1) is then estimated for each country using flows deflated by GDP; it was also estimated using flows in terms of real national currency, but the results are very similar. While the capital flows themselves can influence each other, only the impulse response functions (IRFs) for shocks to the four main macroeconomic variables are presented. IRFs with 8-quarter horizons, as well as bands of ±1.96 standard errors (representing 95 percent confidence intervals), are given in Figure 2.

We first look at Russia. Our main finding is that non-FDI flows, particularly portfolio investment, react more strongly to macroeconomic shocks. Some of these effects are expected: A decrease in the European interest rate results in an increase in portfolio investment (and an increase in the rate would reduce it), implying that foreign capital might seek out Russia as a profitable place to invest. On the other hand, the effect on FDI is small, but becomes positive after about a year. This suggests that Russian FDI is more strongly determined by structural and institutional factors, rather than simply the interest rate.

The three types of flow clearly seem to behave differently with regard to monetary shocks. Portfolio investment sees a decrease a few months after an increase in the domestic money supply. Investors may be withdrawing these short-term investments as inflation or deterioration in the Russian economy becomes apparent. At the same time, FDI and Other investment increase after a positive monetary shock, but to a lesser degree.

Income shocks appear to have an influence on all types of investment, but again, the effects on non-FDI are stronger. Foreign shocks reduce FDI, suggesting that people in other countries may be choosing to save the gains from increased productivity, or invest them at home, rather than make long-term investments in Russia. On the other hand, portfolio investment increases after such an income shock. This again highlights the differences between "hot" money and FDI, and the fact that foreigners might make more short-term investments after they experience growth in their own economies. Investors may be less willing to commit to long-term projects in Russia.

Table 1. Phillips-Perron Stationarity Tests.

		Real (GDP Deflator)		*Share of GDP*	
Country		**Level**	**1st Diff.**	**Level**	**1st Diff.**
Russia	FDI	-2.839	-14.369	-3.904	-15.043
	PORT	-5.922	-14.419	-5.512	-13.454
	OTHER	-4.244	-14.825	-5.464	-15.475
	M	1.227	-7.348		
	Y	1.182	-5.251		
Ukraine	FDI	-4.984	-14.002	-5.472	-14.013
	PORT	-5.79	-16.272	-6.513	-16.708
	OTHER	-1.737	-9.274	-2.837	-11.342
	M	-0.329	-6.203		
	Y	0.414	-7.183		
Belarus	FDI	-3.664	-16.258	-4.483	-13.748
	PORT	-5.673	-11.759	-5.566	-11.979
	OTHER	-4.937	-12.48	-5.29	-12.922
	M	0.285	-7.637		
	Y	0.87	-6.503		
Germany	r*	-1.304	-5.065		
	Y*	-2.441	-3.641		

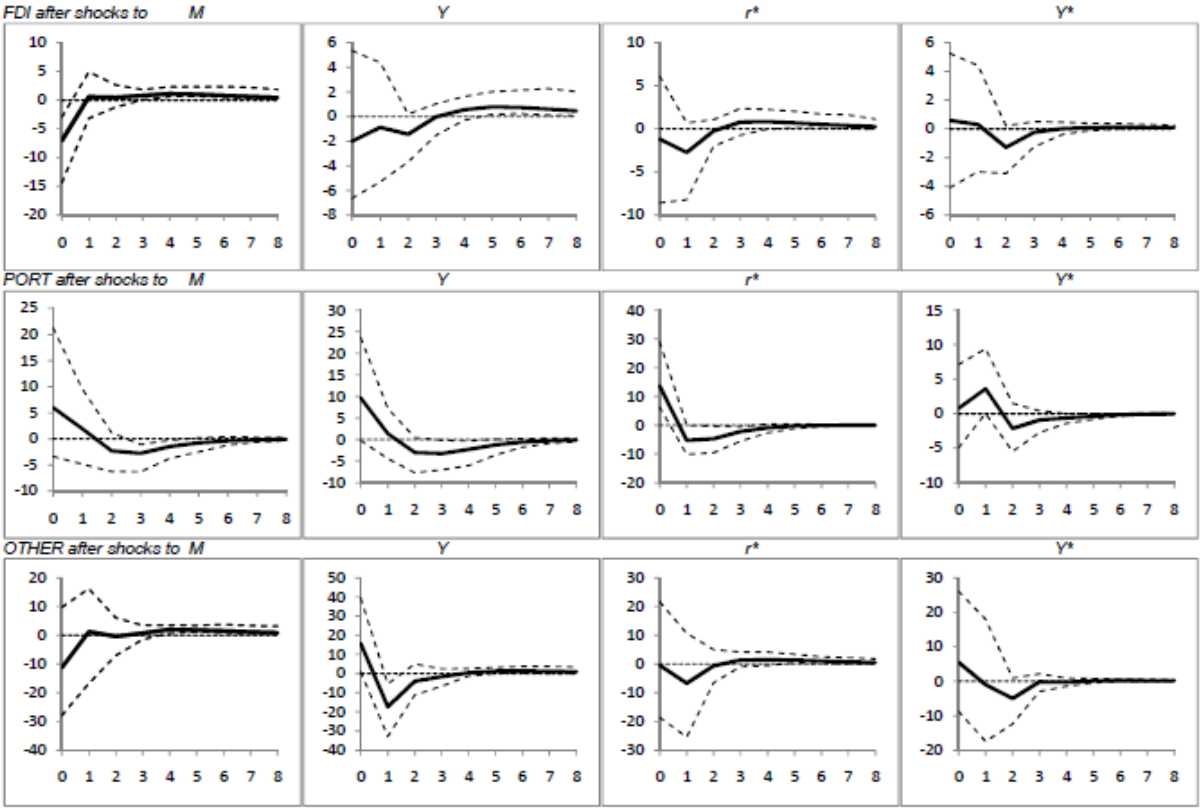
Critical Values: -3.60, -2.90, and -2.60 at 1, 5, and 10 percent, respectively.

The differences among flow types hold for domestic income shocks as well. Growth in Russia reduces all three types of net capital flows in the short term. Portfolio investment experiences a sustained

and significant reduction, which could represent a capital outflow to havens abroad. FDI, however, registers a significantly positive effect after a year. This again suggests that investors are less willing to undertake FDI in Russia; only sustained growth might encourage them to do so.

In Ukraine, however, the effects of these variables are weaker for all flow types. The foreign interest rate has little effect, except on FDI, which responds positively to an interest-rate increase. This finding is difficult to explain, but it is plausible that the "wealth effect" is responsible. The only other significant effect of note is that of domestic monetary shocks, which tend to reduce FDI. This is highly plausible, since Ukraine's experience with inflation will make investors less confident in the country's prospects for macroeconomic stability. This Ukraine's local macroeconomic environment plays more of a leading role in attracting foreign capital than was the case in Russia, and that the country's relative financial isolation makes it less sensitive to global shocks in general.

Figure 2a: Structural VAR Impulse-Response Functions (With ± 1.96 S.E. Bands), Russia.

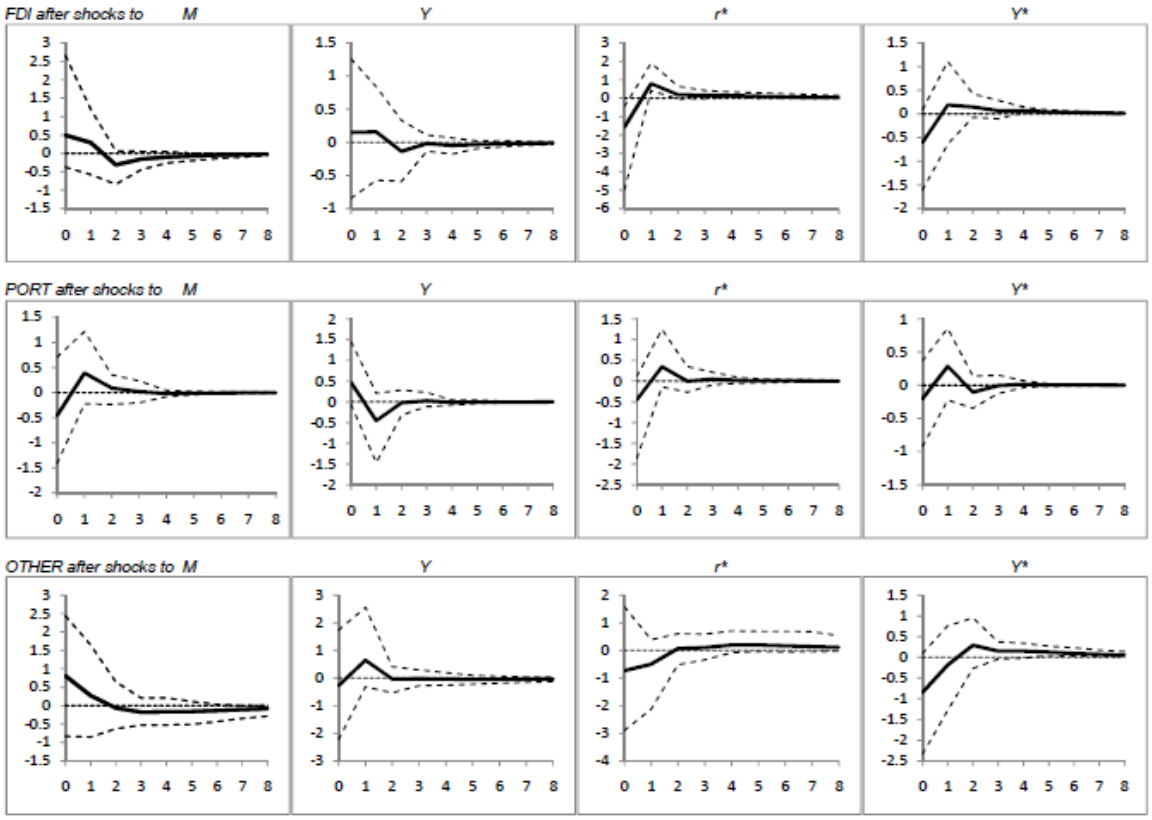


Belarus shows a stronger response to these shocks. Two key results are noted: First, FDI increases after a decrease in the foreign interest rate, while Other investment goes the other direction, and portfolio investment is not affected. Belarus' relatively closed economy means that "hot money" (other than lending) is not drawn in as with Russia. Second, FDI responds negatively to a domestic income increase, while portfolio investment registers a positive effect. The small size of these flows make these results

more difficult to interpret, but this hints at the role of domestic macroeconomic factors (including investor confidence) driving investment in Belarus. Improving economic fundamentals might be more influential in attracting capital into Belarus than are more traditional factors such as returns on investment.

To further assess the influence of each variable, we turn to the forecast error variance decompositions (FEVDs) for the flows in each SVAR. They are reported in Table 2 at one and four quarters. In general, they confirm many of the conclusions from our impulse-response analysis; they also show that the contributions of many of these variables are time-varying.

Figure 2b: Structural VAR Impulse-Response Functions (With ± 1.96 S.E. Bands), Ukraine.



The foreign interest rate makes a large contribution to the forecast error of Russian portfolio investment—this value is as large as 18 percent after four quarters. Domestic income also makes a large contribution to Russian portfolio and Other investment (about nine percent). These results further suggest that investment inflows are helped by relatively high interest rates, while income growth (a domestic factor) might prompt capital (out)flows. In Ukraine, most of the variance of the flows is from the flows themselves, with one main exception: that of *r** on FDI. This matches the key result from the IRFs. In Belarus, the foreign interest rate has a fairly large effect on Other investment, as well as a lesser effect on FDI. In addition, domestic income is a contributing factor in net portfolio flows, corroborating the idea that economic growth in Belarus helps make the country a more attractive place to invest.

Figure 2c: Structural VAR Impulse-Response Functions (With ± 1.96 S.E. Bands), Belarus.

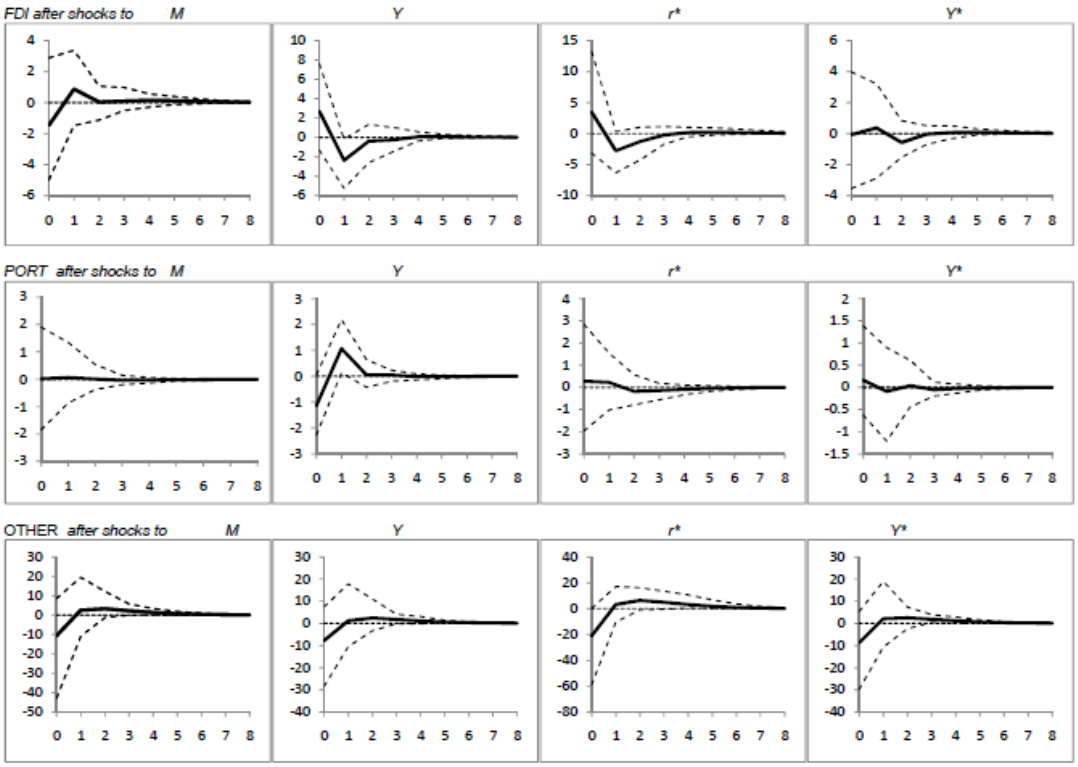


Table 2. Forecast Error Variance Decompositions.

Russia								***Ukraine***							
FDI	*Other*	*Port*	*FDI*	*M*	*Y*	*r**	*Y**	***FDI***	*Oth*	*Port*	*FDI*	*M*	*Y*	*r**	*Y**
1	0.35	0.01	0.52	0.10	0.01	0.00	0.00	1	0.18	0.00	0.62	0.02	0.00	0.16	0.02
4	0.54	0.01	0.36	0.06	0.01	0.01	0.00	4	0.29	0.01	0.50	0.02	0.00	0.16	0.02
Portfolio								***Portfolio***							
1	0.01	0.69	0.02	0.03	0.08	0.17	0.00	1	0.01	0.69	0.02	0.03	0.08	0.17	0.00
4	0.01	0.66	0.02	0.04	0.08	0.18	0.01	4	0.01	0.66	0.02	0.04	0.08	0.18	0.01
Other								**Other**							
1	0.53	0.02	0.36	0.03	0.06	0.00	0.01	1	0.53	0.02	0.36	0.03	0.06	0.00	0.01
4	0.55	0.02	0.30	0.02	0.09	0.01	0.01	4	0.55	0.02	0.30	0.02	0.09	0.01	0.01

Belarus							
FDI	*Other*	*Port*	*FDI*	*M*	*Y*	*r**	*Y**
1	0.35	0.01	0.52	0.10	0.01	0.00	0.00
4	0.54	0.01	0.36	0.06	0.01	0.01	0.00
Portfolio							
1	0.01	0.69	0.02	0.03	0.08	0.17	0.00
4	0.01	0.66	0.02	0.04	0.08	0.18	0.01
Other							
1	0.53	0.02	0.36	0.03	0.06	0.00	0.01
4	0.55	0.02	0.30	0.02	0.09	0.01	0.01

Overall, we find that the IRFs and FEVDs support the same conclusions. Each country—and each flow—responds to different macroeconomic "push" and "pull" factors. Russia, with the largest flows in dollar terms, sees outflows of all types of capital after a domestic income increase. Increases in the foreign interest rate, however, make Russia a more attractive destination for portfolio investment. Russian income shocks also have an effect on its capital flows. In addition, while Ukraine shows comparatively little response to the main macroeconomic determinants, Belarus might see its relatively small capital inflows increase if its macroeconomic fundamentals—particularly economic growth—improve.

CONCLUSION

While not as attractive as some other emerging markets, the former Soviet republics of Russia, Ukraine, and Belarus have drawn in a growing amount of foreign capital in recent years. Since these flows can be highly destabilizing, it is important to study their macroeconomic determinants, and responses to foreign and domestic shocks. Since the underlying theory suggests that the effects of these shocks can be either positive or negative, they must be studied empirically.

This study examines quarterly time series of each country's FDI, portfolio, and Other investment, from 1995 to 2008, and finds that each country and each flow do indeed respond differently from one another. The flow series are placed in a Structural VAR model that was introduced by Ying and Kim (2001). Applying an appropriate set of long-run restrictions, the flows are modeled as functions of domestic monetary and income shocks, as well as foreign interest-rate and income shocks.

Impulse responses and variance decompositions uncover certain key results. First, a distinction must be made between portfolio investment and FDI flows. Second, the impact of macroeconomic shocks can vary over time. Finally, each country shows a unique response for each type of flow. Many of these effects are surprising, but can explain specific issues within each country.

In Russia, for example, certain flows follow what might be expected intuitively. Portfolio investment responds positively to a decrease in European interest rates, suggesting that investors are attracted to a higher rate of return. At the same time, they also increase along with increased foreign income, suggesting that foreigners are choosing to invest this income abroad. On the other hand, these flows decrease after a positive shock to home income. This implies that capital flight is indeed an issue, particularly if Russia continues its recent growth. The differences between the flow types support the "hot money" hypothesis: FDI responds only to income shocks—but only weakly, and temporarily after about two quarters.

The other two countries show unique responses as well, both with regard to the other countries' results and across each type of flow. Ukraine registers very little response for any flow, perhaps due to the fact that it is not yet very well integrated with Western Europe. Belarus might see portfolio inflows increase as a result of economic growth, but this effect does not hold for FDI. As a result, further studies of these countries' capital accounts, now that sufficient time series are available, should take care to pay attention to the country-specific differences between "hot" portfolio flows and FDI.

These results have important policy conclusions. First of all, the region should be aware of the fragility of capital inflows and the possibility of a damaging outflow—but this potential is greater for Russia, and for non-FDI investment. Secondly, this study underscores the fact that these countries cannot be considered a homogenous unit. While Russian investment most closely patterns the "typical" pattern by which foreign capital seeks the highest rate of return, Belarus' closed economy and Ukraine's experiences with partial reform and hyperinflation mean that their capital inflows are driven more by local determinants rather than global ones.

REFERENCES

Chuhan, P., G. Perez-Quiros, and H. Popper. 1996. "International Capital Flows: Do Short-Term Investment and Direct Investment Differ?." World Bank Policy Research Working Paper 1669.

Claessens, S., M.P. Dooley, and A. Warner. 1995. "Portfolio Capital Flows: Hot or Cold?." *The World Bank Economic Review* 9(1): 153-174.

De Vita, G. and K.S. Kyaw. 2008. "Determinants of Capital Flows to Developing Countries: a Structural VAR Analysis,." *Journal of Economic Studies* 35(4): 304-322

Garibaldi, P., N. Mora, R. Sahay, and J. Zettelmeyer. 2001. "What Moves Capital to Transition Economies?." *IMF Staff Papers* 48: 109-45.

Glick, R. and K. Rogoff. 1995. "Global Versus Country-Specific Productivity Shocks and the Current Account." *Journal of Monetary Economics* 35: 159-92.

Hegerty, S.W. 2009. "CapitalFlows to Transition Economies: What Is the Role of External Shocks?,." *Economics Bulletin* 29(2): 1360-1373.

Lane, P.R. and G.M. Milesi-Ferretti. 2007. "Capital Flows to Central and Eastern Europe." *Emerging Markets Review* 8: 106-23.

Lipschitz, L., T. Lane, and A. Mourmouras. 2006. "Capital Flows to Transition Economies: Master or Servant?,." *Czech Journal of Economics and Finance* 56: 202-222.

Lütkepohl, H. and M. Krätzig, eds. 2004. *Applied Time Series Econometrics*, Cambridge, Cambridge University Press.

Melecky, M. 2005. "The Impact of Current Account Reversals on Growth in Central and Eastern Europe." *Eastern European Economics* 43: 57-72.

Ötker-Robe, I., Z. Polanski, B. Topf, and D. Vávra. 2007. "Coping With Capital Inflows: Experiences of Selected European Countries." *IMF Working Paper* #07/190.

Sarno, L., and M.P. Taylor. 1999. "Hot Money, Accounting Labels and the Permanence of Capital Flows to Developing Countries: An Empirical Investigation,." Journal of Development Economics 59: 337-364.

Von Hagen, J. and I. Siedschlag. 2008. "Managing Capital Flows: Experiences from Central and Eastern Europe." ADB Institute discussion paper number 103.

Ying, Y.-H. and Y. Kim. 2001. "An Empirical Analysis on Capital Flows: The Case of Korea and Mexico." *Southern Economic Journal* 67: 954-68.

APPENDIX: DATA SOURCES AND DEFINITIONS

All data are taken from the International Financial Statistics of the IMF. Quarterly GDP series are seasonally adjusted using the Census X-12 procedure. The data spans for the estimations are as follows (difference terms are constructed using data beginning in the previous quarter):

Russia: 1995q2-2008q4; Ukraine: 1996q3-2008q4; Belarus: 1997q1-2008q4

Net capital flows are measured in shares of GDP for each country. They include net FDI (inward minus outward), as well as portfolio and Other investment (liabilities minus assets). These are taken from the Balance of Payments section of the IFS database. As a secondary measure, net flows (in local currency) are also deflated by each country's GDP deflator (2005 = 100).

The explanatory variables are as follows:

M = domestic M2, divided by the country's GDP deflator (in logs)

Y = domestic real GDP (in logs)

r^* = Euro Area Interbank Rate

Y^* = German real GDP (in logs)

In the SVAR, all explanatory variables are entered in first differences.

THE DECLINING EXCHANGE RATE: IMPACT ON THE U.S. ECONOMY 2000-2009

JOHN J. HEIM*

ABSTRACT

Using a simplified Klein/Fair econometric model of the U.S. economy, estimated using 1960 – 2000 data, the paper finds that the 12.9% dollar decline 2000-2009 had a positive effect on exports, but mildly negative effects for domestically produced investment and consumer goods. The estimated overall negative effect on the GDP is modest: 1.7% over the nine years, or about a fifth of a percent per year. It is estimated this decline in the dollar reduced the trade deficit $140.7 billion. This decline is estimated to have increased U.S. net asset position by an $88.6 billion.

1. INTRODUCTION

A decline in U.S. real exchange rate may make goods Americans import more expensive, thereby reducing real income. This *income effect* may reduce U. S. demand for both domestic and imported goods. The decline may also cause a *substitution effect* by making imports more expensive, shifting demand toward cheaper American goods. Also, the cheaper U.S. dollar may have a positive income effect by increasing American exports. The paper also estimates the extent to which exchange rate – induced price increases in imports reduce the trade deficit and thereby reduce transfers of U.S. assets to other nations or their citizens, as required to pay for trade deficits.

2. METHODOLOGY

This study examines how real declining exchange rates 2000 – 2009 have affected U.S. demand for domestic and imported goods. It uses a seven behavioral equation model of the U.S. economy (three consumer demand equations, three investment demand equations and an export demand equation) to estimate the GDP and its components. The three consumption equations are for total, domestically produced and imported consumer goods, the three investment equations are for total, domestically produced and imported investment goods. The seventh equation estimated export demand. The econometric approach is patterned after the more detailed (30 behavioral equations) demand – driven econometric models of Ray Fair (2004).

A significant difference between this study and Fair's models is the way in which autocorrelation is treated. Generally, here it is dealt with by first differencing data. In Fair, it is dealt with by leaving the data in levels and using standard autocorrelation control AR(i) variables.

All data used in those studies was taken from the Council of Economic Advisors' statistical appendix to the *Economic Report of the President, 2002, 2001 and 1997.* Exchange rate values 1960 - 1970 were

* Department of Economics, Rensselaer Polytechnic Institute

assumed constant at 1970 levels, per the Bretton Woods protocols. All data are expressed in real 1996 dollars, or converted to same using the GDP deflator in Table B3 of the President's Report.

Each regression below shows the estimated marginal effect (regression coefficient) for the explanatory variables, the t statistic associated with it, the percent of variance explained and the Durbin Watson autocorrelation statistic. Depending on the particular regression test and the number of lags used, our sample size was 36-38 observations from the 1960-2000 period. With this number of observations, throughout the remainder of the paper, marginal effects with a t-statistic of 1.8 are significant at the 8% level, 2.0 are significant at the 5% level and t-statistics of 2.7 are significant at the 1% level

Because of the simultaneity between the total consumption variable (C) in the GDP accounts, or its component part, domestic consumer goods (C_D), and income (Y) inherent in these equations, two stage least squares estimates of disposable income $\Delta(Y\text{-}T_G)_0$ were used. The remaining right hand side variables were used as first stage regressors. Newey-West heteroskedasticity corrections were also made, generally improving t - statistics. Two Stage least Squares was also used with the investment equations because of simultaneity between investment and the economy's growth rate (the accelerator variable).

Following Heim (2007), we then take as our definition of consumer goods and services imports all imports except for imports of capital goods and industrial supplies and materials. These definitions appear to be reasonable, if not exact, given the data available. Separate regressions were then run on total consumer demand, and separately for imported consumer goods alone. Results for the imports equation were subtracted from the results for the total consumption (C) equation, to estimate demand for domestically produced consumer goods. As noted earlier when discussing autocorrelation, the coefficients obtained in this manner (arithmetically) for each variable are exactly the same as those obtained statistically by regressing these same determinants on domestically produced consumer goods ($C\text{-}M_{m\text{-}ksm}$).

Investment imports were defined using the same process as imports of capital goods plus imports of industrial supplies and materials (M_{ksm}), i.e., total imports minus consumer imports.

Preliminary testing suggested that exchange rates have some lagged effects that go back as far as three years ago, so the average exchange rate for those years (XR_{Av0123}) was used.

3. THE CONSUMER DEMAND MODEL:

Heim (2008b) found that regression results on a modified Keynesian function of the following type explained about 90% of the variance in consumer spending in the 1960 - 2000 period:

$$C = \beta_1 + \beta_2 (Y\text{-}T_G) + \beta_3(T_G - G) - \beta_4 (PR). + \beta_5 (DJ)_{-2} + \beta_6 (XR)_{AV0123} \quad (1)$$

where

$(Y\text{-}T_G)$ = Total income minus taxes

$(T - G)$ = The government deficit

PR = An interest rate measure, the Prime Rate, for the current period.

DJ_{-2} = A wealth measure, the Dow Jones Composite Average, lagged two years

XR_{AV} = The trade - weighted real exchange rate (XR), averaged four years.

Our regression findings for consumer demand model are as follows:

Demand for Domestically Produced Consumer Goods

$$\Delta(C- M_{m\text{-}ksm})_0 = .55\Delta(Y-T_G)_0 + .19\Delta T_{G(0)} + .24\ \Delta G_0 - 1.92\ \Delta PR_0 + .28\Delta DJ_{-2} - .20\ \Delta XR_{AV0123} \quad R^2=74\% \quad (6)$$

(t) (16.2) (1.5) (1.3) (-0.6) (1.9) (-0.2) D.W.= 1.8

4. THE INVESTMENT DEMAND MODEL

The demand for Investment goods may also decline when the exchange rate declines, lowering real business income and raising import prices. How much of the decreased will be for domestic goods compared to imports depends on the marginal propensities to invest (MPI_D or MPI_M) in those goods in response to a change in the economy's real growth rate (i.e., the "accelerator effect") caused by a declining exchange rate. A secondary decrease in Investment may occur due to multiplier effects of the original change, reducing savings, causing increased crowd out effects.

The variables included in the investment equations are

ΔACC = An accelerator variable $\Delta(Y_t - Y_{t-1}) = \Delta\ GDP_t$

ΔDEP = Depreciation,

ΔCAP_{-1} = A measure of last year's capacity utilization level

$\Delta PROF_{-2}$ = A measure of business profitability two years ago

The other variables have the same meanings they had in the consumption equations. The parameters in this investment demand model were estimated to be:

Demand For Domestically Produced Investment Goods

$$\Delta(I-M_{ksm}) = .24\Delta ACC + .91\Delta DEP - .15\Delta CAP_{-1} + .45\ \Delta T_G - .47\Delta G - 9.59\Delta r_{-2} - .40\ \Delta DJ_{-2} + .47\ \Delta PROF_{-2} + 5.37\ \Delta XR_{AV0123} \quad R^2=.88 \quad (11)$$

t= (7.8) (3.0) (-0.4) (6.0) (-2.9) (-7.3) (-1.9) (4.1) (4.1) DW =2.1

5. THE EXPORTS DEMAND MODEL (USING THE REAL BROAD EXCHANGE RATE INDEX)

There is also an increase in income that occurs because of the increase in exports associated with the decline of the exchange rate. A rough estimate of this effect can be obtained by regressing exports on the 4-year average exchange rate above and the growth in the American GDP over the 1960-2000 period. The income variable serves as a proxy for the growth in our major trading partners' incomes over this period. The results are as follows:

$$\Delta X_0 = .12\ \Delta(Y_{AV12}) - 2.86\ \Delta XR_{AV0123} + .68\ AR(3) \quad R^2= 53\% \quad (12)$$

(t) (5.3) (-2.6) (5.4) D.W.= 2.1

6. THE TAX GROWTH MODEL

Part of tax growth is exogenous, i.e., varies with legislative changes in tax rates. However, part is endogenous, i.e., dependent on income growth from year to year. The results of this regression are:

$\Delta T_G = .26\ \Delta(Y)$ $R^2 = 47\%$ (13)

(t) (7.7) D.W.= 1.4

Both the consumption and investment equations above show a positive effect on demand of an increase in tax revenues, presumably by reducing crowd out caused by government deficits. Therefore, in calculating the full effects of a rise in real income due to exchange rate changes, it is important to also measure the secondary boost to income resulting from additional taxes collected as income grows. We might also define tax changes that are government - enacted, i.e., exogenous, as approximately ΔT_{EX} , where

$\Delta T_{EX} = \Delta T_G - .26\ \Delta Y$ (or) $\Delta T_G = .26\ \Delta(Y) + \Delta T_{EX}$ (14)

We say "approximately, because T_{EX} also contains the regression error term.

7. A MODEL FOR CALCULATING MULTIPLIER, ACCELERATOR AND CROWD OUT EFFECTS OF EXCHANGE RATE CHANGES

To illustrate how these terms are used further below, the following definitions of the multiplier and accelerator are presented, using simplified versions of our above consumption and investment equations for ease of exposition:

The GDP (Y) is comprised of consumer goods (C), investment goods (I), goods and services produced for the government (G) and net exports (X-M):

$Y = C + I + G + (X-M)$ (15)

Domestic demand for all consumer goods, imported and domestically produced, can be written:

$C = (c_0 + m_0) + (c_1 + m_{c1})(Y-T_G) + (c_2 + m_{c2})\ T_G + (c_3 + m_{c3})\ G$ (16)

where $(Y- T_G)$ is total income generated producing the GDP minus total taxes; $c_1 + m_{c1}$ are the marginal propensities to consume domestic and imported goods, T_G and G represent the variables measuring the extent to which consumer credit is crowded out by the government deficit. The disaggregated form of the deficit is used (T_G ,G separately) instead of just (T_G - G) because testing above indicates that the effects of the two variables on crowd out are different.

Demand for investment goods in this simple model of the economy might be described as

$I = I_0 + (I_1 + m_{I1})\ \Delta Y - (I_2 + m_{I2})\ r + (I_3 + m_{I3})\ T_G + (I_4 + m_{I4})\ G$ (17)

where ΔY is a Samuelson "accelerator" variable, indicating I grows (accelerates) in response to the general growth in the economy, r is the real interest rate, $(I_1 + m_{I1})$ are the marginal propensities to purchase domestically produced or imported investment goods ut of a change inY. $(I_2 + m_{I2})$ are the marginal propensities to invest when interest rates change. T_G + G represent the investment credit crowd out variables, again disaggregated, and the marginal impact of crowd out is $(I_3 + m_{I3})$ or $(I_4 + m_{I4})$ depending on whether it is caused by taxes or government spending.

Import demand might be expressed as

$$M = M_C + M_I = m_0 + m_{c1}(Y-T) + m_{I1}\Delta Y - m_{I2} r + (m_{c2} + m_{I3}) T_G + (m_{c3} + m_{I4}) G \quad (18)$$

i.e., the demand for imported consumer or investment goods is driven by the same variables as is domestic demand. Substituting (16), (17) and (18) into equation (15) gives

$$Y = (c_0 + I_0 - m_0) + c_1 (Y - T_G) + I_1 \Delta Y - I_2 r + G + X + (c_2 + I_3) T_G + (c_3 + I_4) G \quad (19)$$

using the marginal propensity to consume domestically produced goods from the Section 3 and 4 regressions above. However, if we separate $I_1 \Delta Y$ into its separate components, $I_1 Y$ and $- I_1 Y_{-1}$, and recollect our current year Y terms, we get a modified multiplier (or multiplier/accelerator) coefficient that combines traditional multiplier and accelerator effects.

Again using our regression results above. we can further augment this function by noting that the tax component (T_G) of the "crowd out" variables in both the consumption and investment equation grows as income grows, as shown in our tax growth model above. Also, our consumption and investment regressions above suggest that a rise in taxes depresses consumption spending by decreasing disposable income -$.57B for each billion increase in T_G, but that the same rise in taxes stimulates consumer spending by +$.20B and investment spending by +$.44B, more than offsetting the negative impact of taxes on disposable income, for a net effect of +$.09B. Hence,

$$(-c_1 + c_2 + I_3) T_G = (-.55 + .19 + .45) T_G = (.09) T_G = (.09)(.26\, Y + T_{EX}) = .02\, Y + .09\, T_{EX}$$

Using this formulation and recombining the Y terms gives a further modified multiplier we will call the "Multiplier/Accelerator/Crowd Out" ("M/A/C") multiplier:

$$\Delta Y = \frac{1}{(1-c_1-I_1-[-c_1+c_2+I_3][.26])} [- c_1 \Delta T_{EX} - I_1 \Delta Y_{-1} - I_2 \Delta r + \Delta G + \Delta X + (c_2+I_3) \Delta T_{EX} + (c_3+I_4) \Delta G] \quad (23)$$

where the numerical value of M/A/C multiplier becomes $\frac{1}{(1-.55-.24-.02)} = 5.26$

This is the multiplier we will use to calculate the effect of a change in the exchange rate on income.

9. THREE METHODS FOR CALCULATING THE IMPACT ON THE GDP OF A CHANGE IN THE EXCHANGE RATE

Three separate methods, all yielding the same results, are used to compute the effect of a change in the exchange rate on the GDP (Y):

Method 1: Use marginal effects estimates from the above domestic investment, consumption and export regressions to estimate the initial change in domestic consumption, investment and the GDP resulting from a one index point drop in the trade weighted exchange rate. Apply the M/A/C multiplier (5.26) to the result

Method 2: Use the method favored in many large scale econometric models of the economy (Fair 2003, Pindyck & Rubinfeld 1991). This involves separately estimating ΔC_D, ΔI_D, ΔG and ΔX (using the equations above), and simply summing the results to get ΔY

Method 3: Formally Construct a Keynesian IS curve, and predict ΔY from its determinants and the multiplier implied by the function. It is a slightly more formal presentation of Method 1.

9.1 Method 1

$ + 0.20B (Billion) - Total Estimated Effect (Positive Substitution Minus Negative Income Effect) Of A One Point Exchange Rate Decline On Demand For Domestically Produced Consumer Goods (C_D)

$ - 5.37B - Total Effect (Negative Substitution plus Negative Income Effect) Of Decline In Demand For Domestically Produced Investment Goods (I_D)

$ + 2.86B - Increase In Exports (X)

$ - 2.31B - Initial Net Decline in Real U.S. Income from 1 point Exchange Rate Decline:

x 5.26 - Multiplier/Accel/Crowd Out (M/A/C)Effect

$ -12.15B - Decline in Real Income (Y) after Multiplier/Accel/Crowd Out (MAC)Effects

- 3.16B - Δ Taxes Due To M/A/C Effect @ Historic .26 Rate (.26*12.15 = 3.16B)

$ - 8.99B - $\Delta(Y-T_G)$ = Decline In Disposable Income Associated With A One Point Decline In The Exchange Rate

To see the impact of decreased credit availability (crowd out) due to decreased tax collections:

- $0.60B = ΔC_D Due to Crowd Out Effect, Caused By Decreased Taxes = (.19)($ -3.16B)
- $0.95B = ΔC_M Due to Crowd Out Effect, Caused By Decreased Taxes = (.30)($ -3.16B)

With this information we can summarize the changes in consumption and saving resulting from the decrease in disposable income of $ 8.99B as follows:

$ - 8.99B $\Delta(Y-T_G)$	$ - 8.99B $\Delta(Y-T_G)$	$ - 8.99B $\Delta(Y-T_G)$
x .55 MPC_D	x .11 MPC_M	x .34 MPS (1 -.55 -.11)
$ - 4.94B ΔC_D (Multip. Effect)	$ - 0.99B ΔC_M (Multip. Effect)	$ - 3.06B Δ Savings (Reduction
$ +0.20B Initial ΔXR_{0123} Effect	$ - 3.03B Initial ΔXR_{0123} Effect	in Domestic Funds Available)
$ -0.60B Crowd Out Effect	$ - 0.95B Crowd Out Effect	
$ -5.34B Total ΔC_D	$ - 4.97B Total ΔC_M	

9.2 Method 2:

From the econometric models in Sections 4, we see three variables through which investment is affected by changes in the exchange rate:

1. the decrease in the accelerator income variable in the investment equation, due to the decrease in GDP (including multiplier effects) caused by the one point decline in XR_{AV0123}
2. the decline in tax collections because of the decline in real income caused by the increase in import prices, and
3. through the one point decline in the exchange rate variable

In this case then, the estimated decline in domestic investment will be

$\Delta I_D = \Delta(I\text{-}M_{ksm})$ =.24 ΔACC +.45 ΔT_G + 5.37 ΔXR_{AV0123} = $ - 9.71B

We can also estimate the decrease in demand for imported investment goods as

$\Delta I_M = \Delta(M_{ksm})$ = .05 ΔACC +.07 ΔT_G - 0.40 ΔXR_{AV0123} = $ - 0.43B

By similar reasoning, we see that the changes in the demand for domestic and imported consumer goods are as follows, using the econometric results from Section 3:

ΔC_D = . 55 $\Delta(Y\text{-}T_G)$ +.19 ΔT_G - (0.20) ΔXR_{AV0123} = $ - 5.34B (same result as method 1)

and

ΔC_M = .11 $\Delta(Y\text{-}T_G)$ +.30 ΔT_G + 3.03 ΔXR_{AV0123} = $ - 4.97B (same result as method 1)

So, by Method 2 we have

$\Delta Y = \Delta C_D + \Delta I_D + \Delta G + \Delta X$ = $ - 5.34 – 9.71 + 0 + 2.86 = $ - 12.19B (~Same result as Method 1)

9.3 Method 3: Using the Formal IS Curve Method

$\Delta Y = \Delta C_D + \Delta I_D + \Delta G + \Delta X$

$= (.55\Delta(Y\text{-}T_G) + .19\Delta T_G – 0.20\Delta XR_{AV0123}) + (.24\ \Delta ACC + .45\ \Delta T_G + 5.37\Delta XR_{AV0123}) + 0 - 2.86\ \Delta XR_{AV0123}$

= $ - 12.15 B (Same as Methods 1 and 2)

10. EXCHANGE RATE EFFECTS ON THE TRADE DEFICIT AND U.S. ASSETS

The estimated decline in the U.S. trade deficit resulting from a one point decline in the exchange rate is the sum of the resulting decrease in imports and the increase in exports

$ 4.97B - Decline in C_M

0.43B - Decline in I_M

. 2.86B - Increase in X

$ 8.26B - Decrease in Trade Deficit Associated with a 1 Point Drop in the Exchange Rate

$ - 3.06B - Δ Savings = (.34 MPS)(- 8.99 $\Delta Y\text{-}T_G$)

$ - 3.06B - Decrease In Domestically Owned Wealth (Savings) Resulting From Exchange Rate - Induced Decline in Real Income

11. CONCLUSIONS

The analysis above indicates that when the Federal Reserve's real broad trade - weighted exchange rate index falls by one point, the results are as follows:

Table 3: Exchange Rate Impact on GDP and Trade Balance (Billions of 1996 Dollars)

	Real GDP	Imports	Exports	. Trade Deficit . Dollars (% of GDP)
Actual 2000 Data	$9224.00	$1532.00	$1132.00	$400.00 (4.34%)
Effect of 1Pt. Drop In XR	9211.85	1526.60	1134.86	$391.74 (4.25%)
Effect of 12.9Pt (12.3%)Drop In XR	9066.75	1462.00	1095.11	$293.45 (3.23%)

However, in the perio*d* 2000 – 2009, The U.S. exchange rate dropped even more significantly. The Real Broad Index dropped 12.9 points (12.3%), from 104.8 to 91.9.

Using the 12.9 point drop in the real broad Index during the 2000-09 period, suggests that this would have been associated with a decrease in the GDP over the 9 year period of $157.25 billion, or 1.7% of the GDP. This drop would also have been associated with a drop in the trade deficit of $106.6 billion. As a percent of the GDP, the trade deficit would drop 1.11 percentage points, from 4.34% to 3.23% of GDP *ceteris paribus*. (In 2005 dollars, the GDP decline would have been $207.6 billion, the trade deficit decline $140.7 billion. Percentage changes would remain the same.)

Using the numbers from Method 2, and multiplying them by 12.9, we can disaggregate the total GDP change into its component parts:

$$\Delta Y = \Delta C_D + \Delta I_D + \Delta G + \Delta X$$

$$-\$157.25 \sim= -\$68.88 - \$125.26B + 0 + \$36.89B$$

The $157.25B decrease in GDP associated with the estimated 12.9 point (or 12.3%) 2000-09 decline in exchange rates, would have resulted in a 1.7% decrease in 2000 - level real GDP, *ceteris paribus*. However, Bureau of Economic Analysis data indicated the real GDP grew 15.7% during the 2000-2009 period. Presumably, had the exchange rate decline not occurred, it would have grown 1.7% more, increasing the average annual growth rate slightly - less than one fifth percent per year from 2.74% to 1.93%.. The actual annual growth rate appears to have been lower than it might have been had the exchange rate not declined, but not much. Thus, the evidence indicates that the cheaper dollar of the 2000-2009 period did have a small negative effect on the U.S. GDP, consumption and investment overall, but these effects were swamped by larger scale macroeconomic events going on at the same time (e.g., post 9/11/01 military build up, increased investment spending).

The $8.26 billion decline in the U.S. Trade deficit associated with a one point drop in the real Broad exchange rate index reduces the need for annual transfers of U.S. assets (including dollars) to foreign ownership. Other transfers are still needed to pay for the remaining trade deficit. Hence, there is a decline in the amount of U.S. owned assets that have to be transferred to the rest of the world to pay for the U.S.'s excess of imports over exports. Subtracting the decrease in U.S. assets (decline in new savings of $3.06 billion) associated with the decline of the exchange rate, we estimate each point decline in the exchange rate increases U.S. ownership of assets $5.2 billion, or $67.1 billion for the nine year period. (In 2005 dollars, $88.6 billion).

REFERENCES

_____. *Economic Report of the President*. Washington. U.S. Government Printing Office. 2002 , 2001 and 1997, 2010, 2009 Reports

Fair, R. 2004. *Estimating How the Macroeconomy Works*, Appendix A: U.S. Model. 2004 Available at http://fairmodel.econ.yale.edu/rayfair/pdf/2003A.PDF

Heim, J. J. 2007. "Does A Strong Dollar Increase Demand for Both Domestic and Imported Goods?" *Journal of International Business and Economics.*7(3): 12-22.

Heim, J.J. 2008a. "The Consumption Function." *Review of Business Research* 8(2): 18-36.

Heim, J.J. 2008b. "What Determines Market Demand for Investment Goods?" *International Journal of Business Strategy* 6(1): 67-80.

Heim, J. J. 2008c. "How Falling Exchange Rates 2000 - 2007 Have Affected the U.S. Economy And Trade Deficit." *RPI Working Papers in Economics, No. 803, Department of Economics, Rensselaer Polytechnic Institute*, January 2008.

Heim, J. J. 2009. " A Method For Separating Income And Substitution Effects Of Exchange Rate Changes On Aggregate Demand." *Journal of International Finance and Economics.* Vol. 9(3).

Jorgenson, D. 1971."Econometric Studies of Investment Behavior: A Survey." *Journal of Economic Literature*, 9(4), pp. 1111-47.

Keynes, J.M. 1964. *The General Theory of Employment, Interest and Money*. New York: Harcourt, Brace and World, Inc.

Mankiw, N. G. 2010. *Macroeconomics.* New York: Worth Publishers. p. xxiii.

Samuelson, P. 1939. "Interactions Between the Multiplier Analysis and the Principle of Acceleration." *Review of Economics and Statistics* 21(2): 75-78.

How Much Does "Crowd Out" Reduce the Effects of Keynesian Stimulus

John J. Heim[*]*

ABSTRACT

This paper econometrically examines whether government deficits financed by borrowing reduce credit availability, thereby "crowding out" business and consumer spending. Deficit variables are added to consumption and investment models and tested to see if they negatively impact private spending, are statistically significant, and increase explained variance. U.S. data for 1960 - 2000 is used. A demand-driven econometric model, patterned after the work of Klein and Fair and containing eight behavioral equations is used to estimate effects. This study finds strong statistical evidence, even from within traditional Keynesian models, deficits crowd out private consumption and investment.

INTRODUCTION

The first question we need to ask when analyzing how crowd out affects the economy, is In what theoretical context should we examine this question? What kind of model shall we use to establish the parameters within which we evaluate the effects of crowd out?

In this study, a demand-driven econometric model, patterned after the work of Klein and Fair and containing eight behavioral equations is used to estimate crowd out effects. Demand models are used to provide the context within which crowd out is evaluated because they (1) "provide the foundation of much of our current understanding of economic fluctuations " (Mankiw (2007), (2) because demand fluctuations appear to have caused the recent economic decline (Romer 2010), and (3) because the fiscal policy prescriptions invariably found in demand models are the ones that lose much of their validity if deficits cause offsetting private consumption and investment spending by reducing private credit availability.

In a typical demand driven model of the economy *without* crowd out, the impact of taxes and government purchases can be derived using the GDP identity:

$$GDP = Y = C + I + G + (X-M) \quad (1)$$

where a simple consumption function is given as a linear function of disposable income (Y-T)

$$C = \beta(Y-T)$$

substituting C into (1) gives

$$Y = [1/(1-\beta)]*[-\beta T + I + G + X-M]$$

[*] Department of Economics, Rensselaer Polytechnic Institute

The clear expectation of Keynesian demand theory is that tax changes are negatively related to the GDP, with a multiplier effect -β /(1-β). Changes in government spending and net exports are related to GDP in the positive direction, with a multiplier effect 1/(1-β) and should when tested, have the same coefficients. In Section 2 below, we will test these expected relationships to see if actual econometric estimates yield the predicted results for variables.

HOW CROWD OUT MAY IMPACT BOTH CONSUMER AND INVESTMENT SPENDING

We can expand this model to include any effects of crowd out on investment spending. Assume a simple investment model in which investment is determined by real interest rates (r) and access to credit, which varies with the government deficit (T-G).

$I = \gamma(T-G) - i_{nt} r$

where gamma (γ) indicates the marginal effect of crowd out (the government deficit) on investment spending, and (i_{nt}) represents the marginal effect of interest rates (r).

If we replace investment in the GDP identity with its hypothesized determinants, we obtain a typical Keynesian IS equation:

$GDP = Y = = [1/1- \beta] [(-\beta+ \lambda+ \gamma) T + (1- \lambda-\gamma) G - i_{nt} r + (X-M)]$

In this IS equation, the normal stimulating impact of tax cuts on the GDP (-β) is offset in part by the effects of deficit – induced changes in credit availability (λ+γ). Tax effects may switch from negative to positive if the crowd out effects (λ+ γ) are larger than the disposable income effect (-β). The effect of a change in government spending is also reduced per dollar of expenditure from (1) to (1- λ-γ) times the spending multiplier (1/1- β). Again, the net exports multiplier effect stays the same, now becoming an even stronger stimulus relative to government spending or tax cuts. Results are shown in Table 1.

Table 1: Effects Of Consumer And Investment Credit Crowd Out On The Effectiveness Of Taxes And Government Spending Stimulus

	Without Crowd Out	With Crowd Out		Without Crowd Out	With Crowd Out
Tax coefficient	(-β)	(-β+ λ+ γ)	Government Spending Coefficient	1	(1- λ- γ)
Tax Multiplier	$\frac{(-\beta)}{(1-\beta)}$	$\frac{(-\beta+ \lambda+ \gamma)}{(1- \beta)}$	Government Spending Multiplier	$\frac{(1)}{(1-\beta)}$	$\frac{(1-\lambda- \gamma)}{(1-\beta)}$

The model we shall use for testing later in this paper is equivalent to the model above, but slightly different in form. The model above was based on the usual formulation of the GDP identity

$Y = C + I + G + (X-M)$

We can alternatively write

$Y = C_{D+M} + I_{D+M} + G_{D+M} + (X-M)$, where $M = C_M + I_M + G_M$, or $Y = C_D + I_D + G_D + X$.

This is an important distinction in calculating multipliers because only spending on domestically produced consumer goods generates the multiplier effect on the GDP. Similarly, for investment, a variable like the Samuelson accelerator is likely to affect spending on both domestic and imported investment goods (I). But accelerator effects will only be felt though the growth in domestic investment (I_D). Hence, the last formulation of the GDP identity may be the better form to use when calculating IS curve parameter estimates, since multiplier effects are more correctly estimated. (We abstract from effects on exports of growth in import demand).

Because the data available to us does not allow division of government purchases of goods and services into purchases of domestic and imported goods, the approximate form of the theory we will actually test is

$Y = C_D + I_D + G + X)$

This then presents the standard model of Keynesian mechanics, with crowd out added. Should the problem exist in reality as well as in theoretical conjecture, testing this model should reveal it. Only science and testing can show theoretical conjectures are consistent with empirical reality. After examining some previous efforts to test crowd out theory, we will test the models above, with and without crowd out.

LITERATURE REVIEW

Opinions in the mass media are popularly held, but are not science. To examine the science has to say, a number of prior professional studies examining this topic were reviewed. Some have been entirely, or principally, been reports on other people's science, i.e., literature reviews. Spencer and Yohe, (1970), in reviewing the literature, found that the dominant view the past two hundred years has been that government deficits cause crowding out. Friedman's work (1978), is principally theoretical. He shows portfolio theory suggests the LM curve may shift in response to an IS shift due to a fiscal stimulus like a government deficit. Therefore crowd out effects are indeterminate theoretically: it depends on how much LM shifts relative to IS. Friedman's found his own empirical tests ,based on money demand models, were ambiguous. Gale and Orszag's work (2004) was principally a review of other work, concluding most studies do show a positive relationship between interest rates and deficits (which may cause crowd out), and that most studies that don't show this relationship are VAR types. They note that VAR projections have been shown to be inferior to projections produced by OMB and DRI (p.152). Their paper does includes some empirical testing of a model of the determinants of a consumption, also indicating crowd out matters. In the model demand was hypothesized to be a function of current and one period lagged Net National Product (NNP), government purchases, taxes, transfer payments, interest payments and the size of the government debt. Results of tests of whether tax cuts stimulated consumption were determined by the sign and statistical significance of the coefficient on the tax variable. A negative and statistically significant coefficient on the tax variable was taken as evidence tax cuts were not just saved;

i.e., that Ricardian Equivalence does not hold. (However, the same finding affirms traditional Keynesian theory regarding the stimulus effect of tax cuts without complete crowd out). Other tests indicated a positive relationship between deficits and interest rates, further providing proof for crowd out.

TESTING THE MODEL

If crowd out influences consumption or investment spending

- The deficit should be found to be a statistically significant variable in functions that attempt to include all other factors that can significantly influence consumption and investment, and increases the amount of variance explained in these functions.
- Tests should find the tax variable has a negative coefficient smaller than predicted from traditional Keynesian theory of consumption and investment (which ignores crowd out effects), because reduced spending due to crowd out partially offsets the stimulus effect. If crowd out completely offsets the stimulus, the tax variable will have a positive coefficient.
- The government spending multiplier should be smaller than the exports multiplier if spending generated deficits cause crowd out.

We shall proceed to undertake these tests in the remainder of the paper below.

The single - variable formulation of the government deficit (T-G) should have a positive, statistically significant regression coefficient if crowd out matters. It does in all our tests. However, the single - variable specification implies that if (T) and (G) are estimated separately (the two variable formulation of the deficit), they should be found to have the same coefficients, except for sign: T should be positive, G negative. This is approximately what the investment findings show. However, in the total (C_T) consumption functions, the government spending component (G) typically has a coefficient that is positive in sign, though close to zero in value, and is statistically insignificant. This problem is not found with the tax variable in the same equation. Does this indicate a flaw in the theory of crowd out, i.e., is there some reason crowd out is caused by tax cut deficits, but not government spending deficits? Alternatively, is some variable which affects consumption whose effects are correlated negatively with, increased spending missing from the consumption function? If so, the government spending coefficient should be the net of the two effects, and therefore perhaps close to zero and insignificant. Adding the missing variable (if it exists) to the equation should correct the problem.

There is significant evidence there is such a variable exists: the money supply. The M2 money supply was found to grow in and immediately preceding periods when deficits increased due to government spending. It was not found related to tax cut induced deficits. This M2 growth appears to have offset deficit - related reductions in private consumer credit available. The problem did not appear to affect crowd out coefficients in the investment function, which were close to equal except for sign, as expected.

Heim (2009c) expanded the range of variables found significant in both the consumption and investment equations, finding the following variables also to be significant determinants of consumption:

- **Population Growth (POP)**: a factor found systematically related to growth in consumer demand in addition to the factors previously cited
- **Percentage of Americans 16-24 relative to adults 65 and over ($POP_{16\text{-}24/65}$):** a factor reflecting the fact that younger populations, with lesser incomes, have less to spend on consumer goods, particularly services which account for over half of all consumer spending. They also need to be saving more out of current income for retirement and children's education costs than older adults.
- **Spending on New Housing (HSE):** though an investment item itself, it is an important determinant of consumer durables demand (household appliances).
- **Consumer Confidence Levels (CCI)**: as measured by the Conference Board's Monthly consumer survey (Conference Board, 2009), added because consumer confidence was strongly related to consumer spending, even controlling for income and wealth.

In addition, preliminary testing in this study indicated that that lagged values of the savings components of the M2 money supply, when added to the consumption function, significantly added to the consumption function's explained variance, and restored the negative sign and magnitude of the coefficient on government spending to levels closer to those predicted from one variable (T-G) formulations of the deficit.

- **M2 Money Supply**: Testing indicated that two or three year average M2, particularly the non-M1 parts (savings account deposits, small CD's, money market mutual funds held by individuals and money market deposit accounts) were systematically related to consumer spending. This build up of savings (liquidity) prior to a spending - generated deficit was systematically related to the deficit's effect on consumption. M1 was not found significant. This non-M1 component of M2 probably reflects a dimension of consumer wealth not picked up by our use of the stock market average. It appears that both affect consumption after a lag.

THE COMPLETE MODEL

If the hypothesized IS curve is retested deleting the housing variable (HSE), results are even stronger. The deletion of the housing variable (HSE) seems reasonable on the grounds that though appropriate in the consumption function, we have already accounted for it by including the determinants of investment in the IS curve function. (Housing is the second largest component in investment data).

For the 2 variable deficit model, we have the following consumption function: (the "complete model"). It is estimated without the HSE variable because we use these C estimates (and later I estimates) to predict the IS curve. But HSE is about a third of total I, and the determinants of I are already fully included in the determinants of GDP in the IS function, therefore, including HSE would duplicate variables already in the formulation, and create simultaneous equations bias because HSE is a component of the dependent variable GDP.

Domestically Produced Consumer Goods

$\Delta C_D = .34\Delta(Y-T_G)$ **$+.27\Delta T_G$ $-.74\Delta G$**$-5.56\Delta PR + .34\ \Delta DJ_{-2} + 2.17\ \Delta XR_{AV} + (NA)\Delta HSE - 668.59\Delta POP_{16} + .013\Delta POP + .36\Delta ICC_{-1} + 46.94\Delta M2_{AV}$

(t =) (6.4) **(3.1)** **(-3.2)** (-2.0) (1.4) (2.5) (-2.2) (4.0) (1.1) (5.6)

R^2=88.0% D.W.= 2.0

(C_D - No Crowd Out)

$\Delta C_D = .43\Delta(Y-T_G) - .80\Delta PR + .46\ \Delta DJ_{-2} + .09\ \Delta XR_{AV} + (NA)\Delta HSE - 414.54\Delta POP_{16} + .006\Delta POP + .37\Delta ICC_{-1} + 32.45\Delta M2_{AV}$ R^2=81.3%

(t =) (7.1) (-0.3) (2.3) (0.4) (-1.5) (1.7) (1.1) (4.2) D.W.= 1.8

And the investment equation for the complete model with C-O,

$\Delta(I_D) = .23\Delta ACC + .18\Delta DEP + .18\Delta CAP_{-1}$ **$+.50\ \Delta T$ $-.64\Delta G$** $- 7.54\Delta r_{-2} - .27\ \Delta DJ_{-2} + .44\ \Delta PROF_{-2} + 5.88\ \Delta XR_{AV} + .009\Delta POP$ R^2=.90

(t =) (9.6) (0.6) (0.1) **(7.6)** **(-3.8)** (-6.9) (-1.2) (4.0) (4.8) (3.5) DW =2.3

The revised investment function for the complete model without C-O, also is:

$\Delta(I_D) = .36\Delta ACC + .83\Delta DEP + 2.21\Delta CAP_{-1} - 11.07\Delta r_{-2} + .07\ \Delta DJ_{-2} + .51\ \Delta PROF_{-2} + 4.55\ \Delta XR_{AV0123} - .00\Delta POP$ R^2=.74

(t =) (8.7) (1.5) (1.2) (-3.9) (0.3) (2.9) (4.8) (-0.2) DW =2.5

Predicted IS Curve (No Crowd Out, And No HSE)

$\Delta Y =$ **$-.75\Delta T$ $+1.75\Delta G$$+1.75\Delta X$** $-1.40\Delta PR + .93\Delta DJ_{-2} + 8.12XR_{AV0123} + (NA)\Delta HSE - 725.45\ \Delta POP_{16} + .01\Delta POP + .65\Delta ICC_{-1} + 56.79\Delta M2 + .63\Delta ACC + 1.45\Delta DEP + 3.87\Delta CAP_{-1} - 19.37r_{-2} + .89\Delta PROF_{-2}$

Predicted IS Curve (With Crowd Out, But No HSE)

$\Delta Y =$ **$+.65\Delta T$ $-.57\Delta G$$+1.51\Delta X$** $-8.40\Delta PR + .11\Delta DJ_{-2} + 12.16XR_{AV0123} + (NA)\Delta HSE - 1009.57\ \Delta POP_{16} + .03\Delta POP + .54\Delta ICC_{-1} + 70.88\Delta M2 + .35\Delta ACC + .27\Delta DEP + .27\Delta CAP_{-1} - 11.39r_{-2} + .66\Delta PROF_{-2}$

Actual test Results

$\Delta Y =$ **$+.78\Delta T$ $-.20\Delta G$ $+.61\Delta X$** $-6.69\Delta PR + .30\Delta DJ_{-2} + 4.37XR_{AV} + (NA)\Delta HSE + 505.70\Delta POP_{16} + .05\Delta POP + 1.42\Delta ICC_{-1} + 45.43\Delta M2 + .58\Delta ACC + .16\Delta DEP + 7.97\Delta CAP_{-1} + .04r_{-2} + .21\Delta PROF_{-2}$ R2=97.6%

(t=) (6.0) (-0.7) (-2.1) (2.4) (0.8) (2.4) (1.4) (6.7) (2.8) (3.0) (10.0) (0.3) (2.2) (0.1) (0.8) DW=2.3

9 of 15 IS coefficients in this complete model (without HSE) are better predicted by the 2-variable crowd out model than by the no crowd out model.

For completeness, we also provide estimates of the consistency of IS predictions with actual IS regression results for the 1-variable crowd out case.

Domestically Produced Consumer Goods

$\Delta C_D = .36\Delta(Y-T_G)$ **$+.27\Delta(T-G)$** $-4.49\Delta PR + .24\ \Delta DJ_{-2} + 1.30\ \Delta XR_{AV} + (NA)\Delta HSE - 375.09\Delta POP_{16} + .01\Delta POP + .23\Delta ICC_{-1} + 37.75\Delta M2_{AV}$

(t =) (6.4) **(3.0)** (-1.6) (1.3) (1.3) (-1.9) (2.8) (0.9) (4.9)

R^2=86.0% D.W.= 1.9

Domestically Produced Investment Goods

$\Delta(I_D) = .23\Delta ACC + .16\Delta DEP - .37\Delta CAP_{-1}$ **$+.51\ \Delta(T-G)$** $- 8.22\Delta r_{-2} - .28\ \Delta DJ_{-2} + .44\ \Delta PROF_{-2} + 5.59\Delta XR_{AV0123} + .008\Delta POP$ R^2=.90

(t =) (9.0) (0.5) (-0.3) **(7.6)** (-6.6) (-13) (4.1) (5.6) (3.6) DW =2.3

Predicted IS Curve (No Crowd Out, And No HSE)

$\Delta Y =$ **$-.75\Delta T$ $+1.75\Delta G$$+1.75\Delta X$** $-1.40\Delta PR + .93\Delta DJ_{-2} + 8.12XR_{AV0123} + (NA)\Delta HSE - 725.45\ \Delta POP_{16} + .01\Delta POP + .65\Delta ICC + 56.79\Delta M2 + .63\Delta ACC + 1.45\Delta DEP + 3.87\Delta CAP_{-1} - 19.37r_{-2} + .89\Delta PROF_{-2}$

Predicted IS Curve (With 1-Var. Crowd Out, But No HSE)

ΔY**=+.65**Δ**T + .34**Δ**G+1.56**Δ**X** -7.00PR - .06ΔDJ_{-2} +10.75XR_{AV0123} +(NA)ΔHSE -585.14 ΔPOP_{16} + .03ΔPOP +.36ΔICC_{-1}+58.89ΔM2
+.36ΔACC+ .25ΔDEP - .58ΔCAP_{-1} -12.82r_{-2} + .69$\Delta PROF_{-2}$

Actual test Results (Repeated From Above)

ΔY= **+.78**Δ**T - .20**Δ**G + .61**Δ**X** -6.69ΔPR +.30ΔDJ_{-2} + 4.37XR_{AV} +(NA)ΔHSE +505.70ΔPOP_{16} +.05ΔPOP +1.42ΔICC_{-1}+ 45.43ΔM2

(t=) (6.0) (-0.7) (-2.1) (2.4) (0.8) (2.4) (1.4) (6.7) (2.8) (3.0)

+.58ΔACC+ .16ΔDEP +7.97ΔCAP_{-1} + .04r_{-2} +.21$\Delta PROF_{-2}$ R2=97.6%

(10.0) (0.3) (2.2) (0.1) (0.8) DW=2.3

10 of 15 IS coefficients in this complete model (without HSE) are better predicted by the 1-variable crowd out model than by the no-crowd out model. The empirical tests strongly support the hypothesis that credit crowd out strongly reduces and may completely crowd out the positive effect of fiscal stimulus

However, results do indicate that spending-induced deficits can be offset by increasing the non-M1 components of M2, particularly its savings account and other time deposits, like small CDs, and money market components. To be effective, it must be in the second, third and fourth years preceding the deficit. If done in the deficit year, this implies the M2 offset to crowd out won't be felt until the second, third and fourth years forward.

This summarizes the best science we have been able to bring to bear on the issue of crowd out. The science indicates crowd out systematically reduces or eliminates the anticipated positive impact of Keynesian fiscal stimulus obtained from no-crowd out models.

REFERENCES

Cebula, R.J. 1978. "An Empirical Analysis of the "Crowding Out" Effect of Fiscal Policy in the United States and Canada". *Kyklos*, Vol. 3(3): 424-436

Chan, S. 2010. "Group of 7's Finance Ministers Stand By Stimulus Programs." *The New York Times*, p. A16, February 2.

Consumer Confidence Survey 1967-2009. The Conference Board. Data Available at *The Conference Board.Org*.

Fox, Carl A. *Intermediate Economic Statistics*. New York: John Wiley & Sons. 1968.

Friedman, B. 1978. "Crowding Out or Crowding In? Economic Consequences of Financing Government Deficits". *Brookings Papers on Economic Activity*. Vol. 1978(3): 593-641.

Gale, W.G. and Orszag, P.R. 2004. "Budget Deficits, National Saving, and Interest Rates". *Brookings Papers on Economic Activity*. Vol. 2004(2): 101-187.

Goldberger, A. S. 1961. "Stepwise least Squares: Residual Analysis and Specification Error." *Journal of the American Statistical Association*. 56: 998-1000.

Heim, John J. 2008. "The Consumption Function." *Review of Business Research*, Vol. 8(2): 18-36.

Heim, John J. 2009a. “The Declining American Dollar: Econometric Estimates of Impact On The GDP and Trade Deficit”. *Journal of Applied Business and Economics,* 10(3).

Heim, John J. 2009b. “Consumer Demand For Durable Goods, Nondurable Goods and Services”. *Journal of the Academy of Business and Economics,* Vol. 9 (1).

Heim, John J. 2009c. “The Impact of Consumer Confidence on Consumption and Investment Spending.” *Journal of Applied Business and Economics. Vol.11(2).*

Hicks, John R. 1937. “Mr. Keynes and the Classics: A Suggested Interpretation”. *Econometrica* 5: 147 - 159.

Kuznets, S. 1952. “Proportion of Capital Formation to National Product”, *American Economic Review, Papers and Proceedings,* Vol. 42 (2).

Mankiw, N..G.. 2007. *Macroeconomics.* New York: Worth Publishers. p. xxiii.

Mitra, P. (2006). “Has Government Investment Crowded Out Private Investment In India? *American Economic Review*, Vol. 96(2): 337-341.

Mofidi, A. and Stone, J. 1990. “Do State and Local Taxes Affect Economic Growth?” *Review of Economics and Statistics*, Vol. 72(4): 686-691.

National Bureau of Economic Research 2009. “U.S. Business Cycle Expansions and Contractions”. Available at http://www.nber.org/cycles.html .

Romer, Christiana. 2010. Quoted in Chan, Sewell. “Jobless Rate Tied to Fall in Demand.” *The New York Times*, p. A23, April 18.

Samuelson, P. 1939. “Interactions Between the Multiplier Analysis and the Principle of Acceleration”. Review of Economics and Statistics. Vol. 21(2): 75 - 78.

Spencer, R.W. and Yohe, W.P. 1970. “The Crowding Out of Private Expenditures by Fiscal Policy Actions.” *Federal Reserve Bank of St. Louis Review*. Oct. 1970.

Macroeconomic Policy and Institutional Change in the Age of Limits to Growth

Kent Klitgaard*

ABSTRACT

The early 1970s were a watershed of environmental political economy. Among the most influential work was a 1972 study for the Club of Rome entitled *The Limits to Growth*. The original work and its subsequent updates warned the world of the power of exponential growth and used a computer simulation to project potential scenarios based on the interaction of population, food production, industrial output, resource availability and pollution. Although he study was routinely criticized and even dismissed, the projections are consistent with present data on those variables. The paper develops the interaction between biophysical limits and the internal macroeconomic limits and ends with possible suggestions for living within the earth's carrying capacity.

INTRODUCTION

The years of the late 1960s and early 1970s marked a watershed in thought and action concerning the relation between the environment and the human economy. The nation celebrated the first Earth Day in 1970, and Richard Nixon signed the environmental protection act into law in 1973, beginning a fundamental change in the patterns of government regulation, The peak of domestic (lower 48) oil production peaked in 1970, exactly as predicted by petroleum geologist M. King Hubbert fourteen years earlier in 1956 (when there were only 48 states). The nation began the long slow process of becoming dependent upon stocks of non-renewable foreign fossil fuels. In 1973 this dependency was punctuated by the first "oil shock" since 1956 when Egypt shut off access to the Suez Canal adding to the cost and the time needed to bring Middle Eastern oil to markets in the rich nations of the world (Yergin).

The late 1960s also witnessed the end of the long post war boom as well along with the ability to "fine tune" the economy by means of marginal adjustments in fiscal and monetary policy. The recession of 1974-75 followed on the heels of the 1973 oil price run-up, just as the "double-dip" recession of 1981-82 followed in the wake of the 1979 oil shock precipitated by the Islamic Revolution in Iran. The pattern of oil price run up followed by recession persists to the present day, as the current recession commenced soon after the spot market price of oil reached $147 per barrel in the summer of 2008. The recessions of the 1970s were not of the usual type, where prices fell with aggregate demand. Instead the recession of 1974-75 ushered in the era of stagflation. The efficacy of counter cyclical fiscal policy fell victim to the changed energy base and new political realities. The very notion that government was a vital part of maintaining stability and acting as the vehicle of economic growth was a casualty of the era.

Yet these times also provided the backdrop for a number of articles questioning one of the most hallowed premises of mainstream economics—economic growth. Former AEA president Kenneth

* Economics and Management, Wells College, Aurora, NY 13026

Boulding published, in 1966, “The Economics of the Comings Spaceship Earth.” He argued that the days of unlimited economic growth based on unlimited resources (the cowboy economy) were coming to a close. Rather we were entering the era of the spaceship economy where resources must be shepherded and all wastes must be carefully recycled. This clearly contravened not only liberal and conservative economics, but the majority of socialist thought on the economy. But, as Boulding was fond of saying: “Anyone who believes that an economy can grow geometrically or exponentially inside of a finite ecosystem is either a madman or an economist” (Boulding 1966). In 1971 Nicholas Georgescu-Roegen published *The Entropy Law and the Economic Process*, arguing that economic activity, as the transformation of material inputs into material output, as well as the transformation of finished commodities into trash, was subject to the laws of thermodynamics, especially the second law. While the quantities of energy and matter available to us are constant their qualities decline as high grade energy and matter degrades into waste heat in the process of doing work. Thermodynamic limits imposed economic limits (Georgescu-Roegen 1971). By 1975 in “Energy and Economic Myths” Georgescu-Roegen was calling for the cessation of economic growth itself. In addition 1971 saw the zero economic growth tradition expanded to the literature of the left .Hardesty, Clement, and Jencks published “The Political Economy of Environmental Destruction” in volume 3 #4 of *The Review of Radical Political Economics,* arguing a fundamental contradiction was appearing by the 1970s. The faster the economy grows the faster we deplete resources, increase pollution, and add carbon to the atmosphere, thereby precipitating climate change (this was my first introduction to the subject matter) (Hardesty 1971). Yet if a capitalist economy does not grow depression and large-scale unemployment ensue. In the same issue a paper entitled “Ecology and Class Conflict” England and Bluestone argued that the worst environmental effects were felt by the poorest, both domestically and internationally (England and Bluestone 1971).

In 1973 E.F. Schumacher brought the world *Small Is Beautiful.* In this work he argued that Marx pursued the incorrect path with his commitment to the labor theory of value. Instead of all value being created by labor, nature was the source of wealth, and natural capital should be treated as an original source of value. Interestingly enough, Marx makes a very similar argument in his *Critique of the Gotha Programme*, but Schumacher does not list this work among his references. Schumacher also argues that the commitment towards growth is not only materially destructive but also morally destructive, and advocates the voluntary simplicity of the Buddhist life as a superior pathway towards human satisfaction. In 1973 Georgescu-Roegen’s student, Herman Daly began his quest of advocacy for the need of a steady-state economy. The profession owes Daly a great degree of gratitude for pioneering the pre-analytical vision of an open and growing economy embedded in a finite and non-growing ecosystem, and following that vision with a great deal of theoretical detail. The result is the present-day discipline of ecological economics. Unfortunately the profession largely ignored these early writings. In 1972 a report of a team of economists, environmentalists, and computer scientists commissioned by the Club of Rome issued a report to the Club that was subsequently published as *The Limits to Growth.*

THE *LIMITS TO GROWTH* STUDIES

The report to the Club of Rome originated in the Systems Dynamics Laboratory of the Massachusetts Institute of Technology, Using a program written by Jay Forrester, inventor of random access memory, the team, headed by Donella Meadows projected the effects of exponential growth of five series: population; industrial output per capita; pollution; food production per capital; and resource availability. They introduced four possible scenarios based on the interaction of an exponentially-growing economy and the carrying capacity of the earth to sustain growth. The first scenario is that of continued exponential growth over time, based on the exponential growth of carrying capacity. This forms the cornucopian scenario that all pro-growth theories rely on. The second scenario is that posed by the logistic differential equation. Carrying capacity is limited and fixed. The economy is able to grow exponentially as long as it is distant from the limits of carrying capacity. At some point growth "feels: its limits, and after that inflection point growth increases at a decreasing rate until coming into a steady-state equilibrium at carrying capacity. Alternatively a system could exceed its capacity but learn from the lessons of overshoot and come, eventually, into steady-state equilibrium at carrying capacity Last, and most pessimistically, the system could overshoot its limits and collapse its carrying capacity, finally coming into steady-state equilibrium at a much lower standard of living. Scenarios one and two are purely abstract models, highly unlikely to occur. The Meadows team advocated immediate action to limit economic growth in order to achieve scenario number three instead of scenario number four (Meadows, *et al.* 1972).

The purpose of the initial study was limited. How would these series interact and influence one another through various feedback loops over the next 100 years, or until about 2070. Their computer projections indicated that the availability of non-renewable resources would decline monotonically and exponentially. Action on the scale of human life could replenish neither minerals nor fossil fuels. Once we began to exploit these resources we merely drew down existing stocks. The lack of resources would thereby limit industrial output, which would rise until the mid-21st century then decline. The same phenomenon would exist for food production. Today approximately 7% of the world's industrial energy supply is used to produce fertilizer. Moreover the hybrid grain packages of the "Green Revolution" of the 1960s required an increase in energy inputs: from natural gas to dry crops to mechanical harvesting to fossil fuels to run irrigation pumps, not to mention the fuels needed to transport the crops to market. As energy supply peaks and declines, so to do cost-effective methods of fertilizing and producing crops. Consequently food production peaks and declines by the mid 2000s. As the ability to grow food declines, so too, eventually, does the human population. Finally, the increased use of fossil fuels and petrochemicals escalates pollution until industrial production and food production collapse, at which point pollution abates.

The initial *Limits to Growth* study was not a prediction in the standard sense of the term as used by economists. Rather it was a computer projection based on no fundamental change in economic and social relationships. This is commonly referred to as the "Business as Usual" strategy.

> The "standard" world model run assumes no change in the physical, economic, or social relationships that have historically governed the development of the world system. All

> variables plotted here follow historical values from 1900 to 1970. Food, industrial output, and population grow exponentially until the rapidly diminishing resource base forces as slowdown in industrial growth. Because of natural delays in the system, both population and pollution continue to increase for some time after the peak of industrialization. Population growth is finally halted by a rise in the death rate due to decreased food and medical services (Meadows, *et al.* 1972).

However the study included several runs of the World 3 model under alternative assumptions. One of the most important reflected the concern over technological change. The model assumed a significant (500%) increase in the size of the resource base, but under conditions of exponential growth. Typically the life of a resource is computed by means of a simple static index where one simply divides the remaining resource by current consumption. For example the study cites the remaining availability of chromium reserves (element 24 on the periodic table). In 1972 chromium reserves were estimated to be 775 metric tons while 1.85 metric tons were mined annually. This gives a static index of 775/1.85 or approximately 420 years. However chromium production was growing at the time by 2.6% per year. If consumption is growing exponentially the period would be shorter. A simple computation yields a dynamic time to depletion of a mere 95 years. Moreover a five-fold increase in the size of the resource base increases the time to depletion to only 154 years. Perhaps the greatest message of the *Limits to Growth* study is the power of exponential growth!

On the basis of their computer simulations the Meadows team arrived at three tentative conclusions.

1. If the present growth trends in world population, industrialization, pollution, food production, and resource depletion continue unchanged, the limits to growth on this planet will be reached sometime within the next 100years.
2. It is possible to alter these growth trends and to establish a condition of ecological and economic stability that is sustainable far into the future. The state of global equilibrium could be designed so that the basic material needs of each person on earth are satisfied and each person has equal opportunity to realize his or her human potential.
3. If the world's people decide to strive for this second outcome rather than the first, the sooner they begin working to attain it, the greater will be their chances of success (Meadows 1972:23-24).

The authors completed a 20-year update in 1992 which they entitled *Beyond the Limits.* As the title indicates many of the series they projected in 1972 had already exceeded carrying capacity and had gone into overshoot. The primary question they posed was whether or not human economy and society would realized the natures of overshoot and exponential growth in time to achieve overshoot and oscillation scenario needed to achieve steady-state equilibrium or would the world succumb to the oscillate and collapse scenario. By 2005 they completed a 30-year update. This revision not only updated data sets on resources, population, and food production, provided a more sophisticated analysis of the delays and limits of the early 21st century, but also relied extensively upon the Ecological Footprint model

originally articulated by William Rees and Mathis Wackernagel. They concluded that the overshoot has continued and the world is now more likely to experience overshoot and collapse.

CRITIQUES OF THE *LIMITS TO GROWTH*

Unlike the indifference shown to the aforementioned theories of growth and its limits the report to the Club of Rome was immediately and vociferously subjected to criticism from mainstream and Marxian economists alike. Henry Wallich considered the book "an irresponsible piece of nonsense" (Wallich 1972). As a "new economist" Wallich saw growth as the only method by which the poverty of the "underdeveloped" world could be alleviated. Harvard economist Lawrence Summers declared "There are no...limits to the carrying capacity of the earth that are likely to bind any time in the foreseeable future. The idea that we should put limits on growth because of some natural limit is a profound error" (McKibben 2010). The majority of mainstream critiques rested on the same foundation. The data were limited, which is a proposition that the Meadows team stressed continually throughout the pages of their report. *The Limits to Growth* paid insufficient attention to the role of prices as incentives for environmental improvement. Nobel Laureate Robert Solow devoted the bulk of his 1974 Richard T. Ely Lecture to the belief that sufficient conservation can be achieved by standard neoclassical mechanisms of price incentives and discounting (Solow 1974). However, the Meadows study was directed towards the problem of absolute scarcity. As any student of economics comes to know soon in their career prices reflect only relative scarcities. As Canadian macroeconomic modeler Peter Victor put it, the Meadows project may have understated the role of prices but their critics overestimated them (Victor 2008). Respected environmental economist Allen Kneese criticized the *Limits to Growth* for inattention to the possibilities of technological change. Joan Robinson simply dismissed Forrester as a crank (Robinson 1973).

Critical comments abounded on the left as well, where the study was castigated as an example of Neo-Malthusianism. Malthus has long been despised among the ranks of Marxian economists. He and the Marquis de Condorcet were clearly the first economists to come to grips with the problem of exponential growth in a limited system. If population grew unchecked geometrically while food production grew only arithmetically a critical time would arrive when food production would be insufficient to support the existing population. Famine would soon follow. Condorcet took the radical position that once freed from the inequalities driven by private property the working classes would turn their efforts towards the study of philosophy and the quest towards human perfectibility and away from procreation. Malthus thought that the morally inferior working classes would squander the income debauching at the ale house. Much of the distaste for Malthus stems from his advocacy of "positive checks" that would increase the death rates among the poor.

In other words, if the death of the poor was encouraged the economic surplus devoted to the privileged elite would be sufficient to support the existing organization and institutional structure of society. This position led Marx to dub Malthus "the poisoned parson." Marx also offered his own theory of relative

surplus population entitled “the reserve army of the unemployed.” In the last instance the difference between Marx and Malthus rests in the difference between absolute and relative overpopulation.

Malthus wrote at the very beginning of the industrial revolution, when the application of fossil fuels to industrial uses was in its infancy. Coal did not achieve the status of primary fuels source until 1880. While oil was first drilled in commercial quantities in 1859 it did not replace coal as the primary fuel source for another 100 years. It simply takes a great deal of time to fundamentally alter the primary energy base (Victor 2008). This is a crucial factor in the sense of immediacy felt by advocates of degrowth and the steady-state economy.

The analyses of Marx and Malthus are grounded the analysis in the idea of an economic surplus, or the difference between the value of output and the costs of producing it. The capacity to increase an economic surplus rests on the ability to appropriate an energy surplus. Richard Heinberg provides a framework of five strategies by which humans have historically captured such energy surpluses. The first strategy is that of takeover, where humans simply reserve nature’s photosynthetic capabilities that support myriad species for ourselves. The best example is agriculture, or the growing of field crops, turning a complex ecosystem with multiple trophic layers, or energy pathways, into a simple ecosystem. Any plant in the wrong place is a weed, and any animal that competes with humans is a pest. Both are to be eradicated in the quest for an energy surplus. Takeover is supplemented by tool use, and the productivity of tool use is greatly enhanced by specialization, as per Adam Smith. The penultimate strategy is scope enlargement, made possible by an increase in trade. David Ricardo believed the scope enlargement could proceed more vigorously under a regime of free trade, but in the last instance could fall back on no other way of increasing carrying capacity, at least for England. Transcending the limits of carrying capacity by means of trade remains a staple of neoclassical theology. But trade can only import the carrying capacity of one region into another. Dramatic increases in carrying capacity were made possible only by the final strategy—drawdown. For the vast majority of time humans have lived by appropriating solar flow, from plants, wind, water, etc. By the end of the 19th century we began to drawdown limited terrestrial stocks in the form of coal, oil, and natural gas. This has been the most successful of all human strategies to increase carrying capacity by means of appropriating an energy surplus.

Since Malthus lived and wrote in the era preceding drawdown, and since he opposed scope enlargement by increased trade in food, he could not have predicted the vast numbers that the world could support. For Malthus the critical time at which the exponentially increasing population would overwhelm the food supply was imminent. But the empirical evidence to the contrary is convincing. From 1800 when Malthus wrote to the present day the world population has expanded seven fold, from 1 billion to nearly seven billion. This was enabled primarily by applying cheap oil to the production of food. Natural forms, from crop rotation, to the use of manure could not provide sufficient quantities of nitrogen to allow the mass production agriculture needed to support the burgeoning population. In the mid 1800s islands off the coast of Peru were discovered as a source of nitrogen in the form of bird guano. The deposits of

shore birds, plus the arid climate, resulted in the accumulation of a nitrate-rich source of fertilizer (potassium nitrate). Peru and Chile actually fought a war over access to the islands, but by 1870 the guano was gone—the drawdown strategy reached the limit of resource depletion. Nitrogen hungry farmers turned to the Atacama Desert, a rich source of sodium nitrate. Chile and Peru once again disputed the access to the nitrates and Chile became rich after victory in the nitrate wars. Germany alone imported 900.000 tons per year of Chilean saltpeter, with the United States importing another half a million tons by 1900. By 1940 it was gone.

But the early years of the 20th century saw the development of a viable alternative, when a German chemist by the name of Fritz Haber and an engineer named Karl Bosch developed a process that could create "bread from air." The earth's atmosphere is composed primarily of nitrogen. However the atmospheric form (N_2) cannot be fixed by plants, which require a single atom of nitrogen. N_2 is extremely difficult to break apart and is naturally severed only by lightening, and the bacteria that live on nodules at the base of legumes. Haber and Bosch were able to fix atmospheric nitrogen from ammonia through a combination of heat, pressure and a catalyst. Both won Nobel prizes for their efforts, and the process of expanding agricultural output, and indeed the human carrying capacity, by application of fossil energy had commenced (Hager 2008).

The process continued in the 1960s when plant breeder Norman Borlaug created dwarf varieties of rice that could hold greater quantities of grain in their heads without bending over into flooded paddies and rotting. The rice varieties were subsequently supplemented with wheat. The "Green Revolution" had begun and food output escalated to a level sufficient to feed an exploding population. These new grains however required extensive energy inputs in the form of fertilizers, pesticides and herbicides (they were susceptible to rust and disease from the lack of diversity), as well as irrigation and mechanical harvesting.

But what happens to the ability to grow food when the cheap energy runs out, just as the guano and Chilean saltpeter ran out? Furthermore, if the carbon emission from using the cheap energy in the atmosphere grows to the point where it destabilizes the climate, does that mean another set of biophysical limits has been reached? This was the concern of the Meadows group as early as 1972. Does the evidence available since that time support the hypothesis of biophysical limits to growth?

A NEW CENTURY WITH NEW LIMITS

The biophysical pressures upon the earth's human population have not abated since the 1972 publication of *The Limits to Growth.* Neither has the complexity that gave rise to the interactions of the five series specified by the Meadows group. Every serious scientific measurement indicates that that we are in overshoot. A recent poll of the members of the National Academy of Sciences indicated that all indicated that it was a serious threat to human kind and 97% believed it was of human origins. Specifically the origin of climate change is in the burning of fossil fuels. James Hansen, climatologist at NASA Goddard Space Center estimates that the "tipping point" at which irreversible climate change occurs is an atmospheric concentration of carbon dioxide of 350 parts per million. Current estimates

show a concentration of 389 ppm. Carbon concentrations increase with carbon emissions. For the first eight years of the new century carbon emissions increased at 3.5% annually. A rudimentary knowledge of exponential growth tells one that if this pace were to continue unabated carbon emissions would double in twenty years. If carbon emissions were to grow the time frame would decline. The latest data indicate that the world has not kept that pace. When the world financial system teetered on the brink of collapse and a world recession began carbon emissions declined as fossil fuel consumption fell. Carbon emissions are now rising at a mere 2.5% per year. This indicates a doubling time of 28 years. In essence by reducing our carbon emissions we have bought ourselves an additional eight years to figure out a solution. But the price we have paid so far is a terrible one: increasing unemployment, bankruptcy and an increase in poverty.

In addition the world's oil supply is at, or near, it peak. When the US oil supply peaked in 1970 conditions fundamentally changed. The gap between declining domestic supplies and growth-driven demand was filled in two ways: increased imports and increased drilling. The oil price run-up following the 1973 Saudi boycott caused prices to nearly double. The rising price served as an incentive to increase drilling. Drilling spiked from 2 million feet per year in 1970 to more than 7 million feet per year in 1980. Despite the increased drilling effort, and new discoveries on the North Slope of Alaska, production fell from 7 billion barrels per year in 1970 to less than six billion in 1980. When oil prices collapsed after the early 1980s drilling fell to two million feet per year and output to five billion barrels. (Hall and Day 2009). Moreover oil discovery in the continental United States peaked in 1930. Forty years later production peaked. World oil discovery peaked in 1960. If similar time lags exist on a world scale as they did domestically, we are at or near the peak of world oil production. One cannot extrapolate exactly, as not all oil producing countries keep as accurate records as do US companies.

If one looks at oil production country by country one finds that many of the individual oil producers are at or post peak, just like the United States. The world's second largest oil producer, the former Soviet Union peaked in the 1990s. Iran, Mexico, and Canada have also peaked. Indonesia recently left OPEC because they are no longer an oil exporter. Domestic oil consumption, driven by rising population and income has absorbed the country's additional capacity. Energy analyst Matthew Simmons presents data that indicate domestic consumption, as well as population is rising rapidly in the oil producing Middle East, from 66 million tons of oil equivalent in 1970 to 388 million in 2000. His estimates indicate a further increase, to 1652 by 2030 (Simmons 2000).

As conventional oil supplies peak and decline an oil hungry world searches for alternative sources of hydrocarbons: deep water drilling; tar sands; heavy oils; coal gasification, etc. Unfortunately the new sources will be harder to find, more expensive to drill and process and likely to be is less stable areas of the globe, either climatically or politically. This can be measured by the Energy Return on Investment (or EROI). This is merely a physical measure (in joules or kilocalories) of the ratio of energy out to energy in. In the 1930s the EROI for the newly discovered oil in Texas and Oklahoma was approximately 100:1. Today domestic oil has an EROI of less than 10:1. It simply takes more energy to construct a platform like

the *Deepwater Horizon* than it does to drill down a few hundred feet in the high plains. Having found the best oil half a century ago, the remainder is more difficult to extract and more expensive. The age of oil is far from over, but the age of *cheap oil* is coming to a close.

There are biological as well as physical indicators that the human economy is in overshoot. The ecological footprint, which measures the biocapacity needed to support a given lifestyle and provide food, fiber, and non-fossil energy compared to the biocapacity the earth can produce. The latest measurement shows that the average ecological footprint uses 2.2 hectares of land to support the existing standard of living. However average biocapacity is only 1.8 hectares. The world is in overshoot. The only way one can appropriate more net primary productivity than the world produces is to deplete the stock that produces it. But this is a short-term and limited solution. Currently North Americans consume the amount of food, fiber and biomass that 9.7 hectares produce. If the world was to live at this standard we would need five planets to produce the subsistence. But we only have one!

My point is simple. If you are already in overshoot you cannot grow your way into sustainability.

WHAT THE *LIMITS TO GROWTH* LEFT OUT

To achieve a sustainable human economy, defined as living well within nature's limits, according to the Meadows study we will need to voluntarily reduce our growth rate, and we need to start now, rather than later. Climate scholar and activist put the matter bluntly: "Basically, it turns out, they were right. The Club of Rome, E.F. Schumacher, the MIT crew with their whiz-bang program on their 1970 mainframe, Jimmy Carter. They were right. You grow too large, and then you run out of oil and the Arctic melts" (McKibben 2010: 97).

It is highly unlikely that the United States, let alone the world, will arrive at a *voluntary consensus* to reduce growth. Since the 1940s economic growth has been the vehicle by which the country has achieved laudable social goals such as the increase in living standards and the reduction of poverty. Growth as the primary vehicle of well being is even codified into law by the Employment Act of 1946. The law called for "reasonably" full employment, stable prices, and economic growth. Growth was to become the tool that enabled the others. Growth was the domain of the new economists of the 1960s, based largely on the liberal growth agenda of deficit spending, easy money, rising wages, and augmented consumption.

This strategy came to its demise in the climate of stagflation that followed the 1973 oil price spike, declines in productivity growth, and rising deficits. Since that time it has been followed by a conservative growth agenda of wage cuts, deregulation of production and finance, and easy money. If one wants to conduct a policy of consumption-driven growth in an era of declining wages another source of spending must be available. In the 1980s and beyond this role has fallen to expanding debt and credit. Unfortunately this strategy too has limits. The investment process is still uncertain. The percent of total profits from manufacturing has been in decline since 1970, while profits from financial endeavors have

risen since 1985 with a slight downturn following the aftermath of the short-lived financial collapse of 1987 (Foster and Magdoff: 55) Decade by decade the growth rate of the economy has also declined.

Decade	Annual Average % Growth
1930s	1.3
1940s	5.9
1950s	4.1
1960s	4.4
1970s	3.3
1980s	3.1
1990s	3.1
2000s	2.6

The nation has never seen growth rates in the postwar era as those that were found during the stimulus package known as World War II. Despite Herculean efforts on the part of the FED and the stimulus of subsequent wars and financial deregulation the economy has been growing more slowly for decades. The economy suffers from internal limits to growth as well as external biophysical limits. We are growing too slowly to absorb the unemployed into the economy while growing too rapidly to live within nature's limits. Moreover, even slower growth from a larger base produces a larger environmental impact while, as Domar told us, final demand must increase at an increasing rate in order for investment to increase. And investment drives prosperity. We may well be approaching the era of peak debt, just as we approach the era of peak oil. Clearly world financial markets are reacting to the expansion of government debt. Witness the recent financial troubles in Europe surrounding Greece and now Portugal. The financial panic which comprised the first stage of the current crisis has given way to a fiscal crisis of the state. State and local governments, unable to expand deficits are cutting back, often substantially, and are unable to provide the fiscal stimulus needed to stabilize housing and consumption.

But what are the ramifications of degrowth? David Trauger estimated that to stay within nature's limits per capital income in the US would need to fall to $18,500 (Trauger 2007). It is clear that under our present institutional structures that it would be difficult to accomplish this. I have trouble imagining a growth dependent financial system voluntarily adapting to a regime of no growth, and deficits will be much harder to manage in the absence of economic growth. But the era of degrowth is upon us, if only in the initial stages. Growth rates have been in decline for four decades now, the internal limits manifest as periodic recessions have not disappeared despite the claims of the new economists and the efficient market theorists, while the biophysical limits are approaching exponentially. At this point I have only hopes and not analyses. Some basic suggestions include:

1. A national dialog on income distribution. In *The Republic* Plato commented that the difference in income between a philosopher king and a slave should be five to one. Today the average corporate executive makes 430 times the salary of their average worker. Have a national dialog

on what the gap between the richest executive and the poorest worker should be. Come to a consensus by democratic means. Then figure out the minimum income that anyone should make. Tax all incomes in excess of the difference at 100%. This would reduce economic growth, but that is the objective.

2. Restore meaningful work. Most college professors are very lucky in that our work is meaningful. Most of those who work in factories and offices can't say the same. Restoring the unity between head and hand and giving workers control over the pace of work will reduce productivity, and productivity is the basis of economic growth. This, once again is the goal. If work is meaningful conspicuous consumption (to use Veblen's term) becomes less important. Let satiable biophysical needs and subsistence goods predominate over positional goods. Get out your bicycle. Ditch your suburban assault vehicle.
3. Practice voluntary simplicity. Live in a small house. Walk to work. Eat lower on the food chain. Engage with your community. Plant a garden. Voluntarily reduce your productivity. Take time to watch the sunset.
4. Believe in the hopes offered by technological change but don't place blind faith in it. Nearly every technological change of the 20th century was accomplished by means of an increase in energy use. But we are living through the end of the age of cheap oil and alternatives such as wind and water comprise only about 1% of our energy mix. A ten-fold increase does not replace the oil that we have. But it is running out. We have to come to grips with the idea that using less must be part of the solution.
5. Have fewer children. If Georgescu-Rogen (1975) is correct we need to have a population small enough to feed by means of organic agriculture. Voluntary control over fertility is vastly superior to the other alternatives: genocide or mass starvation.

If by these and other similar means the annual growth rate were reduced we might all live meaningful fulfilling lives within natures limits, and yet few be absolutely starved.

References

Boulding, K. 1966. *The Economics of the coming Spaceship Earth.* Washington D.C.: Resources for the Future.

England, R. and Bluestone, B. 1971. "Ecology and Class Conflict." *Review of Radical Political Economics.* 3(4): 31-55.

Georgescu-Rogen,N.1971. *The Entropy Law and the Economic Process.* Cambridge, MA.: Harvard University Press.

Georgescue-Rogen, N. 1975. "Energy and Economic Myths."*Southern Economic Journal. 41(3).*

Hall, CAS and Day, J. 2009. "Revising the Limits to Growth After Peak Oil. *The American Scientist.* 97 (May June: 230-237.

Hardesty, J., Clement, N., and Jencks, C. 1971. “The Political Economy of Environmental Destruction.” *Review of Radical Political Economics.* 3(4): 82-102.

Marx, K. 1974. “Critique of the Gotha Programme.” In *The First International and After,* David Fernbach (Ed.) New York: Random House.

McKibben, B. 2010. *Eaarth.* New York: Times Books.

Meadows, D, Meadows, D, Randers, J., and Behrens, W. 1972. *The Limits to Growth.* New York: Universe Press.

Meadows, D., Meadows, D. and Randers, J. 1992. *Beyond the Limits.* Post Mills, Vermont. Chelsea Green Publishing Company.

Meadows, D., Meadows, D., and Randers, J. 2004. *The Limits to Growth Thirty-Year Update.* White River Junction, Vermont: Chelsea Green Publishing Company.

Robinson, J. 1972. “The Second Crisis of Economic Theory.” *American Economic Review,* 62(2): 1-10.

Schumacher,E.F. 1973. *Small Is Beautiful.* London: Blonde and Briggs, Ltd.

Simmons, M. 2005. *Twilight in the Desert.* Hoboken, NJ: John Wiley and Sons, Inc.

Solow, R. 1974. The Economics of Resources or the Resources of Economics. *American Economic Review.* 64(2): 1-14.

Trauger, D. 2007. “Ecological Footprints: Tracking Economic Steps Towards Sustainability and Biodiversity Conservation.” *Annual Meetings of the Wildlife Society.* Tucson, AZ.

Yergin, D. 1991. *The Prize.* New York: Simon and Schuster.

Job and Worker Flows: Evidence from New York State Counties

Arindam Mandal*

Abstract

The purpose of the paper is to understand the behavior of the labor market variables, especially job and worker flows in the New York State counties. The paper identifies some strong regional trends in the data. Geography plays an important role in the behavior of these variables. Especially, counties surrounding big cities have generally not so favorable labor market outcomes. This finding can have very strong implication in terms of how the policy makers implement the labor market policies. Labor market policies taken at the national or the state level may not be that efficient or effective in delivering these policies. Policy makers need to take a more localized view of the labor market problems.

INTRODUCTION

Do labor market policies need to be more local? This question plagued the policy makers for longtime. Views varied widely across the spectrum of academic and policy making. With the advent of current recession, this issue again becomes a major contention in the delivery of policy goals.

Aggregate labor market data at the national or state level can give a broad overview of the state of labor market at a particular point in time. Breaking the data at the county or the regional level may give a very different understanding of an event. In standard economic theory, general assumption is perfect labor mobility. In reality, economic and social barriers may separate local labor markets. If these barriers restrict the mobility of labor, then weak labor market demand in particular geographical area or region may increase unemployment rate there above its level in the areas with strong labor demand

Most studies used unemployment as the primary variable to understand labor market behavior. This is understandable given the fact that this is the key variable in understanding the performance of a region and as a measure of the wellbeing of the labor force. At the same time, for the United States and for most of the countries, this is the only variable which available at more localized level. But modern economies are characterized by extremely dynamic labor markets. Each month thousands of workers get separated from their current employment, either to move into unemployment or take up a new job. At the same time, each month firms destroy thousands of existing jobs and create new jobs. For example, in the New York State, during the period between 2001 and 2009, in a typical month on average 0.49 million people were unemployed and actively looking for jobs. In the same time period, about 0.18 million vacancies were posted every month and about 0.4 million workers hired per month. To complement these labor market flows, average 0.53 million workers have been separated every quarter. All these flows are significant, when compare with the average 9.4 million labor force in the New York State. Though, the unemployment

* Department of Economics, School of Business, Siena College, 515 Loudon Road, Loudonville, NY-12211

behavior has caught the interest of the economists for a long time, understanding the important labor market dynamics related to hiring, separations and vacancies are fairly new in the profession. These job and worker flows entail a very interesting dynamics about the labor market reallocation.

In this study, I have used Local Employment Dynamics data to study the job and worker flows across the counties in New York State (NYS). The study finds that job and worker flows along with wages and unemployment have a strong regional character in NYS. The counties with a big city of population 90,000 or more typically have substantially higher vacancy and lower unemployment and separation rate than the counties surrounding the big city counties. Also, average and new hire wage much higher in the counties with big cities.

The paper is divided as follows: section 2 give a short literature review. Section 3 and 4 describe the data and the descriptive statistics. Section 5 gives the spatial patterns in data and finally section 6 and 7 analyze policy implications and conclusions.

LITERATURE REVIEW

Empirical literature on local labor market dynamics primarily focused on understanding unemployment behavior. The magnitude of unemployment disparities among regions within countries is almost as large as the magnitude of unemployment disparities among countries themselves (see Taylor and Bradley, 1997; European Commission, 1999). Interestingly, macroeconomics gives no explanation for the existence of this regional disparity in unemployment. Many macroeconomic studies have attempted to explain unemployment disparities between countries (OECD, 1994; Phelps, 1994; Scarpetta, 1996). From these studies it emerged that the major explanation of unemployment disparities between countries is found in differences in labor market institutions, such as the wage bargaining, social security, retirement and tax systems. However in many countries labor market institutions do not differ to any great extent between regions, therefore other explanatory variables of regional unemployment disparities must be found (Elhorst, 2005). Clearly broad economic and social factors can explain the persistence of unemployment in geographical areas. For example, studies found that high unemployment tends to be in areas with attractive climates and amenities, high wages and high unemployment insurance (Marston, 1985). To complement the unemployment, there is a literature on labor flows. Recently, understanding labor flows caught the interest of the economists more (Pissarides, 2000). The empirical literature on jobs and worker flows primarily focused on the flows at the national level. Lack of data is the primary reason for lack of empirical studies at the regional level. Lack of studies does not mean that understanding regional labor flows is not important.

DATA DESCRIPTION

Quarterly Workforce Indicators (QWI) provides detailed local estimates of variety of employment and earnings indicators based upon information from Local Employment Dynamics (LED). Employment, earnings, gross job creation and destruction, and worker turnover is available at different levels of

geography, typically down to the county or metro area. Detailed information regarding the data is available in Stevens (2007). The definition of the measures in QWI is as follows: -

- **Total Employment:** Total number of workers who were employed by the same employer in both the current and previous quarter.
- **Job Creation:** The number of new jobs that are created by either new area businesses or the expansion of employment by existing firms.
- **New Hires:** Total number of accessions that were also not employed by that employer during the previous four quarters.
- **Separations:** Total number of workers who were employed by a business in the current quarter, but not in the subsequent quarter.
- **Average Monthly Earnings:** Total quarterly earnings of all full-quarter employees divided by the number of full-quarter employees, divided by 3.
- **Average New Hire Earnings:** Total quarterly earnings of all full-quarter new hires divided by the number of full-quarter new hires, divided by 3.

The county level data on employment, unemployment and labor force is collected from the Local Area Unemployment Statistics (LAUS) as provided by the Bureau of Labor Statistics (BLS).

DESCRIPTIVE STATISTICS

We have a panel of sixty two counties of New York State for the time period 2001 Q1 to 2009 Q2. The primary variables of interest are job creation rate (Vacancy rate now onwards), separation rate, new hire rate, unemployment rate, average monthly earnings and average new hire earnings. The data on average new hire earnings and average monthly earnings obtained from LED, whereas separation rate, new hire rate, job creation rate (vacancy rate) and unemployment rate are calculated from LED and LAUS. Separation rate and new hire rate is calculated as the proportion of separations and new hires to total employment as defined in LED. Vacancy rate is obtained by dividing job creations by sum of employed and job creation. The panel allows us to analyze variations both across time and across cross sections. To do cross section analysis, we took the average for each county over the time period 2001 Q1 to 2009 Q2.

Table 1 provides the summary statistics for the cross section data. Vacancy rate varied substantially across counties with 12.79% in New York County to 2.63% in Richmond County, with standard deviation of 1.5. New Hire Rate varied between 22.73% in Hamilton County to 11.49% in Kings County, with standard deviation of 1.8. Separation varied between 36.15% in Hamilton County to 15.55% in Kings County with standard deviation of 2.9. Wage varied more across counties. For example average new hire wage varied from $4,509 to $1,331 in New York County and Yates County respectively, with standard deviation of 454.01. Similarly, average monthly earnings varied between $7,575 to $2,140 in New York County and Hamilton County with standard deviation of 761.33.

Average New Hire Earnings varied between $4,509 and $1,331 in New York County and Yates County respectively. Similarly, average monthly earnings varied between $7,575 in New York County to $2,140 in the Hamilton County.

Table 1: Descriptive Statistics for Cross Section Data

	Unemployment Rate	Vacancy Rate	Hire Rate	Separation Rate	Average New Hire Earnings	Average Monthly Earnings
New York State Average	5.39%	4.96%	14.98%	19.88%	$1,821	$3,087
Max	8.30%	12.79%	22.73%	36.15%	$4,509	$7,575
Min	3.88%	2.63%	11.49%	15.55%	$1,331	$2,140
Standard Deviation	0.83	1.5	1.8	2.9	454.01	761.33
Correlations with…						
Unemployment Rate	1					
Vacancy Rate	-0.1521	1				
Hire Rate	-0.0944	0.3522	1			
Separation Rate	-0.0145	0.3354	0.9217	1		
Average New Hire Earnings	-0.1541	0.4661	-0.0534	-0.1705	1	
Average Monthly Earnings	-0.1689	0.4609	-0.1589	-0.2693	0.9746	1

Source: Author's calculation based upon LED

There is strong positive correlation between average new hire earnings and average monthly earnings. This is because counties with higher average earnings also typically have higher average new hire earnings and vice versa. The correlations between unemployment rate and earnings are negative. This alludes to the fact that tight labor market often leads to lower earnings. The strong positive correlation among hire rate, separation rate and vacancy rate confirms the fact that there is substantial labor market churning.

SPATIAL PATTERNS IN LABOR FLOWS

The descriptive statistics above shows lot of variation in all the labor market variables across counties and within counties. In this part, we will analyze if there are any regional or geographical patterns in these variations. We will primarily focus on the effect of big cities on the movement of labor market variables.

Historically, cities being the hub of economic activities, it is expected that labor market behavior will be substantially different from areas surrounding them. For the analysis purpose, we considered the cities with population of 90,000 or more. According to the latest census data, Albany, Rochester, Syracuse, Buffalo and New York City are the cities in NYS with population of more than 90,000. To understand the regional effects, we analyzed the counties adjoining the county in which the selected city is situated. For example, the city of Rochester is situated in Monroe County; hence it is compared with the adjoining Livingston, Genesee, Ontario, Orleans and Wayne County.

Typically, unemployment rate and the separation rates are much lower in counties with a big city than its level in adjoining counties. On the other hand vacancy rate, average new hires wage and average monthly wage is much higher. Fig. 1-5 and Table 2-6 show the behavior of unemployment and vacancy rate in the counties surrounding Albany, Rochester, Syracuse, Buffalo and New York City. In all the cases, unemployment rate is lower than average in the big city counties. This is expected given the fact that often big cities are the hub economic activities. Vacancy rate is higher than the adjoining counties for all the cities except for Rochester, where Ontario and Genesee counties have higher vacancy rate. NYS do not have any big city or metropolitan area in the northern part of the state. These upstate counties are Clinton, Essex, Franklin, Jefferson and St. Lawrence. Fig 6 and Table 8 show the behavior of unemployment and vacancy rate in the upstate NYS counties. Both vacancies and unemployment rates along with average new hire earnings and average monthly earnings are very similar across these counties. This confirms the fact that big cities have impact on the behavior of labor market variables.

Fig 7-12 and Table 2-6 show the behavior of hiring rate and separation rate. Separation rates are typically lower in the counties with a big city, on the other hand hiring rate do not show any persistent trends based upon geography. Fig 13-18 shows the behavior of average new hire wage and the average monthly wage. Typically, big city counties have substantially higher new hire wage and average monthly wage. Only exception is the Albany County, where average monthly earnings are highest in the Schenectady County. This can be explained by the existence of GE in Schenectady.

POLICY IMPLICATIONS

Above analysis of data clearly points to the fact that geography plays an important role in the labor market. Tackling the unemployment or any other labor market issues needs to have a strong local component. Often policies taken at the national or state level may not be appropriate or efficient in tackling the labor market issues. Recently policy makers with the help of LED data started understanding this issue. For example, LED has been used for the targeted response to economic shocks (Saleh, 2009). In the state of New Jersey, information from LED has been used to provide relief to dislocated workers in the wake of financial crisis. LED data can also be used for economic assessment and industry targeting. The gender and age specific information can be used for better targeting of the groups in need.

CONCLUSION

Understanding the labor market variables at the local level is the key in understanding the problems related to the labor market. Appreciating the fact that the geography plays an important role in understanding the behavior of labor market variables may help the policy makers to address the labor market issues more effectively. The existence of spatial disparities in the market confirms the markets are not efficient. If markets were efficient, then the equilibrating forces of capital and labor mobility and change in relative prices would eventually eliminate unemployment above frictional levels. With the advent of data sets like LED will help the policy make's to understand the labor market outcomes much better and take the market inefficiencies under consideration before taking any economic policy.

.

REFERENCES

Bean, C. R. 1994. "European Unemployment: A Survey." *Journal of Economic Literature*, 32: 573–619.

European Commission. 1999. *Sixth Periodic Report on the Regions*, Brussels/Luxembourg.

Elhorst, J.P. 2005. "The Mystery of Regional Unemployment Differentials: Theoretical and Empirical Explanations." *Journal of Economic Surveys*, 17(5): 709-748.

Marston, S. T. 1985. "Two Views of the Geographic Distribution of Unemployment." *The Quarterly Journal of Economics*, 100: 57–79.

OECD. 1994. *The OECD Jobs Study*. Paris: OECD.

Phelps, E. S. 1994. *Structural Slumps*, Cambridge, MA: Harvard University Press.

Pissarides, C.A. 2000. *Equilibrium Unemployment Theory*, Second Edition, Basil Blackwell, Oxford.

Scarpetta, S. 1996. "Assessing the Role of Labour Market Policies and Institutional Settings on Unemployment: A Cross-Country Study." *OECD Economic Studies*, 26: 43–98.

Saleh, Y. 2009. *Targeted Response to Economic Shocks: The Case of New Jersey*, Department of Labor and Workforce Development, New Jersey, http://lehd.did.census.gov/led/library/workshops/2009/LED_03_05.pdf.

Stevens, D.W. 2007. *Employment that is not covered by state unemployment insurance laws.* United States Census Bureau Technical Paper TP-2007-04.

Taylor, J. and Bradley, S. 1997. "Unemployment in Europe: A Comparative Analysis of Regional Disparities in Germany, Italy and the UK." *Kyklos*, 50: 221–245.

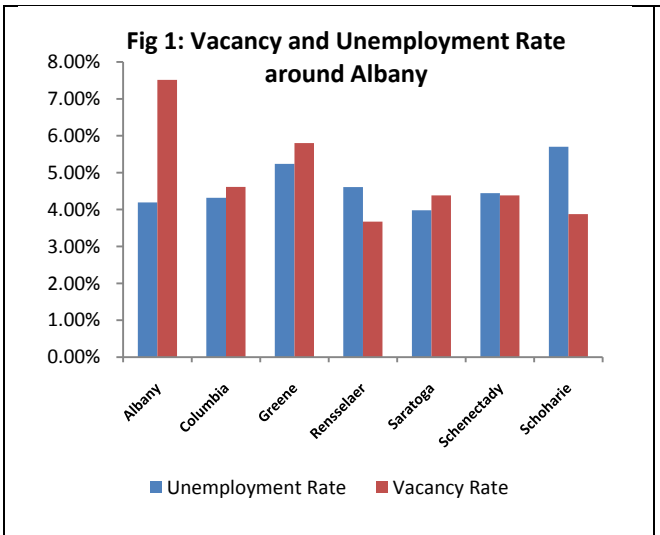


Fig 1: Vacancy and Unemployment Rate around Albany
8.00%
7.00%
6.00%
5.00%
4.00%
3.00%
2.00%
1.00%
0.00%
Albany
Columbia
Greene
Rensselaer
Saratoga
Schenectady
Schoharie
Unemployment Rate
Vacancy Rate

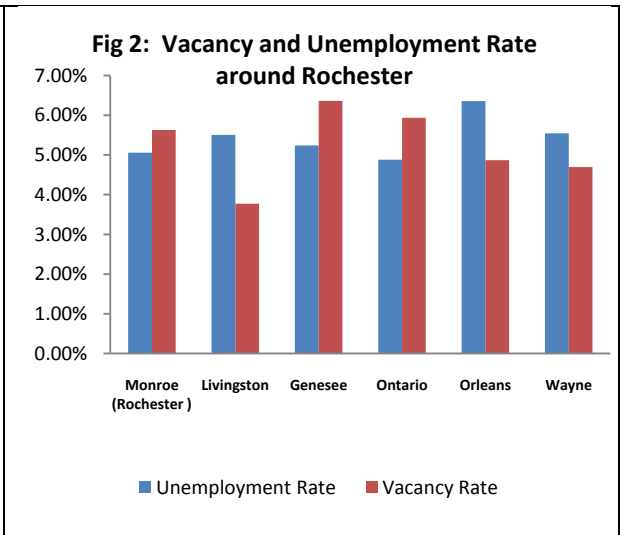


Fig 2: Vacancy and Unemployment Rate around Rochester
7.00%
6.00%
5.00%
4.00%
3.00%
2.00%
1.00%
0.00%
Monroe (Rochester)
Livingston
Genesee
Ontario
Orleans
Wayne
Unemployment Rate
Vacancy Rate

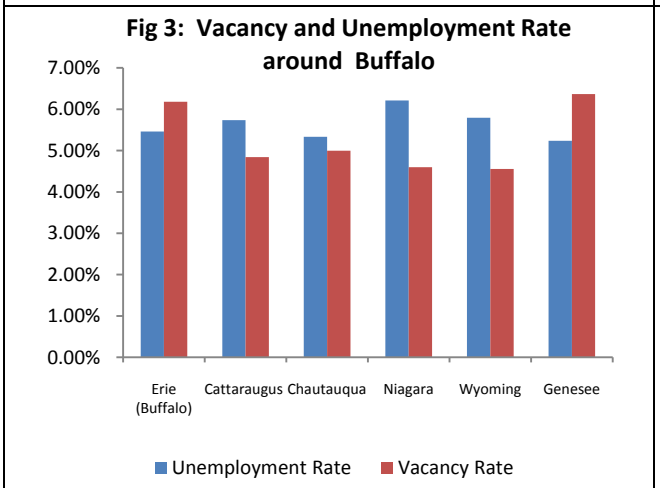


Fig 3: Vacancy and Unemployment Rate around Buffalo
7.00%
6.00%
5.00%
4.00%
3.00%
2.00%
1.00%
0.00%
Erie (Buffalo)
Cattaraugus
Chautauqua
Niagara
Wyoming
Genesee
Unemployment Rate
Vacancy Rate

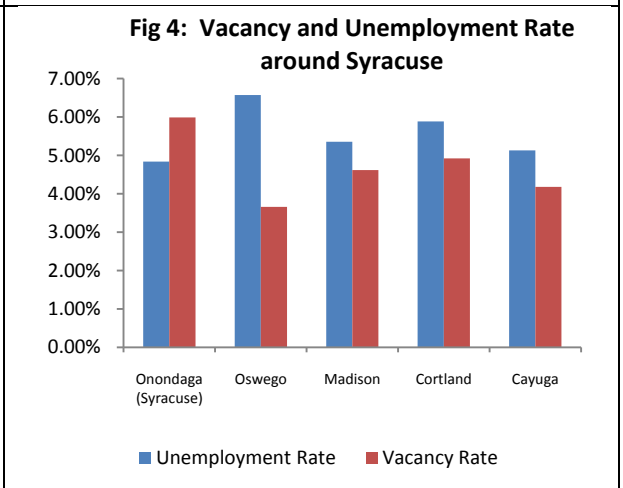


Fig 4: Vacancy and Unemployment Rate around Syracuse
7.00%
6.00%
5.00%
4.00%
3.00%
2.00%
1.00%
0.00%
Onondaga (Syracuse)
Oswego
Madison
Cortland
Cayuga
Unemployment Rate
Vacancy Rate

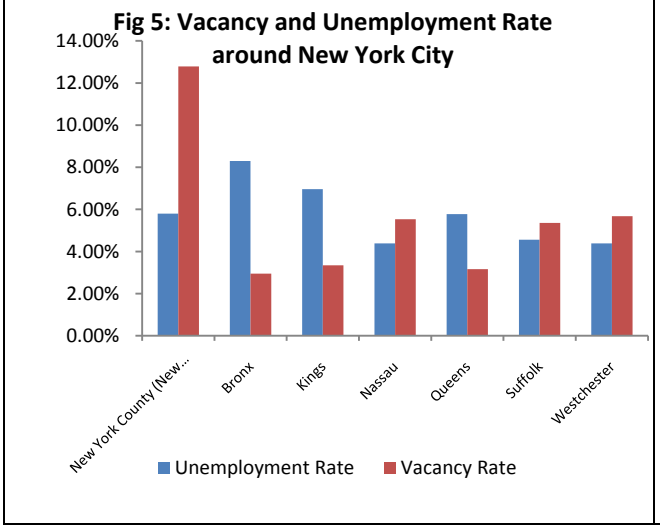


Fig 5: Vacancy and Unemployment Rate around New York City
14.00%
12.00%
10.00%
8.00%
6.00%
4.00%
2.00%
0.00%
New York County (New...
Bronx
Kings
Nassau
Queens
Suffolk
Westchester
Unemployment Rate
Vacancy Rate

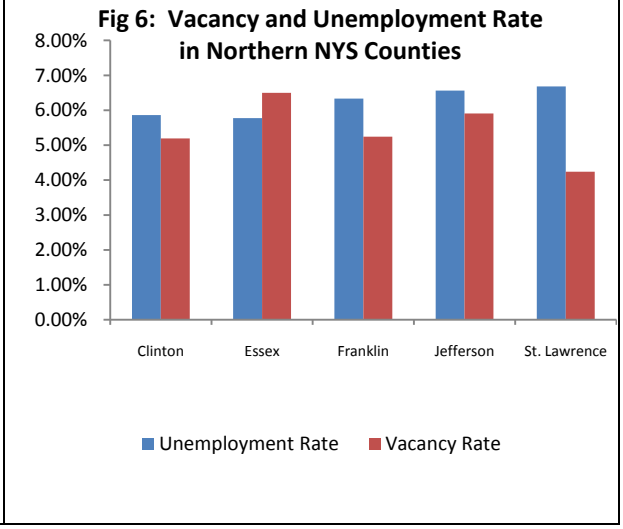


Fig 6: Vacancy and Unemployment Rate in Northern NYS Counties
8.00%
7.00%
6.00%
5.00%
4.00%
3.00%
2.00%
1.00%
0.00%
Clinton
Essex
Franklin
Jefferson
St. Lawrence
Unemployment Rate
Vacancy Rate

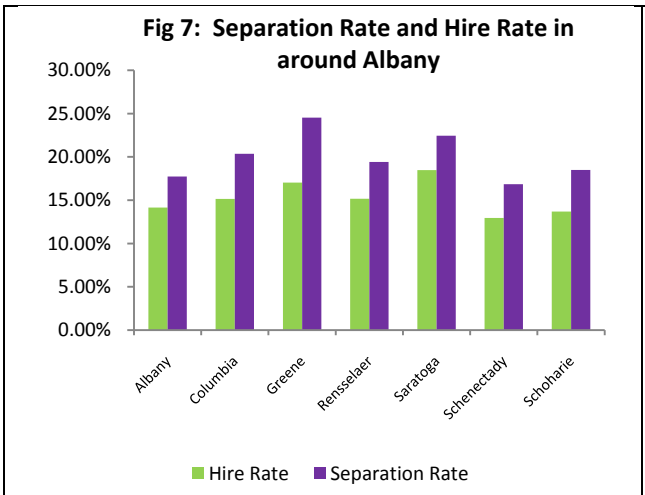


Fig 7: Separation Rate and Hire Rate in around Albany
30.00%
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
Albany
Columbia
Greene
Rensselaer
Saratoga
Schenectady
Schoharie
Hire Rate
Separation Rate

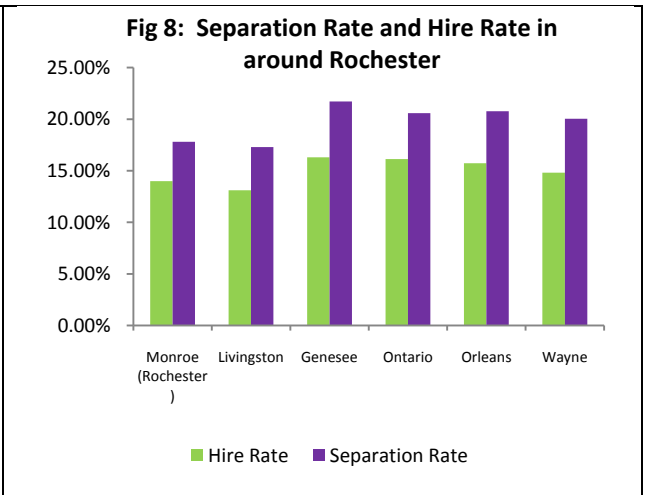


Fig 8: Separation Rate and Hire Rate in around Rochester
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
Monroe (Rochester)
Livingston
Genesee
Ontario
Orleans
Wayne
Hire Rate
Separation Rate

Fig 9: Separation Rate and Hire Rate in around Buffalo
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
Erie (Buffalo)
Cattaraugus
Chautauqua
Niagara
Wyoming
Genesee
Hire Rate
Separation Rate

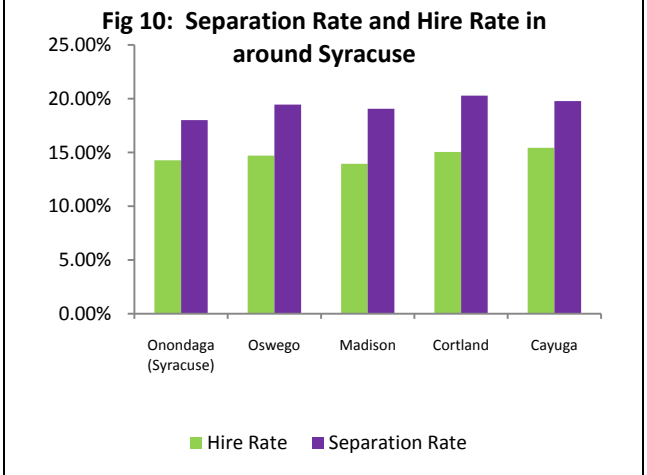


Fig 10: Separation Rate and Hire Rate in around Syracuse
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
Onondaga (Syracuse)
Oswego
Madison
Cortland
Cayuga
Hire Rate
Separation Rate

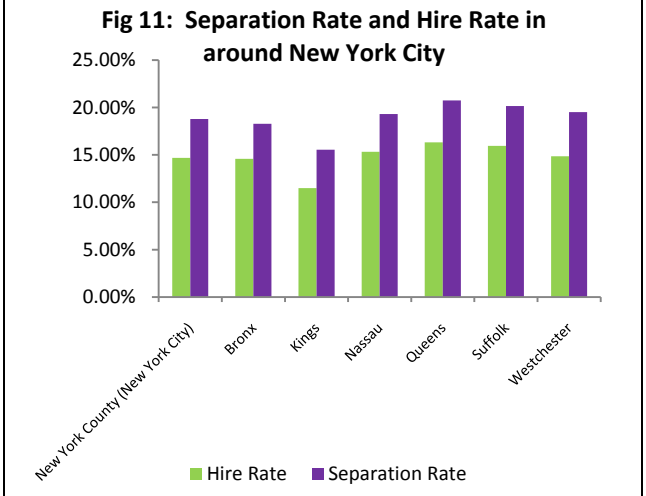


Fig 11: Separation Rate and Hire Rate in around New York City
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
New York County (New York City)
Bronx
Kings
Nassau
Queens
Suffolk
Westchester
Hire Rate
Separation Rate

Fig 12: Separation Rate and Hire Rate in Northern NYS Counties
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
Clinton
Essex
Franklin
Jefferson
St. Lawrence
Hire Rate
Separation Rate

Fig 13: Average New Hire and Average Monthly Earnings around Albany

Fig 14: Average New Hire and Average Monthly Earnings around Rochester

Fig 15: Average New Hire and Average Monthly Earnings around Buffalo

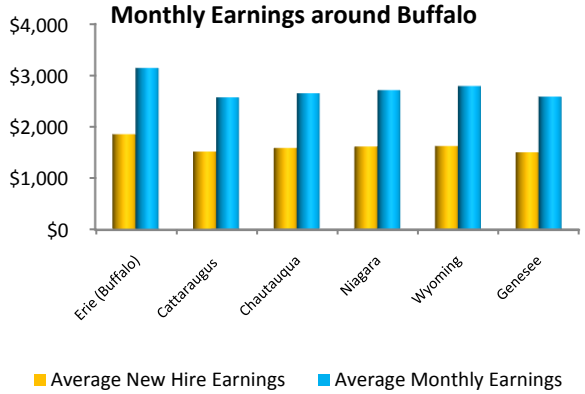


Fig 16: Average New Hire and Average Monthly Earnings around Syracuse

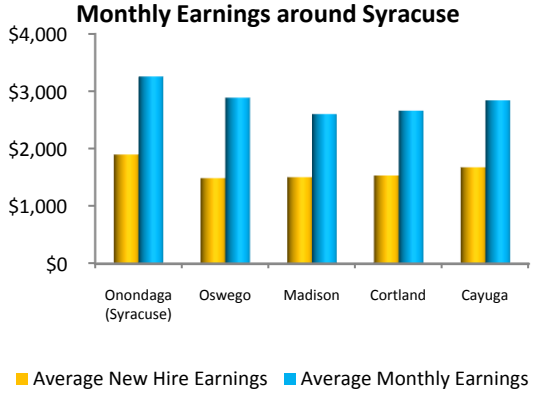


Fig 17: Average New Hire and Average Monthly Earnings around New York City

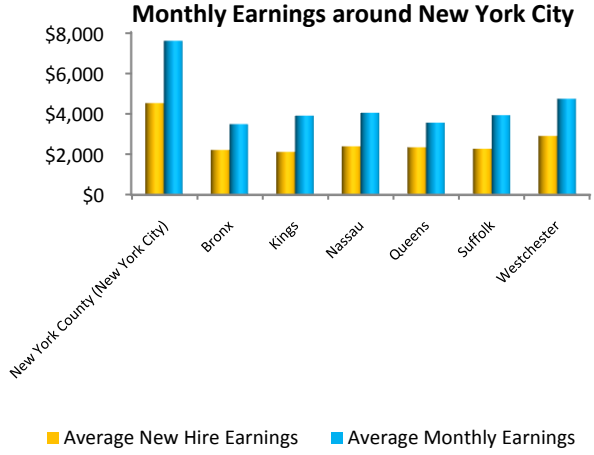


Fig 18: Average New Hire and Average Monthly Earnings in Northern NYS Counties

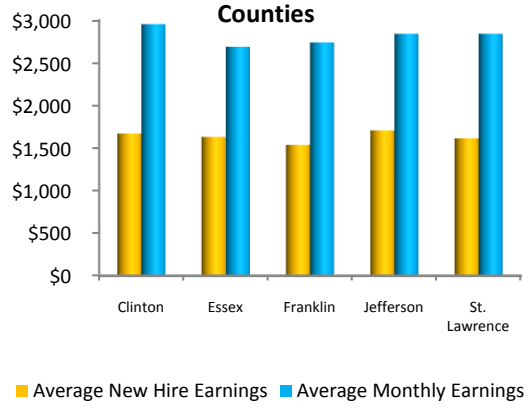


Table 2: Summary Statistics of Counties adjoining Albany

County	Average New Hire Earnings	Average Monthly Earnings	Unemployment Rate	Vacancy Rate	Hire Rate	Separation Rate
Albany	$2,118	$3,676	4.19%	7.51%	14.15%	17.73%
Columbia	$1,669	$2,737	4.32%	4.61%	15.15%	20.35%
Greene	$1,683	$2,889	5.24%	5.80%	17.03%	24.52%
Rensselaer	$1,852	$3,030	4.61%	3.67%	15.16%	19.42%
Saratoga	$1,785	$3,035	3.98%	4.38%	18.47%	22.44%
Schenectady	$2,159	$3,803	4.44%	4.38%	12.95%	16.84%
Schoharie	$1,549	$2,733	5.70%	3.88%	13.69%	18.50%
Average	$1,831	$3,129	4.64%	4.89%	15.23%	19.97%
Std. Dev	231.12	436.04	0.61	1.34	1.93	2.72
Max – Min	$610	$1,071	1.72%	3.84%	5.52%	7.68%

Table 3: Summary Statistics of Counties adjoining Rochester

County	Average New Hire Earnings	Average Monthly Earnings	Unemployment Rate	Vacancy Rate	Hire Rate	Separation Rate
Monroe (Rochester)	$1,908	$3,334	5.06%	5.63%	13.99%	17.80%
Livingston	$1,533	$2,758	5.50%	3.77%	13.10%	17.30%
Genesee	$1,507	$2,590	5.24%	6.37%	16.30%	21.72%
Ontario	$1,709	$2,908	4.88%	5.94%	16.14%	20.58%
Orleans	$1,563	$2,680	6.35%	4.87%	15.74%	20.77%
Wayne	$1,755	$2,837	5.54%	4.69%	14.82%	20.04%
Average	$1,662	$2,851	5.4%	5.2%	15.0%	19.7%
Std. Dev	156.14	261.74	0.52	0.95	1.28	1.76
Max – Min	$401	$744	1.47%	2.59%	3.19%	4.43%

Table 4: Summary Statistics of Counties adjoining Buffalo

County	Average New Hire Earnings	Average Monthly Earnings	Unemployment Rate	Vacancy Rate	Hire Rate	Separation Rate
Erie (Buffalo)	$1,861	$3,147	5.46%	6.18%	15.32%	19.26%
Cattaraugus	$1,521	$2,574	5.73%	4.84%	13.68%	18.18%
Chautauqua	$1,592	$2,657	5.33%	5.00%	14.73%	19.35%
Niagara	$1,622	$2,719	6.21%	4.60%	15.03%	19.99%
Wyoming	$1,631	$2,797	5.80%	4.55%	12.44%	16.48%
Genesee	$1,507	$2,590	5.24%	6.37%	16.30%	21.72%
Average	$1,622	$2,747	5.63%	5.26%	14.6%	19.16%
Std. Dev	127.57	212.72	0.36	0.81	1.35	1.76
Max –Min	$354	$574	0.97%	1.81%	3.86%	5.24%

Table 5: Summary Statistics of Counties adjoining Syracuse

County	Average New Hire Earnings	Average Monthly Earnings	Unemployment Rate	Vacancy Rate	Hire Rate	Separation Rate
Onondaga (Syracuse)	$1,903	$3,264	4.84%	5.99%	14.27%	18.01%
Oswego	$1,491	$2,895	6.57%	3.66%	14.70%	19.46%
Madison	$1,510	$2,610	5.35%	4.62%	13.94%	19.06%
Cortland	$1,539	$2,665	5.88%	4.92%	15.05%	20.29%
Cayuga	$1,680	$2,848	5.13%	4.18%	15.44%	19.78%
Average	$1,625	$2,856	5.55%	4.67%	14.68%	19.32%
Std. Dev	172.27	257.54	0.69	0.88	0.60	0.86
Max – Min	$412	$654	1.74%	2.33%	1.50%	2.28%

Table 6: Summary Statistics of Counties adjoining Rochester

County	Average New Hire Earnings	Average Monthly Earnings	Unemployment Rate	Vacancy Rate	Hire Rate	Separation Rate
New York County	$4,509	$7,575	5.80%	12.79%	14.68%	18.79%
Bronx	$2,202	$3,478	8.30%	2.95%	14.58%	18.29%
Kings	$2,115	$3,887	6.96%	3.35%	11.49%	15.55%
Nassau	$2,380	$4,039	4.39%	5.53%	15.33%	19.32%
Queens	$2,338	$3,541	5.77%	3.16%	16.32%	20.75%
Suffolk	$2,269	$3,916	4.56%	5.36%	15.96%	20.16%
Westchester	$2,899	$4,726	4.39%	5.68%	14.86%	19.52%
Average	$2,673	$4,452	5.74%	5.54%	14.75%	18.91%
Std. Dev	848.20	1436.67	1.48	3.41	1.58	1.69
Max – Min	$2,394	$4,097	3.91%	9.84%	4.84%	5.20%

Table 7: Summary Statistics of Counties in Northern NYS

County	Average New Hire Earnings	Average Monthly Earnings	Unemployment Rate	Vacancy Rate	Hire Rate	Separation Rate
Clinton	$1,669	$2,953	5.86%	5.19%	15.63%	19.88%
Essex	$1,630	$2,687	5.78%	6.50%	16.97%	23.56%
Franklin	$1,535	$2,738	6.33%	5.24%	12.22%	17.49%
Jefferson	$1,705	$2,841	6.56%	5.91%	15.90%	19.97%
St. Lawrence	$1,612	$2,840	6.68%	4.24%	13.97%	18.98%
Average	$1,630	$2,812	6.24%	5.42%	14.94%	19.98%
Std. Dev	64.27	103.00	0.41	0.85	1.86	2.24
Max - Min	$170	$266	0.91%	2.26%	4.75%	6.07%

Source: Author's calculation based upon LED

THE POLITICAL ECONOMY OF EDUCATION REFORM: THE CASE OF *A NATION AT RISK*

Johanna Mitchell*

ABSTRACT

This paper explores the power structure behind education policy through analysis of the political and economic affiliations of the authors of A Nation at Risk, an educational reform report. Analysis of historical documents such as the 1983 reform report provides an explanation for contemporary policy and practice. Many researchers have analyzed the contents of the report and the historical circumstances following its release, but the connections authors shared with the state and their class backgrounds have been ignored as has the class-biased nature of the education policy-making process. Power structure research and a Marxist theoretical perspective are employed in this essay to demonstrate the interlocking network of authors and their shared organizational affiliations and interests. The networks and exploration of the policy-making process demonstrate that education policy resulting from reports such as A Nation at Risk is formulated by a power elite, and serves to enhance the interests of the dominant economic class.

INTRODUCTION

The *No Child Left Behind Act* (NCLB) of 2002 still causes teacher educators serious concerns. It is a rare day that passes on campus when no prospective or practicing teacher expresses anxiety about NCLB requirements. Teacher educators are sometimes frantic to develop new programs or alter existing ones to address teacher-licensure and test requirements without sacrificing goals of equity and access. To some, NCLB is an unprecedented and unexpected threat upon the educational shores, an Act guided by specific economic objectives. Yet, it would be a mistake to see the provisions of the Act as unique in the history of American education. Indeed, little more than 25 years ago, another attack on public schooling surfaced that fostered many of the same recommendations formulated to ostensibly address the national economic situation. Unlike NCLB, *A Nation at Risk* was issued by the executive branch instead of Congress. It did not become law through actions by the federal government. *A Nation at Risk*, however, was built upon many of the same assumptions about the national economic situation and had similar effects on the education community. Indeed, NCLB is one of the indirect outcomes of that earlier effort to reform American public schooling. *A Nation at Risk* laid the foundations for debate and posited the national educational and economic values that eventually came to legislative fruition with NCLB. Thus, it is important to look closely at the origins of the earlier report and examine it from a class-power perspective since it was written by individuals associated in a variety of ways with a dominant economic class.

* Hartwick College, 1 Hartwick Drive, Oneonta, NY 13820

The essence of the education literature regarding the report has been content analysis seeking answers to questions about "what" was written and speculations at to "why." This study addresses the failure of education researchers to deal with the agents and mechanisms employed in the construction of *A Nation at Risk* which contributes a substantially different answer to the question of "why" it was written. The theoretical perspective employed in the study emphasizes questions of class power and examines the concrete relationships between class forces and policy formation. As an empirical and historical study, this essay provides additional information about policy formation heretofore lacking in the education literature. As a theoretical treatise, it posits explanations for the class-specific recommendations made in *A Nation at Risk*.

THEORETICAL AND EMPIRICAL QUESTIONS

A Nation at Risk assumed a direct relationship between education and the condition of the economy.[1] It linked a supposedly ailing economy with decreasing student achievement and lowered educational standards. The flamboyant phraseology and educational prescriptions of *A Nation at Risk* are commonplace in discussions of that education reform initiative, and may even have become cliché. The contents of the report are familiar, as are the many debates surrounding its political and economic goals, but its subtle reactionary and class-specific prejudices may be less well-known. Recent efforts to assess the impact of *A Nation at Risk* have been disappointing because they unequivocally accept the notion that academic achievement was at the heart of the report rather than economic concerns. These assessments continue to neglect the political and economic roots of the report by failing to address questions about the interests and affiliations of the authors or the nature of the policy-making process through which the report was developed.[2]

The empirical questions that should be asked in order to get to the biased root of this reform effort are: Who were the members of the National Commission on Excellence in Education (NCEE), the authors of *A Nation at Risk*? What were their political and class backgrounds? What were their shared connections to the state, corporations, universities, policy-making bodies, and the social elite? Theoretical questions include: To what extent did the construction of the 1983 educational reform report reflect interests ascribable to a dominant class? How did the policy-making process itself contribute structural limitations and biases to the final results? Much of the literature dealing with *A Nation at Risk* has neglected examination of the political and economic backgrounds of the authors, and does not theorize the class-specific education policy-making process that initially set the boundaries for prescriptions for reform.

DEFINITIONS

For purposes of this essay, it is necessary to operationally define a set of terms within a Marxist perspective. The essay deals with "essentially contested concepts" (Hanson, 1985) such as *power*, *dominant class*, *the state*, and the *power elite.*[3] First, *power* is defined in structural terms as a set of "relation-ships in society and exercised by individuals and groups based on their location in a given structure" (Isaac, 1987, 28).[4] This definition includes practical norms or customs and traditions, covert official and unofficial connivance, and enduring social relationships that are structurally set. The exercise of power must be seen as contingent. Its limits are set by the way particular individuals and groups choose to deal with their given circumstances.

Second, the *dominant class* is a set of individuals and families (including, but not exclusively, a social elite) who compose an "economic class that has successfully monopolized other power sources to dominate a state-centered society at large" (Mann, 1986, 5). The claim that a particular class is dominant, however, is not meant to imply that it is ultimately successful. G. William Domhoff claims that dominance does not leave other levels of society in a position of total powerlessness. He argues that "[D]omination does not mean total control, but the ability to set the terms under which other groups and classes must operate" (1983, 2).[5] It is crucial to emphasize the power of the dominant class to set the terms for debate, because with respect to policy formation, what is really important about *A Nation at Risk* is not its particular recommendations as much as the fact that it established the philosophical boundaries within which education would operate. This process eventually led to NCLB.

The third term to be defined, *the state*, is a multifaceted concept that refers to "an interconnected network of territorial, institutional, and ideological phenomena" (Barrow, 1993, 24). Under the theoretical perspective employed in this essay, the state can frequently be depended upon to further specific interests; those of the dominant class, which is a capitalist class.

Fourth, the dominant class is the pool from which the *power elite* is most often pulled. The power elite is,

> "…composed of men [and women] whose positions enable them to transcend the ordinary environments of ordinary men and women; they are in positions to make decisions having major consequences .... they are in command of the major hierarchies and organizations of modern society .... not solitary rulers. Advisers and consultants, spokesmen and opinion-makers are often the captains of their higher thought and decision" (Mills, 2000, 3-4).

They should be seen as members of a dominant class or "members of a top social stratum, as a set of groups whose members know one another, see one another socially and at business, and so, in making decisions, take one another into account" (11). However, this is not always the case. Not all members of this dominant class are politically active and are therefore not considered part of the power elite. Individuals from other classes sometimes move in and out of the power elite. They occupy a structural or

institutional position of power for a time, and then move out of it, leaving the structural position intact. The extent of this group's power is “subject to historical variation” (20).

POWER STRUCTURE RESEARCH

The Table below lists the authors of *A Nation at Risk* along with their primary affiliations. The list includes Terrel Howard Bell, Secretary of Education, who was appointed by President Reagan and assigned the duty of establishing the NCEE which produced the report. Bell established the Commission, set the agenda and goals, and determined the time-table. The essay includes several Figures that graphically demonstrate the voluntarily listed organizational and interpersonal affiliations of Bell and the other 18 authors of *A Nation at Risk*. The data contained in the Table and Figures include the names of the individual authors of the report followed by his or her initials in parentheses. The Figures reflect the connections that the authors of *A Nation at Risk* had with the state, corporations, universities, foundations, and the social elite. The data provide empirical evidence of membership in the power elite as defined above. The organizations are loosely categorized as follows: (1) The “STATE” includes the executive, legislative, and judicial branches of the federal government and its agencies, committees, and councils. It also includes the United Nations affiliations and agencies of state governments such as Departments of Education. Of course, there are great power differentials between these different groups, but this category follows the definition of “state” as including the governmental, coercive, administrative, and ideological subsystems. (2) The category titled “CORPORATIONS AND BUSINESS” encompasses Chief Executive Officers of corporations and their Boards of Directors, corporate law firms, banks, and any major corporate-sponsored policy groups. It also includes individuals who own their businesses, even if those enterprises are rather small. (3) “UNIVERSITIES AND THEIR INSTITUTES” incorporates all universities attended by individuals as well as those in which the individual taught or was an administrator, trustee, or visiting lecturer. It also includes universities that awarded honorary degrees, and any research institutes directly affiliated with a specific university. (4) The category titled “FOUNDATIONS AND ASSOCIATIONS” encompasses research institutes that could not be tied to a university, and includes private foundations, think tanks, and societies. Organizations that could not be categorized but were affiliated with more than one individual were listed in this category. (5) The “SOCIAL ELITE” lists private clubs attended by a select group, fraternities, sororities, and private societies.

As noted above, all listings in biographical sources are voluntary so there exists the possibility of numerous false negatives. The guiding thesis for assigning individuals to organizational categories was to err on the side of negative affiliations. It was the case that many affiliations could not be isolated under any of the categories applied, so those affiliations were not added to the networks. It was also the case that some institutions appeared several times for a particular individual. In that instance, the institution was drawn into the networks only once.[6]

Table 1: Authors and Primary Affiliations

1. T. H. Bell (THB) Secretary of Education, United States Department of Education, Washington DC.
2. D. P. Gardner (DPG) Chair of NCEE, President of the University of Utah, President-Elect, University of California
3. Y.W, Larsen (YWL) Vice-Chair of NCEE and Immediate Past President of the San Diego School Board, CA
4. W.O. Baker (WOB) Chairman of Board (Retired) Bell Telephone Labs
5. M.A. Campbell (MAC) Former Commissioner of Education, NE
6. E. A. Crosby (EAC) Principal, Northern High School in Detroit, MI
7. C.A. Foster, Jr. (CAF) Immediate Past-President of Foundation for Teaching Economics, San Francisco, CA
8. N.C. Francis (NCF) President of Xavier University, New Orleans, LA
9. A. B. Giamatti (ABG) President of Yale University in New Haven, CT
10. S.B. Gordon (SBG) President Highline Community College Midway, WA
11. R.V. Haderlein (RVH) Immediate Past-President of the National School Boards Association
12. G.Holton (GH) Mallinckrodt Professor of Physics and of the History of Science, Harvard University
13. A.Y. Kirk (AYK-RAK) Kirk Associates in Mecosta, MI
14. M.S. Marston (MSM) Member of the Virginia State Board of Education
15. A.H. Quie (AHQ) Former Governor of MN
16. F.D. Sanchez, (FDS) Superintendent of Schools Albuquerque, NM
17. G.T. Seaborg (GTS) University Professor of Chemistry and Nobel Laureate, University of California- Berkeley
18. J. Sommer (JS) National Teacher of the Year for 1981-1982, New Rochelle High School, New Rochelle, NY
19. R. Wallace (RW) Principal of Lutheran High School East, Cleveland Heights, OH

INTERLOCKING CONNECTIONS AND POINTS OF DIVERGENCE

In a nation where the government is closely tied to large corporations and a social elite, policy would be class specific. The cumulative effect of upper class power would establish a situation wherein the policy-planning process itself is institutionalized through time so that its systematic conduct would rarely be questioned. In such a context, education policy would be more likely to serve the interests of a dominant class. If this is the case with *A Nation at Risk*, there should be ample evidence that indicates multilayered connections between the individuals who actually construct policy and the state, corporations, universities, foundations and think tanks, and a social upper class. This essay provides evidence that can serve as a starting point for class analysis.

At first glance, the authors of the report appear to be a diverse group of educators and political officials, with one corporate chairman and a parent of school children. Indeed, *The New York Times* reported that Annette Kirk (AYK-RAK) was "a mother of four from Mecosta, Mich. who was the parent representative," [7] leaving readers to infer that even typical or common folk whose children attended public schools were involved. Bell (THB) referred to the reform movement spurred on by *A Nation at Risk* as "grass-roots" (Bell, 1984, 531), and the report itself discussed the different ideas and arguments it heard during several public hearings and in NCEE commissioned papers. It stated that "In many ways, the membership of the Commission itself reflected that diversity and difference of opinion during the course of its work" (NCEE, 1983, 3). One is left with the impression that this diverse group must have all reached

the same conclusion through a democratic process, and therefore that conclusion must be accurate and unbiased.

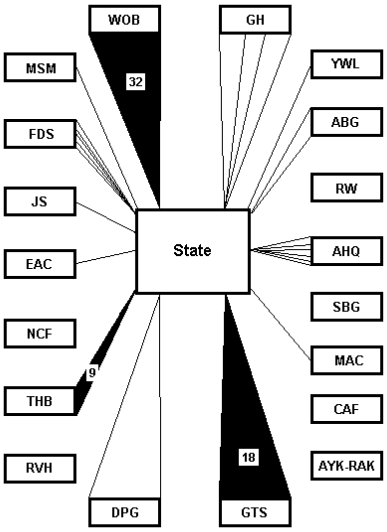
Contrary to what was portrayed, *A Nation at Risk* did not represent a cross-section of opinion or views of people from different positions in American society. Rather, it reflected the point of view of an inter-connected group of policy elites. Figure 1 is a graphic representation of the organization network that illustrates the connections between the authors of *A Nation at Risk* and the state. Baker (WOB), for instance, had 32 connections to the state including membership in military groups such as the Advisory Board of Military Personnel and Supplies, Air Force Systems Command, and Oakridge National Laboratory. He was involved with other state organizations such as the Commission on Critical Choices for Americans, Energy Policy Office, Federal Energy Administration, National Bureau of Standards, National Institute of Health, National Research Council, National Security Agency, several presidential advisory groups, and many more. Quie (AHQ) was a former member of Congress and Governor of the State of Minnesota. Seaborg (GTS) was Chairman of the Atomic Energy Commission for a time, and was a member of the President's Science Advisory Committee, Federal Radiation Council, National Aeronautics and Space Council, Federal Council on Science and Technology, and more. Figure 1 illustrates that *some* of the authors of *A Nation at Risk* were highly and intricately connected to the federal government.

Some of the federal connections are especially significant with regard to policy formation. Membership in these organizations indicates membership in the power elite. Baker's (WOB) connections to the National Security Agency and the Federal Emergency Management Advisory Board are indicative of a highly placed, influential individual who was actively engaged in federal policy construction. Giamatti (ABG) was a member of the Council on Foreign Relations (CFR) which is the largest policy formation body in the United States. Indeed, the CFR was a primary actor in the creation of the post-World War II international economy, and its relationship to the upper class and corporate community has been clearly demonstrated.[8] Some of the authors were connected to international organizations such as the United Nations Education, Science, and Cultural Organization (UNESCO), International Atomic Energy Agency, and the Limited Test Ban Treaty. These relationships that some authors had with federal policy boards, the executive branch of the government, and international organizations support the inference that they were politically powerful. These authors made decisions with far-reaching consequences; their positions allowed them to transcend the environments of ordinary men and women; and, they had access to major hierarchies.

On the other hand, there was a minority of individuals with few to no connections to the state other than their membership on the NCEE and their role as State and local educators. A small number who wrote the reform report were isolates; that is, they were probably not affiliated with the power elite on a continuing basis. No connections between Foster (CAF) and the state or corporations, for instance, could be located in voluntary sources. Although Foster once served as president of the Foundation for Teaching Economics, he is considered an isolate since this affiliation is not shared with any of the authors of the

same report. The Foundation for Teaching Economics, however, is an important policy-discussion group, but to err on the side of caution, Foster is not considered part of the power elite.

Figure 1: Organization Network: *A Nation at Risk*, State



Figures 2 through 5 (available from author) illustrate the affiliations authors had with corporations and business, universities and their institutes, foundations and associations, and the social elite, respectively. It is important to note that several of the individuals who were connected to the state, were also involved in corporations. This is not the case for all of the authors of *A Nation at Risk*. Those individuals who did not have numerous connections to the state, did not have numerous connections to corporations and business. Several authors had multiple connections to universities and their institutes, but few had any connection to the social elite. It is important to note that Baker (WOB) and Seaborg (GTS) had numerous connections with the state, corporations and business, universities and their institutes, foundations and associations, and the social elite. Those connections total 117 for Baker and 97 for Seaborg.

Most important with respect to membership in the power elite are the connections that authors had with corporations and business and the social upper class. Only 5 members of the NCEE were affiliated with corporations. Baker was the most intricately connected with the corporate community. Baker, Seaborg, Holton, and Gardner (WOB, GTS, GH, DPG) had the most connections to the social elite. Membership in Sigma Xi, the Bohemian Club, and Colonial Dames of America is especially indicative of upper class standing, as is attendance at elite private schools. It is important to note that affiliation with an elite university is not enough to qualify for membership in the power elite. Combined with private club membership, corporate affiliations, and relationships with think tanks and policy-making groups, however, attendance at or teaching in an elite university allow an individual access to seats of power that ordinary persons do not have.

Figures 6 through 10 (available from author) illustrate the interpersonal networks or *shared* contacts among the authors and organizations. Shared connections are extremely important because they frequently serve to strengthen the social ties that augment class-consciousness. Figure 11 demonstrates the complexity of these interlocking relationships.

The intricate interlocks demonstrated by the data are augmented by other relationships with the state. Some of the shared connections include the Federal Emergency Management Agency (FEMA), National Aeronautics and Space Administration (NASA), National Institute of Health (NIH), National Science Foundation (NSF) which is directed by the National Science Board, the Office of Science and Technology (OST), the National Research Council (NRC), and the Presidents Science Advisory Committee (PSAC).

It is crucial to point out that there is a significant disparity between the number and type of affiliations claimed by the highly connected authors of the report and other authors. In the Figures, a triangle of connections is visible between one set of three authors and another for all of the others. The gulf that divides the two sets of individuals, those intricately connected from those who have few to no connections, is significant. The intricate connections allow for the inference that few individuals were highly influential and would have had a great deal of power with respect to policy formation.

Figure 10: Interpersonal Network – Social Elite

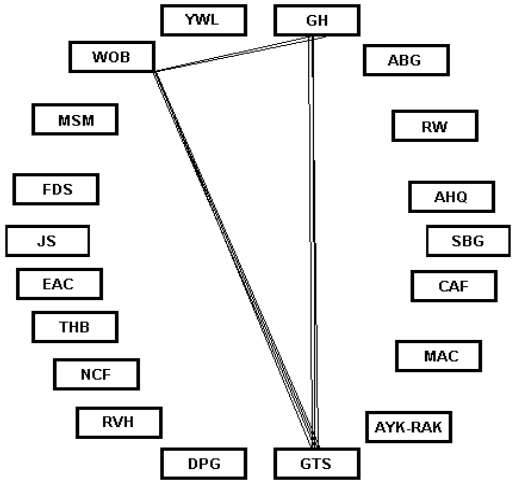
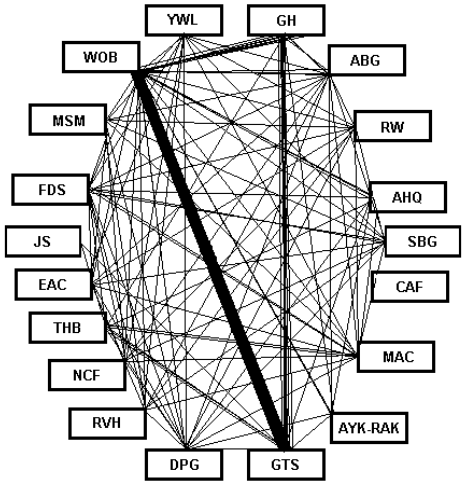


Figure 11: Shared Connections



INTERPRETATION OF POLICY CONSTRUCTION

In a letter to Carl D. Perkins, Bell wrote that he was establishing a "blue-ribbon" commission "comprised of 19 respected citizens from the education community, State and local government, and the general public."[9] The wording in the letter was notably similar to the wording of the final report, employing words such as "mediocrity" and "excellence." It provided Perkins with information about the problem to be studied, the sources the NCEE would use to make its final recommendations, the individuals who would write the report, and the time expected to complete it. These elements of the policy formulation process were decided in advance.

The policy-planning process begins with the identification of specific problems so that the parameters of discussion are established from the outset. This limits the range of possible solutions that are considered. The policy-formation process ends in government where solutions are constructed into policy. Between the beginning and the end “there is a complex network of people and institutions that plays an important role in sharpening the issues and weighing the alternatives” (Domhoff, 2002, 71). It is during this process that interests are ascertained, articulated, and refined. Policy constructors from the dominant class and subordinate groups either clarify and solidify their interests, or those interests are amended, altered, or silenced. Either way, a consensus is reached through compromise and particular interests are transcended during the process. This point becomes most obvious in the records of the public hearings that the NCEE held, and in the selective emphasis of specific recommendations made in the NCEE commissioned papers.

Between 1981 and 1983 the NCEE held six public hearings in different cities across the nation. Each hearing was directed at a particular topic which promoted specific educational reforms, and was chaired by a member of the commission. Each hearing began with testimony from a number of *invited corporate and academic experts*, some of whom had submitted commissioned papers for the NCEE to review. In the Denver meeting, Linda Sorrento, an invited guest speaker, stressed that:

> “Perhaps most importantly, we need to establish some kind of vehicle where we can blend the three worlds of government, education, and private industry. We need to support each other more and empower each other on levels that we really have not” (Sorento, Denver, 77).

Sorrento recommended that the three groups work together to develop curriculum because educators need to know "what the needs of industry are" so they can help enhance the marketable skills of students.

The NCEE believed that poor educational performance was at the root of economic problems. Expressions of concern about values or virtue came to the forefront in the public hearings, and those concerns were tied to the educational decay that was seen to be responsible for economic crises. Instruction in academic subjects was a way of improving character or behavior. Some of the commissioned papers relied upon declining test scores as evidence of decay in values, thereby accepting the definition of the problem advanced by elites. Lauren and Daniel Resnick recommended more emphasis on standards, tracking, and examinations. These authors argued that there was a “lack of stringent course requirements” and “weak content and poor instruction within courses” in many cases (Resnick & Resnick, 1982, 35). Alexander Astin imparted an elitist attitude when he cited declining achievement skills and claimed that college students were less well-prepared than 10-15 years ago due to the “watering down” of public high school curricula (Astin, 1982, 46).

Most of the commissioned papers, however, were optimistic and supportive of education as providing significant gains in a variety of areas. Many of them cast doubt on cross-national comparisons of test scores, and advocated a more open or liberal approach than what was finally suggested in *A Nation at Risk*. Martin Maehr, for instance, wrote that there was little evidence that the current generation had lost

the achievement ethic (Maehr, 1982), and Donald Holsinger reported that there were no significant variations in achievement between the U. S. and other countries (Holsinger, 1982). Barbara Burn and Christopher Hurn wrote that a comparison between achievement scores in different countries was not valid because America had different values and expectations of education. The American system was a highly diversified one, according to these researchers, and valued equality, utility, and practicality. Burn and Hurn stated that American students were individualistic and suspicious of government authority, so comparisons between countries were not very useful (Burn & Hurn, 1982).

Several commissioned papers recommended more freedom of choice in schools rather than an emphasis on basics. In her paper, Deborah Stipek advocated a somewhat Deweyan approach to education with hands-on experiences (Stipek, 1982). Beatrice Ward, John Mergendoller, and Alexis Mitman recommended more practice for high school students in making elective choices, more cooperative learning experiences, and more emphasis on social skills (Ward, Mergendoller, & Mitman, 1982). Frederick Rudolph also suggested extensive choices for students and faculty and celebrated the ways in which American education disperses control (Rudolph, 1982). Opposition to tracking and pull-out programs was expressed by Thomas Good (Good, 1982), and Zelda Gamson condemned overzealous budget-cutters because working people and women suffered most from cuts (Gamson, 1982). Several authors questioned the value of imposing standards by external examinations as other countries do because Americans value encouraging the average and motivating the less-capable rather than shoring up a set of centrally-determined, unyielding standards. These more liberal views and recommendations were nevertheless ignored in *A Nation at Risk*. The selective use of evidence and contentions found in the commissioned papers was accompanied by a selective use of testimony from public hearings. As stated earlier, the issues to be discussed were limited at the outset.

The Commission also selectively addressed concerns of its own members. Pull-out programs were attacked extensively in the second public hearing, and when corporate influence on schools was criticized by Crosby (EAC) in the Denver hearing, Campbell (MAC) responded as the chairperson with “It's a complex problem I think being pointed out. Are there other questions?”[10] which effectively cut off further discussion. In addition, Sommer (JS) questioned whether the educational situation was really as bad as it had been portrayed. Yet, when *A Nation at Risk* was released, it did not mention the hesitation on the part of these educators to increase testing, involve corporations in decision-making, or depict the educational situation as seriously flawed.[11]

The first public hearing held by the NCEE was devoted to science, mathematics, and technology education. Seaborg (GTS) stated in his opening remarks that “the strength of our technological and scientific enterprise will do much in the coming decades to determine the economic well-being, security, as well as the health and safety of Americans.” Although Seaborg was interested in the “competence of professional scientists and engineers,” he was also concerned with,

> “… the overwhelming remaining proportion of the population who do not become professional scientists and engineers, but who need to understand science and

> technology if they are to function effectively as technicians, repairmen and technology users in business, government and the armed services."

A definite degree of elitism crept into his account that bemoaned the lack of scientific and technological knowledge of working people and the middle class. Seaborg employed phrases that differentiated professional scientists from average citizens. References to the education of the former included such phrases as "sophisticated" and "most capable," but the latter was frequently referred to as "the rest of the population."[12]

In the fourth public hearing, Clifford Sjogren, Dean of Admissions at the University of Michigan, contended that those who design curricula must "give them [gifted and talented students] a more rigorous educational program. And the curriculum, then, should be less flexible for those students than it is now."[13] Baker (WOB) agreed, and in the public hearing held in Cambridge he spoke about encouraging genius in science and engineering. Elitism was visible in remarks made throughout this hearing that focused on gifted and talented education. Marcel Kinsbourne, stated that "The important point is that there is no kind of experience which can push up the potential set by the biological limitations of the brain, but there are many kinds of experiences that can pull it down." [14] There was a great deal of elitist rhetoric at this hearing expressed by William Oliver Baker about the importance of productively using the "best minds," the intellectual "lighthouses," the "great and glorious Miltons." [15] Despite the efforts of some, not much about race and class issues was explored in *A Nation at Risk* beyond a slight mention. The NCEE recommended that the federal government assume responsibility for educating gifted and talented, socioeconomically-disadvantaged, minorities, and handicapped students. Protecting the right of minorities to be educated was mentioned several times, but beyond this, little was said.

Virtue was mentioned prominently in the NCEE public hearing held in Denver in 1982. Invited guest speakers and Commission members discussed the problem at length, and Roy Forbes, the Assistant Director of Programs for Education Commission of the States and former Director of the National Assessment of Educational Progress, claimed that the first thing business asks for is "send us the people with the correct attitudes." [16] In the same meeting, a staff member representing the Department Chief of Naval Education and Training for Research and Development, Martha Brownlee, argued that attitudinal development was essential with respect to the military because people need to know how to handle a disciplined environment. Crosby (EAC) suggested that the problem is not necessarily limited to schooling:

> "I keep hearing attitude, attitude .... you might find that we [educators] had the same problem with that youngster, too, that we had a problem with attitude, and along with a poor attitude he probably had very poor attendance, very poor punctuality and consequently, he managed to get out not really having those particular skills." [17]

A long discussion about the importance of discipline and structure ensued. Later in the hearing, the Executive Director of the American Institute of Banking, Kathy Collins Smith, emphasized the need for "communication skills; customer relations, telephone etiquette; personal financial skills; interviewing skills

and personal appearance skills."[18] These "virtues" would contribute to a better "attitude" which in turn would benefit those in control of the workplace.

In reviewing values education in other countries, Judith Torney-Purta and John Schwille from the University of Michigan, stated in a commissioned paper for the NCEE that the Commission should "give extensive attention to values in its report" and what and how values should be taught in schools. A major issue for such an agenda, as for this paper," they claimed, "is how to reconcile striving for consensus with tolerance for dissent" (Torney-Purta & Schwille, 1982, 66). The authors argued that although teaching shared values might be important in elementary school, in the junior high schools, criticism should be taught. These suggestions for the teaching of criticism were not found in *A Nation at Risk*.

During the Chicago public hearing, the Assistant Vice President of Academic Services at Berkeley, Alice Cox stated that her university only admitted the upper 12 and a half percent of high school graduates in the state, and admission requirements were being raised still. "Some people would accuse us of being elitists." "In the 1950s," she said,

> "when there was a certain degree of excellence, at least in science and mathematics, in education, people who fed into institutions such as the University of California, came from traditional high schools. And by "traditional," I mean certain high schools. The university was not accessible to all people. That is one of our goals." [19]

Cox implied that university admissions are part of the ongoing class struggle, as did Clifford Sjogren who recognized that his support for Advanced Placement classes benefitted only certain groups and that it may be seen as "an elitist attitude." [20]

Numerous disparaging remarks were made by guest speakers and Commission members in the NCEE public hearings about specific class groupings that indicate class-specific concerns. For instance, Martha Brownlee stated in Denver, "Today, all of the [military] services are enjoying a productive recruiting environment, caused perhaps by our economic problems .... But it is disquieting to find that those reading below the ninth-grade level, which is the minimum to respond to safety instructions and administrative requirements and not fall over the pointy end of the ship, will be about 24 percent." [21] Crosby (EAC) broached the idea that there may be a conflict of interest between big business and education, though, since young people are playing video games and arcade games at all hours and business fights to keep those arcades open during school hours. Martha Brownlee stated that the conflict of interest is seen in the military as well since they have to work to separate their people from learned family values. Crosby said

> "We probably need to stop the corporations because they want kids from all over having good basic skills but then all of a sudden they decide to move down in lower Mississippi where they can get cheap labor because the people don't have those skills and they pay less taxes." [22]

This tendency for Commission members and guest speakers to stereotype based on class and race was visible in a number of disparaging remarks made by Kirk (RAK-AYK). With respect to immigrants during the first NCEE public hearing, she stated:

> " … just as you have people who have come here from other countries, immigrants who never learn to speak the language, and so consequently they are always with a problem and then their children go on. Are we going to have a generation of people who never will be able to use computers, and then the next generation, in a sense, will be trained in that thought process?"

"There's a lot of backwoods people in this country," she argued.[23] In the third hearing, Kirk also stereotyped women and remarked that once women reviewed all of their opportunities and discovered that they were capable of going into engineering, law, or medicine, "they may decide that they are really happier teaching." [24]

The interlocking network that connects some of the authors of *A Nation at Risk* demonstrates the workings of the power elite in a class-specific context of policy-construction. Most important, the selective use of evidence provided by commissioned papers and participants in the public hearings promotes the inference that the development of education reform was class-specific.

A CLASS POWER ANALYSIS

Education policy was, in 1983, constructed within a specific political and economic climate, not within a social vacuum. Policy was intricately tied into class power, and the state agents and mechanisms that developed it cannot be seen as neutral. State actions and the individuals involved in them frequently serve, intentionally or unintentionally, the dominant economic class. Under these circumstances, education policy in 1983 was shaped by an interlocking elite with particular views of economic problems and educational solutions to those perceived problems. The views espoused ruled out dissenting opinions and ultimately established the terrain for debate about American education which indirectly resulted in NCLB.

In 1983, education policy that did not need Congressional approval presented what appeared to be an inexpensive, easily accomplished, relatively conflict-free way of addressing economic problems. The conservative economic agenda was solidified under Reagan into a pro-big-business, anti-labor agenda. The reactionary political and economic climate of the 1980s established conditions conducive to reactionary ideologies that supported education reform for the sake of corporate and elite interests.

The education reform report is an artifact of an on-going class struggle. The role of the state, at this time, was to arrange for conditions conducive to economic growth which benefitted a specific economic class and enhanced productivity for the sake of a specific class. The role of education was ideological in the sense that the correct values must be taught; it was material in the sense that stability and productivity are essential to continued accumulation, and those would require an obedient and productive work force; and, it was political in the sense that certain groups needed specific conditioning, and

government could play a part in that. To maintain its legitimacy and its democratic character, however, the state must devise policy that does not look biased.

The content of *A Nation at Risk* was informed by the papers that were commissioned and the public hearings that were held, along with the testimony of the expert community. Some papers carried more weight than others in the final recommendations, however, and some testimony was relied upon more readily. This information helped the authors of the report to clarify and flesh out their interests during the policy-formation process. The conservative political and economic context of the 1980s helped promote a conservative ideology that worked to the detriment of working people, women, and people of color; it helped individual authors, affiliated with the dominant class, develop their consciousness in conjunction with those of capitalism; it made it easier to silence or convert interests that might be contentious; and, it supported a specific set of material interests.

The report was a product of a relatively unified power elite rather than a product of some sort of wide-spread, public consensus. A class power analysis puts into perspective the policy-formation process as part of a class struggle. Future efforts to analyze education reform may be more robust if augmented with power structure research and interpreted through a Marxian lens. These analyses may be necessary, if not sufficient, for this task. *A Nation at Risk* should be seen as an important step in the redesigning of education debates to support dominant class interests. Viewing *A Nation at Risk* in this context allows education researchers to understand the connections between NCLB and the ways in which the terms of contemporary debate have historically been determined by specific interests.

ENDNOTES

1. Several reports issued that year made similar connections. See especially, Task Force of the Business-Higher Education Forum, *America's Competitive Challenge: The Need for a National Response*, (Washington, D. C.: Business-Higher Education Forum, 1983); Task Force on Federal Elementary and Secondary Education Policy, *Making the Grade*, (NY: Twentieth Century Fund, 1983); and, Task Force on Education for Economic Growth, *Action for Excellence: A Comprehensive Plan to Improve our Nation's Schools*, (Denver: Education Commission of the States, 1983).

2. *A Nation Reformed? American Education 20 Years After A Nation at Risk*, (Cambridge: Harvard Education Press, 1983) is a case in point. The volume is a collection of essays, edited by David T. Gordon, that, for the most part, do not approach questions about political and economic roots of reform.

3. W. B. Gallie, "Essentially Contested Concepts" in *Proceedings of the Aristotelian Society 56*, 1955-1956, and Russell L. Hanson, *The Democratic Imagination in America: Conversations with our Past* (Princeton: Princeton University Press, 1985). Both emphasize the historical contingency of the terms.

4. Although Issac does not define himself as a Marxist, his definition of power is relevant for this essay.

5. Domhoff emphasizes that his perspective is not Marxist.

6. It is important to note that simply because an individual received a grant or fellowship from a specific corporation, university, or foundation, interest alignment with the goals of the granting agency should not be inferred. It must also be noted that not all affiliations are equally influential in developing class interests. This essay does not attempt to precisely measure the relative influence of diverse institutions on establishing class interests.

7. Fiske, 1983, B-6. Referring to Kirk as the parent representative is highly misleading. Annette Kirk was the wife of Russell Amos Kirk, a friend and colleague of William F. Buckley, Jr., a well-known and highly respected conservative. Russell Amos Kirk authored several books and articles about conservatism; held meetings, dinners, and weekend retreats for young conservatives at the family residence; supported many young conservatives' endeavors; and, had a remarkable impact on the perspective. (See *The Intelligent Woman's Guide to Conservatism*, (NY: Devin-Adair, 1957) and *Prospects for Conservatives* (Washington, D. C.: Regnery Gateway, 1989). This is not meant to imply that Annette, as his wife and business partner, unequivocally shared his perspective. It would be naïve to assume, however, that Russell had no impact on her views given many of the conservative remarks Annette made in the NCEE meetings. In this essay, Annette's initials are combined with Russell's (AYK-RAK), but some of the organizational connections were his.

8. See, for example, Domhoff, 2002, 83-88.

9. A copy of this letter and all citations from the NCEE public hearings can be found in Bell's personal papers housed in Special Collections in the Marriott Library at the University of Utah in Salt Lake City. For this letter, reference August 26, 1981. Hereafter, references to hearings will cite only author, city in which the hearing was held, and page number embedded in text. Commissioned papers are listed in the references.

10. Campbell, Denver, 59.

11. After the report was issued, Representative Carl D. Perkins (a Democrat from Kentucky and Chairperson of the House Labor and Education Committee) conducted an informal survey of the authors of *A Nation at Risk*. Eight out of the 10 members who replied to Perkins' survey disagreed with the interpretation offered by White House. The respondents said that there was a need "for more---not less federal support for education." Perkins also asked whether the members "backed the president's proposals for reduced spending and abolition of the Department of Education." The answer was an "emphatic" no. See *The Washington Post*, August 5, 1983, A-19.

12. Seaborg, Stanford, 10-11.

13. Sjogren, Chicago, 24.

14. Kinsbourne, Cambridge, 33.

15. Baker, Cambridge, 89.

16. Forbes, Denver, 29.

17. Crosby, Denver, 54-55.

18. Smith, Denver, 83-84.

19. Cox, Chicago, 42, 112-113.

20. Sjogren, Chicago, 26.

21. Brownlee, Denver, 42-43.

22. Crosby, Denver, 59.

23. Kirk, Stanford, 259, 289.

24. Kirk, Atlanta, 78.

REFERENCES

Astin, Alexander W. 1982. "Excellence and Equity in American Education." Paper submitted to NCEE (Washington, D. C.: National Institute of Education).

Bell, Theodore H. (1984). "American Education at a Crossroads," *Phi Delta Kappan*, April:531.

Burn, Barbara B. & Hurn, Christopher H. 1982. “An Analytic Comparison of Educational Systems.” Paper submitted to NCEE (Washington, D. C.: National Institute of Education).

Gamson, Zelda. 1982. “A Little Light on the Subject: Keeping General and Liberal Education Alive.” Paper submitted to NCEE (Washington, D. C.: National Institute of Education).

Good, Thomas L. 1982. “What is Learned in Schools: Responding to School Demands in Grades K-6.” Paper submitted to NCEE (Washington, D. C.: National Institute of Education).

Holsinger, Donald B. 1982. “Time, Content, and Expectations as Predictors of School Achievement in the U. S. A. and Other Developed Countries: A Review of I. E. A. Evidence.” Paper submitted to NCEE (Washington, D. C.: National Institute of Education).

Maehr, Martin L. 1982.. “Motivational Factors in School Achievement.” Paper submitted to NCEE (Washington, D. C.: National Institute of Education).

Mann, Michael. 1986. “The Ideology of Intellectuals and Other People in the Development of Capitalism” in *Stress and Contradiction in Modern Capitalism,* eds. Leon Lindberg et. Al. (Lexington: Lexington Books).

Mills, C. Wright. 2000. *The Power Elite*. (Oxford: Oxford University Press).

Resnick, Lauren B. & Resnick, Daniel P. 1982. “Standards, Curriculum, and Performance: An Historical and Comparative Perspective.” Paper submitted to NCEE (Washington, D. C.: National Institute of Education).

Rudolph, Frederick. 1982. “Education Excellence: The Secondary School-College Connection and Other Matters: An Historical Assessment.” Paper submitted to NCEE (Washington, D. C.: National Institute of Education).

Stipek, Deborah. 1982. “Motivating Students to Learn: A Lifelong Perspective.” Paper submitted to NCEE (Washington, D. C.: National Institute of Education).

Torney-Purta, Judith & Schwille, John. 1982. “The Values Learned in School: Policy and Practice in Industrialized Countries. Paper submitted to NCEE (Washington, D. C.: National Institute of Education).

Ward, Beatrice, John R. Mergendoller, & Alexis L. Mitman. 1982. “The Years Between Elementary School and High School: What Schooling Experiences do Students have?” Paper submitted to NCEE (Washington, D. C.: National Institute of Education, 1982).

United States Monetary Policy during the Financial Crisis, Credit Facilities Development and Potential

Marwan El Nasser and Richard Robinson*

ABSTRACT

In response to the financial crisis, the Federal Reserve developed new credit instruments during late 2008 and early 2009 for the purpose of supporting the liquidity of financial institutions, and restoration of credit flows. These instruments represented a new method of directly funding targeted financial markets. Bank credit flow, however, has been anemic during the crisis and recovery period. It is argued here that the precedent established by the Federal Reserve for direct credit market intervention is currently needed to support the substantially lagging bank loans. Specific rationales for the evolution of the suggested tools of credit easing are presented.

INTRODUCTION

The U.S. and global financial markets became illiquid and volatile after August 8, 2007. During the third quarter of 2008, the normal efficient operation of the financial markets came to a halt, private intermediation among banks, and also between banks and depositors, virtually stopped. The portfolio composition of U.S. commercial banks, dominated by agency and GSE securities, became illiquid. Banks became unable to securitize loans. Bank capital came under pressure because of the increased write-downs on securities and rising loan- loss provisions.

In the beginning of the financial crisis, banks managed by borrowing from their foreign branches. This strategy helped as internal borrowing from foreign operation contributed more than 20 percent of domestic asset growth in the second half of 2007, a portion twice as large in the previous year. An additional major source of bank liquidity came from bank borrowing from Federal Home Loan Bank. This source also contributed more than 20 percent of total bank asset-growth in the second half of 2007. The Federal Reserve also tried to provide funds through the discount window where the spread was dropped 50 basis points, and banks were explicitly invited by the Fed to use the discount window. Banks were reluctant to use the discount window, however, because of the “stigma” traditionally attached to using this source, and due to easier borrowing conditions from the Federal Home Loan Bank. Discount Window borrowing was negligible in 2007.

The deteriorating bank liquidity resulted in the Federal Reserve’s bold movement of creating numerous new arrangements beyond the traditional tools of open market operations and discount loans. The Fed aggressively expanded its lending and purchasing of assets with special focus on impacting selected financial markets. The aggressive Federal Reserve lending through the nontraditional facilities

* Professor Robinson is Professor of Finance and Professor El Nasser is a Professor of Economics, both of the School of Business, State University of New York at Fredonia, Fredonia, NY 14063

resulted in more than doubling of its assets. In August, 2007, the Fed's assets were $869 billion, but in December 2008, these assets had grown to $2,268 billion.

Fed's assets remained at high levels during 2009 and 2010. In April, 2010, the Fed's assets were $2,357 billion, but bank excess reserves grew to over $1 trillion during this period. Money supply measures M1 and M2, however, were virtually unchanged during most of 2008 and only slightly larger beginning the third quarter of 2008 up to the present (December 2010). The massive growth of bank reserves and the monetary base, and the small growth of the money measures, indicate that bank credit and the money supply multipliers virtually collapsed. Clearly, the data show that the U.S. economy has plunged into a Keynesian type bona fide liquidity-trap.

On November 2008, the Federal Reserve began unwinding some of the nontraditional credit facilities from its balance sheet, but in the meantime, it began adding mortgage-backed securities. Bank total reserves and excess reserves, however, remained roughly constant at historically high levels.[1]

This recent history prefaces the current and important debate regarding monetary policy. The Federal Reserve faces the dilemma of how and when to unwind the extraordinary monetary accommodation and absorb the huge quantity of excess reserves that it created to combat the financial crisis. Doing this prematurely risks tipping the economy into another financial crisis, but delaying too long risks substantial inflation.

The purpose of this article is to re-examine the Fed's current options for monetary policy, especially with respect to an extension of its credit facilities rather than additional increases in the monetary base. We argue here that renewal of the specially designed credit facilities is not only feasible, but would avoid further increases in bank excess reserves. We argue that this extension of the new policy tool of credit facilities would be a more viable method of stimulating the flow of bank credit, which up to now has been the key factor in inhibiting a more robust economic recovery. In order to fully compose our argument concerning the viability of an extension and development of the Fed's credit facilities, however, we must review in more detail the recent history of the Fed's involvement in this new instrument of monetary policy.

COLLAPSE OF THE MONEY SUPPLY PROCESS

Beginning in August, 2007, the rise in delinquencies on subprime mortgages triggered cascading effects that led to the collapse of the financial markets. As would be expected, the Federal Reserve responded to the crisis by employing traditional monetary tools. The Fed cut the discount rate in August of that year, followed by the Federal Open Market Committee easing monetary policy in September by reducing the target for the federal funds rate by 50 basis points. The continued financial turbulence led the Committee to reduce its target for the federal funds rate by a cumulative 325 basis points by the Spring of 2008. Though this policy response was rapid and bold and may have slowed down the financial turmoil and the deteriorating economy unfortunately, the intensification of the financial instability led to

further deterioration in the economic outlook. The Committee responded by further cuts in the target for the federal funds rate to 0-25 basis points in December.[2]

The exceptional circumstances of 2007-2008 interfered with the Open Market Desk's ability to implement monetary policy. Credit spreads widened, lending standards became more restrictive, and credit market led to tighter financial conditions overall. In particular, many traditional funding sources for financial institutions and markets vanished, and banks and other lenders found their ability to securitize mortgages, auto loans, credit card receivables, student loans, and other forms of credit greatly curtailed. As extensions of credit came to a halt, the normal money supply process collapsed, and the money supply measures of M1 and M2 remained virtually unchanged despite bank reserves and the monetary base drastically increasing. This is to say that the money supply multipliers with respect to the monetary base collapsed. Any reserves that were added to the banking system were held as excess reserves.

The process of additional bank lending has necessary prerequisites:

- The central bank must create additions to the monetary base.
- The required reserve ratio against deposits must be less than 1. (This allows the possibility of excess reserves.)
- Banks desire to lend their excess reserves.
- There is a demand for new loans.

Post the initiation of the financial crisis; the first two of these conditions were clearly met by the Fed. There has, however, been some debate about the roles of the latter two. Bank reserves, excess reserves and the monetary base all drastically increased since August 2008. The money supply, however, remained virtually unchanged. This indicates the collapse in bank credit and the money supply multipliers for M1 and M2. The consensus, however, is that banks' unwillingness to lend caused the collapse of the money-supply transmission process. (See Carpenter and Demiralp, 2010, and Berrospide and Edge, 2010.) Figures 1, 2 and 3 illustrate this phenomenon.[1]

Figure 1: M1 and M2 Multipliers

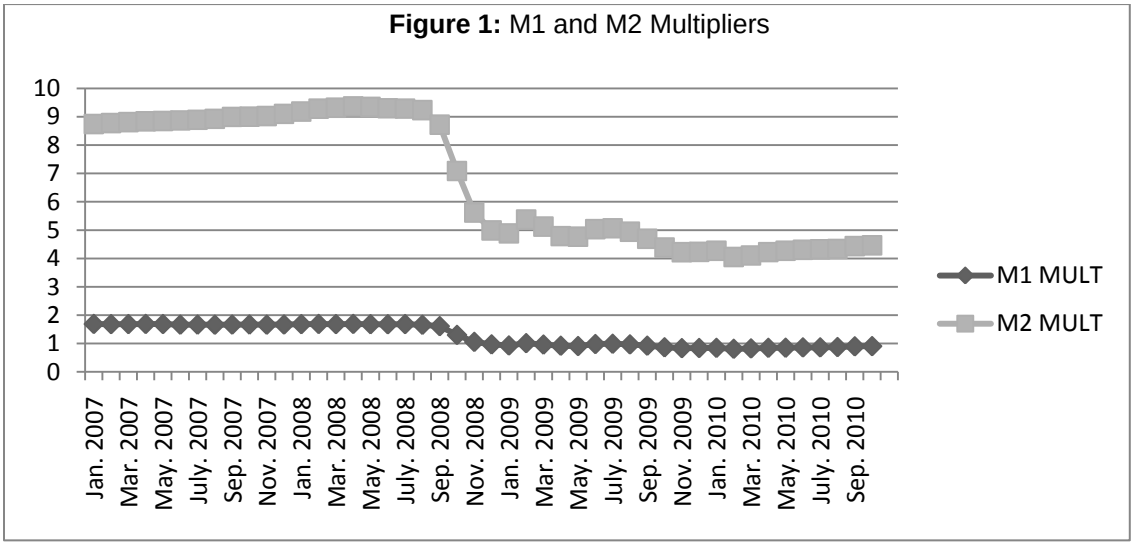


Figure 2: Excess Reserve Ratio= Excess Reserves/checkable Deposits

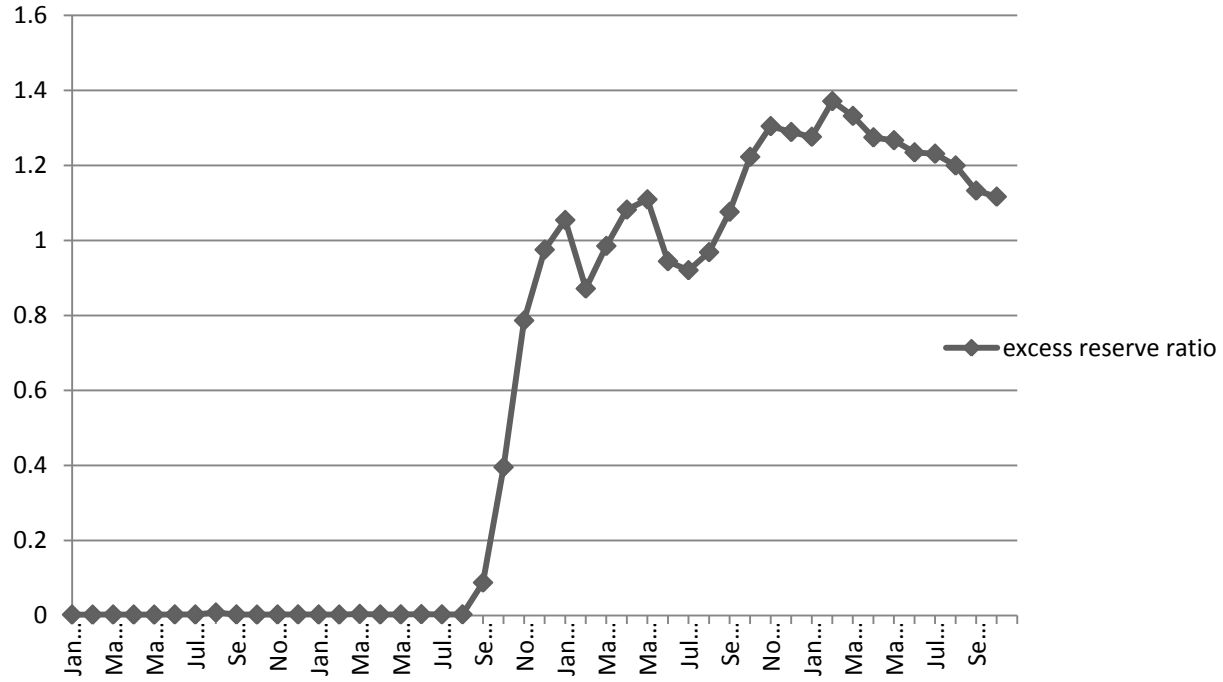
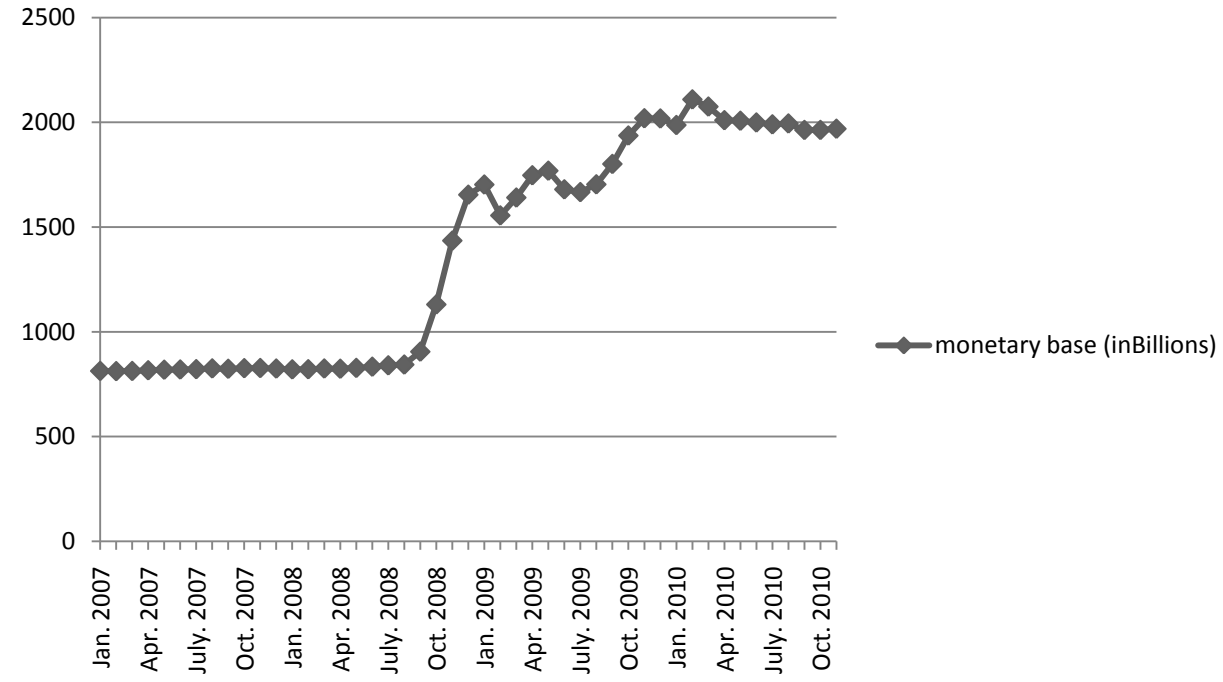


Figure 3: Monetary Base (in Billions) Jan 2007-Non. 2010



THE FEDERAL RESERVE'S RESPONSE TO THE FINANCIAL CRISIS

Bank creation of money came to a halt after the start of the financial crisis because of bank unwillingness to lend. The immediate challenge to policy makers was to reestablish the channels for bank credit, and to induce banks to lend the mountainous amount of excess reserves that the Fed supplied to the banking system.

The Fed's reduction in the target federal funds rate from 5-1/4 percent to effectively zero was an extraordinarily rapid easing in the stance of monetary policy.[2] In addition, however, the Fed implemented a number of programs designed to support the liquidity of financial institutions and foster improved conditions in financial markets. These new programs led to a significant change to the Fed's balance sheet.

The Federal Reserve explicitly identified two policy approaches to support credit markets: *quantitative easing* measures and *credit easing* measures.[2] Both approaches are similar in that they involve expanding the Federal Reserve balance sheet. However, the focus of the *quantitative easing* approach is on the expansion of bank reserves or the monetary base. The focus is on the liabilities side of the central bank balance sheet without regard to how the liabilities are created on the asset side. This policy approach was used by the Bank of Japan from 2001 to 2006. In contrast, the *credit easing* approach focuses on the type of loans and securities that the Fed holds, and on the importance of the mix of assets in improving credit conditions. The Federal Reserve focused its *credit easing* policies on reducing credit spreads and improving the functioning of credit markets. This approach bypassed the traditional gauges of the stance of monetary policy, i.e. the *quantitative easing* approach. This change in policies was necessary because the existence of large amounts of excess reserves indicates the ineffectiveness of *quantitative easing* because its link to money expansion has malfunctioned.

For this reason, the Federal Reserve created or redesigned many new lending facilities or tools.[1] The purpose of the new lending facilities was to address the funding difficulties of financial intermediaries and to improve market efficiency in short-term money markets. The new lending facilities were intended to strengthen the effectiveness of monetary policy.

The first set of tools consisted of enhancing or modifying the central bank's traditional role as the lender of last resort. This involves the provision of short-term liquidity to depository institutions. A second set of tools involve the provision of liquidity directly to key credit markets. The new facilities under this category include Commercial Paper Funding Facility, the Asset-Backed Commercial Paper Money Market Mutual Fund Liquidity Facility, the Money Market Investor Funding Facility, and the Term Asset-Backed Securities Loan. Each of these facilities is reviewed below. The third set of instruments includes the change in implementing the open-market operations. On November 25, 2008 the Fed initiated the outright purchases of Agency and Agency MBS debt. On March 2009 the Fed initiated the outright purchases of longer term Treasury Securities. For example, on November 25, 2008, the Federal Reserve announced plans to purchase up to $100 billion in government-sponsored enterprise (GSE) debt and up to $500 billion in mortgage-backed securities. On March 18, 2009, the Fed announced plans to purchase up to

$300 billion of longer-term Treasury securities in addition to increasing its total purchases of GSE debt and mortgage-backed securities up to $200 billion and $1.25 trillion, respectively.[3]

THE FED'S NEW POLICY TOOLS AND CREDIT FACILITIES

The new credit facilities were aimed at:

1. Broadening the eligible set of borrowers and/or expanding the types of collateral accepted against extensions of Federal Reserve credit.
2. Directing credit directly to particular markets that are under severe stress.
3. Purchasing assets directly so as to influence key longer-term interest rates such as mortgage rates.

The following presents a listing of the nine new credit facilities and policy tools created by the Fed in response to the credit market collapse:

(1) Term Auction Facility (TAF): On December 12, 2007, in view of the serious illiquidity of term bank-funding markets, and the hesitation of banks to use the discount window because of the associated stigma, the Federal Reserve announced the creation of the TAF. Under this program, the Fed auctions term funds of either one-month or three-months, up to $150 billion per auction for depository institutions. All depository institutions that are eligible for discount loans are eligible to participate in TAF auctions. All advances are fully collateralized.

(2) Term Auction Facility (TAF): The TSLF was initiated by the Fed to deal with the illiquid repo markets.

(3) Primary Dealer Credit Facility (PDCF): The PDCF provides an alternative source of financing for dealers, and has proven useful in the wake of the takeover of Bear Stearns and the bankruptcy of Lehman Brothers. Its future in normal times is an important policy issue that has not yet been fully resolved.

(4) Swap Line (SL): This facility was created to deal with illiquid money markets that became segmented across countries and time zones. Under the swap agreements with various central banks, the Federal Reserve lends money to other central banks, collateralized by the other countries currency, where those central banks lend to commercial banks in their jurisdictions.

(5) Commercial Paper Funding Facility (CPFF): The Federal Reserve created the CPFF to deal with illiquidity in the short-term commercial paper funding markets. Under the CPFF, the Federal Reserve Bank of New York, through its primary dealers, financed the purchase of both unsecured and asset-backed commercial paper from eligible issuers. The CPFF financed only highly rated, U.S. dollar-denominated, three-month commercial paper.

(6) Asset-Backed Commercial Paper Money Market Mutual Fund Liquidity Facility (ABCPMMMFLF): The Federal Reserve initiated this facility to help money market funds that hold asset –backed commercial paper to meet demands for redemptions by investors. This facility provided funding to

U.S. depository institutions and bank holding companies to finance their purchases of high-quality asset-backed commercial paper from money market mutual funds.

(7) *Money Market Investor Funding Facility (MMIFF):* As money market mutual funds and other investors increased their liquidity positions by investing in shorter-term (frequently overnight) assets, short-term debt markets became under considerable strain. In response to easing the short-term debt strain, the Federal Reserve created the MMIFF. Under this facility, the Federal Reserve Bank of New York provided senior secured funding to finance the purchase of certain money market instruments from eligible investors. Eligible assets included U.S. dollar-denominated certificates of deposit, bank notes and commercial paper issued by highly rated financial institutions. By facilitating sales of money market instruments in the secondary market, the MMIFF gave money-market mutual-funds the confidence that they could extend the terms of their investments and still maintain appropriate liquidity positions.

(8) *Term Asset Backed Securities Loan facility (TALF):* The TALF is a funding facility that helped market participants meet the credit needs of households and small businesses by supporting the issuance of asset-backed securities collateralized by loans of various types. Under the TALF, the Federal Reserve Bank of New York loaned $200 billion on a non-recourse basis to holders of AAA-rated loans backed by newly originated consumers and small businesses. The U.S. Treasury Department, under the Troubled Assets Relief Program (TARP) of the Emergency Economic Stabilization Act of 2008, provided $20 billion of credit protection to the Fed in connection with the TALF.

(9) *Agency Mortgage-Backed Securities Purchase Program:* On November 25, 2008, the Fed announced plans for the outright purchases of Agency and Agency MBS debt. It purchased direct obligations of Fannie Mae, Freddie Mac, and the Federal Home Loan Banks, and mortgage-backed securities (MBS) guaranteed by Fannie Mae, Freddie Mac, and Ginnie Mae. The Agency and Agency MBS purchase phase was completed on March 31, 2010. As of May12, 2010, the Fed's holdings of MBS were $1.068 trillion, and $168 billion of Federal Agency Debt.[3]

THE EFFECT OF NEW NONTRADITIONAL TOOLS ON THE FEDERAL RESERVE BALANCE SHEET

Extending Federal Reserve credit through the new lending facilities expanded the Federal Reserve balance sheet. The factors supplying reserve more than doubled from $902,993 on August 8, 2007 to $2.3 trillion on September 1, 2010. Almost all of the expansion in reserves occurred in excess reserves. Bank reserves, together with currency, make up the monetary base which increased significantly as the Fed's balance sheet expanded. However, bank credit declined or remained stationary as the money multiplier collapsed because banks chose to leave the great bulk of their reserves as idle excess reserves.

The bloated total assets of the Federal Reserve and the more than one trillion dollars in bank excess reserves present two important issues:

1. The longer-term issue concerns the Fed exit strategy of reducing excess reserves to normal levels so as to avoid inflationary bursts. Robinson and El Nasser (October, 2009), Robinson and El Nasser (Summer, 2010), and Wheelock (September, 2010) concerned this potential difficulty.
2. The shorter-term issue concerns the persistently almost stagnant or slow economic recovery with high rate of unemployment in the face of the dormant huge excess reserves. This is the subject of our current examination in this paper.

LOAN MARKET COLLAPS AND DIRECT INTERVENTION

The financial crisis and consequent recession was particularly marked by a collapse of loan markets. Table 1 presents the supporting data for this claim. Although economic recovery began in 2009, it continued at an anemic rate through 2010, at least anemic relative to historical growth rates for recoveries.

Table 1: Percentage Change in Loans (Seasonally Adjusted) Generated by US Commercial Banks 2005-November 2010*

Time Period	Commercial & Industrial	Real Estate	Consumer
2005	13.9%	15.4%	8.2%
2006	13.7	10.4	.8
2007	18.3	6.8	9.6
2008	14.3	.2	7.1
2009	-18.6	-5.4	-3.9
2009-I	-12.7	-1.0	7.9
2009-II	-16.4	-2.3	-9.5
2009-III	-27.3	-8.7	-4.8
2009-IV	-23.7	-9.9	-9.0
2010-I	-19.5	-8.1	-18.3
2010-II	-15.6	-6.5	-3.3
2010 April	-21.2	-5.9	.6
2010 May	-9.9	-4.9	-9.0
2010 June	-7.0	-5.9	-1.1
2010 July	-2.0	-8.3	-.1
2010 August	-1.5	-2.7	-5.1
2010 September	-5.5	-4.5	-17.1
2010 October	1.3	-5.2	-6.2
2010 November	2.3	-1.1	-9.1

*Series H8 Federal Reserve Bank Release, Board of Governors of the Federal Reserve.

In his Jackson Hole address, Chairman Bernanke cites the increased risk aversion of financial intermediaries as the cause of the collapse of these loan markets. He cites the continuation of this collapse as due to both the increased risk aversion and the decrease in loan demand on the part of business and consumers. Note that the savings rate is currently 6% as compared to an expected 4% as pointed out by the Chairman. The anemic economic recovery cannot be rectified without a resolution of this business and consumer loan collapse.

Under Term Asset-Backed Securities Loan Facility (TALF*)* and the Commercial Paper Funding Facility, the Fed funded short-term loans originated by dealers. Through these intermediaries, the Fed purchased securities that were collateralized by high credit-worthy loans. In essence, the Fed provided loanable funds to targeted credit worthy customers through intermediaries. We suggest that this process could be utilized again with commercial banks being the intermediaries, and credit worthy consumer and businesses being the targeted customers.

THE FUTURE USE OF CREDIT FACILITIES

The Fed is currently paying extremely low interest rates on reserves (approximately 15 basis points). The excess reserves certainly provide a long-term problem for control of the money supply as cited by Robinson and El Nasser (2009 and 2010) and Wheelock (2010). Directly providing more funds to the loan market would theoretically exacerbate this longer-term monetary control problem, i.e. longer-term inflation rates are a concern. The Fed, however, currently holds approximately $.8 trillion in Treasury securities which could be used to counteract any direct loan-market facility intervention. The Fed could simply sell at least some of these Treasury securities to fund purchases of securities from commercial banks, securities which are backed by (1) commercial and industrial loans, (2) real estate loans, and (3) consumer loans, the three loan areas that particularly collapsed and are continuing at subpar levels. The precedent for this sort of program was established for commercial paper and other asset markets during 2008-09. A similar program for the proposed targeted credit-worthy consumer, commercial, industrial and real estate loans which are clearly inhibiting recovery due to their low levels is currently warranted. Furthermore, such a program could not only be instituted by the Fed, but conducted in such a way as to be neutral with respect to the monetary base, i.e. leave the base unchanged and hence not exacerbate any future problem of inflation control.

This sort of program was established for commercial paper and other asset markets during 2008-09. A similar program for the proposed targeted credit-worthy consumer, commercial, industrial and real estate loans, which are clearly inhibiting recovery due to their low levels, may currently be warranted. This would be an extension of a newly created tool of monetary policy, i.e. the use of credit facilities rather than traditional quantitative easing. Furthermore, such a program could not only be instituted by the Fed, but conducted in such a way as to be neutral with respect to the monetary base, i.e. leave the base unchanged and hence not exacerbate any future problem of inflation control.

ENDNOTES

1. Credit and Liquidity Programs, and the Balance Sheet, Monetary Policy, Board Of Governors of the Federal Reserve, http://www.federalreserve.gov/monetarypolicy/bst.htm
2. Ben S. Bernanke. “Chairman of the Federal Reserve Board, the Crisis and the Policy Response,” At the Stamp Lecture, London School of Economics, London, England, January 13, 2009 http://www.federalreserve.gov/newsevents/speech/bernanke20090113a.htm

3. Detailed information about the Federal Reserve's balance sheet is published weekly as part of the H.4.1 release. For a summary of Fed lending programs, see *Forms of Federal Reserve Lending to Financial Institutions* (229 KB PDF).

REFERENCES

Altig, David. 2009. "Federal Reserve Policy During the Financial Crisis," presentation to the *Symposium Examining the Cause, Effects,& Consequences of the Financial Crisis,* sponsored by Financial Decisions, *et al,* Birmingham Alabama, October, 2009.

Bernanke, Ben Sb January, 2009. "The Crisis and the Policy Response." Stamp Lecture, London School of Economics, http://www.federalreserve.gov/newsevents/speech/bernanke20090113a.htm.

Bernanke, Ben S.a. August, 2010. "The Economic Outlook and Monetary Policy," address presented at the Federal Reserve Bank of Kansas City Economic Symposium, Jackson Hole, Wyoming.

Mishkin, Frederic. 2010. *The Economics of Money, Banking & Financial Markets*, 9th edition, Addison Wesley, Imprint of Pearson, Boston, Mass.

Robinson, Richard.a October, 2009. "Escaping the Trap," presentation to the *Symposium Examining the Cause, Effects , & Consequences of the Financial Crisis,* sponsored by *Financial Decisions*, *et al,* Birmingham Alabama, October, 2009.

Robinson, Richard.b and Marwan El Nasser. Summer, 2010. "Escaping the Trap: Prospects for Federal Reserve Policy during the Recovery," *Financial Decisions*, Volume 22, Number 1.

Robinson, Richard.c and Marwan El Nasser. April, 2010. "The Liquidity Trap and the Federal Reserve Exit Strategy," presented at the Business Research Consortium of Western New York, Fifth Annual Conference, at SUNY Geneseo, and at the Pennsylvania Economics Association Annual Meetings, June, 2010. As earlier version was presented at the Social Science Research Colloquium, SUNY Fredonia, January, 2010.

Torres, Craig. March, 2009. "Bernanke , Kohn pledge Fed to Withdraw Credit when crisis ends," *Bloomberg*, *March 30, 2009,*

Warsh, Kevin M. September, 2009. "The Fed's Job is only Half Over", *Wall Street Journal*, September 25, 2009

Wheelock, David C. September, 2010. "The Monetary Base and Bank Lending: You Can Lead a Horse to Water," *Monetary Trends*, Federal Reserve Bank of Saint Louis, Research.stlouisfed.org.

A Chair's Guide to Student Evaluations of Teaching

William P. O'Dea*

ABSTRACT

This paper attempts to analyze the factors that influence student evaluations of teaching effectiveness. The student evaluation instrument at SUNY-Oneonta contains thirteen items. The thirteenth item is summative and asks the students to provide an "overall evaluation of teaching effectiveness." The first twelve items represent the different qualities that a good teacher should possess. Using ordinary least squares analysis, we find that 96 percent of the variation in the summative "overall effectiveness" item can be explained by the twelve qualities of good teaching. Adding additional variables such as the instructor's gender, years of experience and grade distribution does not improve the model's explanatory power and none of these variables are statistically significant. The results indicate that our students value instructors who deliver well-organized lectures, answer questions clearly and are available for consultation. Importantly, the results show that our students do not penalize faculty members who are rigorous and maintain demanding grading standards.

I. INTRODUCTION

As department chair, one of my major responsibilities is writing evaluation letters for my faculty members who are requesting contract renewal, tenure or promotion. In these letters, I am expected to assess the faculty member's performance in the standard areas of teaching, research and service. However, since the primary mission of the College at Oneonta is undergraduate education, the section of my letter dealing with teaching is especially important.

In assembling the folders that accompany their requests, we require our faculty to provide a portfolio of materials that can be used to evaluate their teaching. They are expected to provide copies of syllabi and examinations, their grade distributions, a self-evaluation of their teaching and two peer reviews based on classroom visits. Last, and by no means least, they must submit summaries of the student perception of instruction (SPI) questionnaires administered in all courses taught during the preceding four semesters.

By examining syllabi and reading examinations and peer reviews, it is possible to form a reasonably accurate impression of the level at which courses are being taught and the demands being placed on students. However, reading through this material takes a considerable amount of time and the impression formed is unavoidably subjective. The SPI data, by contrast, come in a readily digestible form that gives the illusion, if not the reality, of being scientifically precise. In addition, our SPI form has a bottom line question that asks the student to provide an "overall evaluation of teaching effectiveness." In my letter, I am expected to address the SPI data. Not to do so would place my faculty members at risk.

* Department of Economics, Finance and Accounting, State University of New York College at Oneonta, Oneonta, NY 13820. (607) 436-2127 odeawp@oneonta.edu

I am also a member of the College Promotion and Tenure Committee. Attempts to parse the meaning of the SPI data are at the heart of our discussions of teaching performance. Indeed, it has struck me that even members of the Committee who profess to place little credence in the SPI results still use them to support their arguments for or against a candidate's teaching effectiveness.

Having looked at the SPI results from thousands of course sections, I assume that in filling out the questionnaires our students are trying to tell us something. Whether that something is what we as an institution want to know is what is at issue. We want to evaluate the quality of teaching. Can the faculty member effectively convey a body of knowledge? Can the students apply the concepts developed in a course to relatively unstructured problems? Of course, we would also like our students to retain what they learn. Rephrased, do the SPI data provide insight into the quality of an instructor's teaching rather than into the instructor's popularity or entertainment value?

The outline for the reminder of the paper is as follows. The next section is devoted to a brief literature review. The literature on student evaluations of instruction is enormous. Unfortunately, for a chair looking for guidance, this enormous effort has produced relatively few hard answers. The third section discusses how the SPIs are administered at the College at Oneonta and presents summary SPI data along with data on instructor and course characteristics. The results of preliminary data analysis are also presented in the third section. The fourth section concludes.

II. AN IMPRESSION OF THE LITERATURE

The literature on student evaluations of instruction is huge. As of 1999, Becker and Watts (1999) report that "nearly 2000 studies have been conducted (p. 344)." Work in this area did not cease in 1999. Having neither the time nor the inclination to wade through 2000 academic papers, I decided to limit my attention to papers appearing in the *Journal of Economic Education* during the last 25 years. My premise is that, since the JEE is the leading journal in the field of economic education, reading articles appearing there should provide an accurate impression of the state of the art. By limiting my attention to the JEE I reduced my reading list to approximately 30 articles. Of these 30, I have read about 15. Consequently, this section represents my impression of the literature.

The literature appears to have three major concerns. The first is to identify the important explanatory variables that drive the student evaluations. The second is to establish the relation between grading standards and the student evaluations. Siegfried and Fels (1979) put the matter very bluntly: "By far the most controversial issue is whether instructors can 'buy' higher evaluations by lowering the (effort) price to students of achieving a given grade (p. 931)." SPIs became common in the middle to the late 1970s. Before this time, few colleges asked students to evaluate the quality of teaching. Now, very few colleges don't ask their students to rate their instructors (Becker and Watts, 1999). The period during which SPIs became increasingly common coincided with an increase in student grades. The temptation to infer a causal relation is obvious (Zaganzadeh, 1998). The third focus of the literature is to identify proper estimation techniques. The least cost solution to any estimated problem is to resort to ordinary least

squares (OLS). The problem is that OLS may not be the best technique to employ in the statistical analysis of students evaluations.

In addressing the first two concerns, Needham (1978) provides a simple theoretical framework that is useful. Students have utility functions whose arguments are leisure time and some academic result. For Needham, the academic result is the course grade. But there is no reason why it could not be the amount learned. We can visualize classes as being populated by a mixed group of students some of whom view the purpose of a college education as being the acquisition of a body of knowledge and skills and some of whom view the point as being the acquisition of a marketable credential. The utility function has the conventional convex shape. The student also has a concave production possibility frontier (PPF) that describes the rate at which leisure time can be converted into a grade (learning). The optimal point is where the PPF is tangent to the highest reachable indifference curve. Instructors are able to influence this process by rotating the PPF to the right (assuming that leisure is placed on the vertical axis). In other words, for a given sacrifice of leisure, the student can receive a higher grade (learn more). A rightward shift of the PPF enables the student to reach a higher indifference curve. An increase in student welfare would then lead to better evaluations.

Instructors can shift the PPF by fair means and foul. Instructors can improve the quality of their teaching. For example, they can deliver better organized and clearer lectures, provide more vivid examples, give more precise and direct answers to student questions, hold more frequent and convenient office hours, and tell better jokes to keep the students engaged and awake. These strategies would work for both types of student. Higher SPI scores in this scenario would indicate an improvement in the quality of teaching. The problem is that it is also possible to enable students to obtain higher grades with a smaller time commitment by relaxing standards. For example, instructors can give easier examinations or grade them more leniently. This strategy would be appreciated by the “education as a credential” group. However, the “body of knowledge” group would not appreciate the decrease in the quality of their education. Depending on which group of students is larger, the SPI ratings could either go up, down or remain constant.

In a setting where faculty members make up and grade their own examinations, which is the norm at the College at Oneonta, grades cannot be interpreted as an unambiguous measure of what the students have learned. High grades can be the product of effective instruction or lax standards. Low grades could reflect rigorous instruction or ineptitude.

With regard to the question of what instructor qualities are valued by students, the literature is reassuring. Boex (2000) reports that “from the students’ point of view, organization and clarity was the single most important attribute of effective economics instruction (p. 213).” Bosshardt and Watts (2001) find that students care about enthusiasm and preparation. DeCanio’s (1986) results highlight the importance of organization and structure. Nelson and Lynch (1984) find “clarity of communication and instructor enthusiasm” to be the most important factors explaining favorable student evaluations (p. 21).

With regard to the impact of grades on evaluations, the literature is somewhat reassuring. Giving higher grades does not appear to be a foolproof strategy for obtaining higher SPI ratings. Nelson and Lynch (1984) report that a one point increase in the average grade in a course (on a four point scale) only raises the instructor rating by .15 (on a five point scale). DeCanio (1986) finds "no evidence that the expected grade influences the instructor's teaching effectiveness score (p. 172)." On the other hand, Mirrus found that a one point increase in the average grade would lift the SPI rating by .85 of a point (cited in Siegfried and Fels, 1979). In Boex's (2000) study, the average grade has a statistically significant impact on the student ratings. However, other factors such as clarity of instruction and organization tend to be more important.

Of course, even if grades and SPI scores are positively related, this is not necessarily a problem if high grades are due to effective instruction that results in more learning. Again, the problem is that grades do not provide an unambiguous measure of what students have learned, since examination design and grading are under the control of the individual instructor. Of particular interest, then, are the papers by Carrell and West (2010) and Gramlich and Greenlee (1993). Both papers consider settings (required math and science courses at the United States Air Force Academy and the introductory and intermediate economic theory courses at the University of Michigan, respectively) where course content and examination design and grading are under department control rather than under the control of the individual instructor. In addition, at both institutions, students are randomly assigned into course sections. These are settings where the "teacher as coach" metaphor might be accurate. Carrel and West found that instructors whose students performed better on the common examinations tended to receive better student evaluation scores. Gramlich and Greenlee found a "statistically significant but small" relation between student grades and faculty ratings (p.4). Interestingly, Carrel and West found that instructors whose students performed well in introductory level courses did not tend to perform as well in upper-level follow on courses. They found that "student evaluations are positively correlated with contemporaneous professor value added and negatively correlated with follow-on achievement (p 412)." "Deep learning", which is the object of the educational enterprise, was not recognized and rewarded by the student evaluations.

The nature of the data sets used in the statistical analysis of student evaluations raises some interesting econometric issues. For studies where the data point is an individual student evaluation the discrete nature of the responses suggests that some sort of logit model rather than OLS might be the proper estimation techniques (DeCanio, 1986). If the data point is a course section, the average of student responses to the items on the SPI questionnaires can be continuous between the floor and the ceiling response. However, the existence of the floor and the ceiling still argues in favor of a logit approach. In addition, the existence of upper and lower limits might result in heteroskedasticity because the error distribution is truncated (Siegfried and Fels, 1979).

Moreover, a single equation estimation approach might not be appropriate (Mehdizadeh, 1990). For example, student grades or expected grades are commonly used as an explanatory variable in the

equation accounting for student ratings. However, given the control that instructors have over the grading process, it would be inappropriate to treat grades as an exogenous. Since the grades are endogenous, an equation characterizing the process by which grades are determined would be needed. These points having been made, I am going to use a simple OLS model for my initial data analysis.

III. BACKGROUND AND DATA ANALYSIS

At Oneonta, it is College policy that the SPI questionnaires be administered in every course every semester. Individual departments can use a College questionnaire or substitute an instrument of their own design. The Division of Economics and Business, which houses the Departments of Economics, Finance and Accounting (EFA) and Management, Marketing and Information Systems, uses its own questionnaire.

The timing and mechanics of administering the SPI instrument are up to the individual faculty member. Typically, the SPIs are administered during the last two weeks of the semester. Some brave souls administer the evaluations during the final examination period. Instructors are not supposed to administer the questionnaires. Some instructors recruit other members of the faculty to hand out and collect the forms. Other instructors explain the process and hand out the forms designating a student in the class to collect the forms and return them to the department office. The department secretary then types up the written comments. The forms are then sent to the Office of Institutional Research which generates summary statistics and provides analysis of the individual questions.

The questionnaire used by the Division of Economics and Business is relatively compact. It asks the students to answer thirteen questions and provide their GPAs. At other institutions, the students are asked to answer many more questions. For example, the Penn State form that provided the data for the Nelson and Lynch (1984) study contained 33 questions. The Georgia State University questionnaire asked the students to answer 35 questions (Boex, 2000). The questions we ask are fairly standard and deal with issues of classroom management, course organization and testing. We don't have a question dealing with instructor enthusiasm but we do ask about poise. Unlike the instruments at other institutions, we don't ask the students to provide their expected grade for the course or to rate the value of the course. Our bottom line question focuses on instructor effectiveness. On our form, one is excellent (or the equivalent) and five is poor. On the back of the form, the students can provide written comments. The student comments tend to be terse. If our students write more than a sentence or two, it is generally a sign that they are not pleased.

Table 1 provides data on the average scores for the thirteen questions for the course sections taught in EFA from the fall of 2008 to the spring of 2010. There were a total of 147 course sections. In general, the scores are slightly over two which means that on the" overall evaluation of teaching effectiveness" and most of the individual components of instructor performance our students think their instructors are (almost) very good.

One way to approach the form is to regard the first twelve questions as describing the qualities that we think that an excellent instructor should possess. To determine which of these qualities have the greatest impact on the bottom line "overall effectiveness" rating, I regressed "overall effectiveness" on the twelve qualities using OLS. The Koenker-Bassett test (Gujarati, 2003) indicates that heteroskedasticity is not present. Some of the correlations between the questionnaire items are high enough to suggest that multicollinearity could be a problem, but at this stage I have made no effort to address the issue.

Table 1: Summary Statistics Teacher Evaluation Questionnaire

Question	Mean	Standard deviation
Rigorousness of the course	2.28	.52
Organization of the Course	2.23	.61
Teaching skill is evidenced by classroom presentation	2.24	.69
How well instructional materials were coordinated with lecture	2.2	.64
Poise and self confidence	1.86	.54
Planning and clarity of examination questions	2.34	.66
Ability to answer questions and clearly explain concepts	2.14	.68
Tolerance of attitudes and opinions of others	1.88	.5
Ability to maintain control of class discussion	2.07	.53
Availability for consultation	1.96	.47
Maintains demanding grading practices	2.06	.43
Fairness in grading	2.11	.43
Overall evaluation of teaching effectiveness	2.25	.66

The results are shown in Table 2. Overall, our twelve qualities of effective teaching account for 96 percent of the variation in the "overall evaluation of teaching effectiveness "ratings, which suggests that our questionnaire includes the major qualities that our students believe that an effective instructor should possess. The single most important explanatory variable is "teaching skill as evidenced by classroom presentations." My interpretation is that our students value an instructor's ability to deliver a coherent presentation. The ability to answer questions, the planning and clarity of examination questions, the coordination of materials with the lecture, and availability for consultation were also very important. The results indicate that delivering a rigorous and demanding course and maintaining demanding grading practices tend to result in higher (worse) ratings on the "overall evaluation" item. However, the results are small and statistically insignificant. The results indicate that our students do not overly penalize instructors who hold them to high standards. It could be that our students expect economics, finance and accounting courses to be technical and difficult and are not shocked when this expectation is met.

To determine whether the existence of a floor and a ceiling for the ratings affected the results, I reran the regression dropping the observations at the very top (ratings of 1.5 and below) and the very bottom (ratings of 4 and above). The coefficient estimates did not change very much nor did the standard errors.

It is also possible that course and instructor characteristics affect the student evaluations. (Although the results in Table 2 indicate that there is not a lot of variation left to account for.) Table 3 presents summary data for course and instructor characteristics that have been employed as explanatory variables in other studies. The number of students enrolled in the course captures the possibility that students might prefer smaller classes because they allow more interaction between the instructor and the students. The percentage of surviving students who complete the SPI form accounts for the possibility that the sample of students completing the SPI forms might not be representative of the entire population of students in the course. Thus, the SPI data are subject to sample selection bias. Boex (2000) raises the possibility that students who are disappointed with a course are likely to stop attending class and thus will not fill out the SPI questionnaire. The result would be a higher overall SPI rating. Becker and Watts (1999) suggest that instructors might manipulate the response rate by deciding when to administer the form. For example, at our institution, an instructor might administer the evaluation on a Friday, which is a day with a relatively high absenteeism rate. Of course, if the group of students completing the questionnaires includes the most talented and hardworking students in the course, these would be the students who are best able to evaluate an instructor's effectiveness.

TABLE 2: OLS Regression Results

Variable	Estimated Coefficient	Standard Error	t-statistic
Constant	-.0742	.0814	-.9115
Rigor	-.0117	.04	-.293
Organization	.0758	.0652	1.1631
Teaching skill	.2982	.0846	3.5232**
Coordination of materials	.1528	.0628	2.4324*
Poise	.01984	.0505	.393
Planning and clarity of exam questions	.1635	.0515	3.1767**
Ability to answer questions	.2	.0561	3.5664**
Tolerance	.0133	.0419	.3185
Maintain Control	-.0104	.0352	-.2965
Availability for consultation	.1283	.0434	2.9539**
Demanding grading	-.0583	.053	-1.0994
Fairness of grading	.0938	.0441	2.12*

Number of observations=147 R^2 = .9657 adjusted R^2 = .9627 F = 315.0679
** significant at the 99 percent confidence level * significant at the 95 percent confidence level

Table Three also contains data on the level of the course. The introductory 100 level courses have very heavy enrollments of students who come from outside the Division of Economics and Business and are taking the course to satisfy general education requirements or a related work requirement for their majors. These students might be more interested in an easy grade than in what they learn. The 200 and 300 level courses are taken mainly by Business Economics, Economics, and Professional Accounting majors. The 200 level required courses are the intermediate economic theory courses, corporate finance, and business law, which are taken by Business Economics and Professional Accounting students to satisfy major requirements. Presumably, some of these students would not be very disappointed if, say, intermediate microeconomic theory were dropped from the list of major requirements. The 200 and 300 level major courses are taken by economics and accounting majors and students pursuing a concentration in finance, i.e. these course are taken by volunteers who should know what they are in for.

TABLE 3: Summary Statistics Class and Instructor Characteristics

Variable	Mean(Proportion)	Standard deviation
Enrollment	34.63	14.12
Withdrawals	2.32	3.05
Number of SPI forms completed	25.46	9.91
Percentage of survivors completing SPI form	80.45	13.4
100 level	42.86 (63/147)	
200 level required	31.29 (46/147)	
200 level major	8.84 (13/147)	
300 level major	17.12 (25/147)	
Median grade	2.82	.45
Male	65.36 (96/147)	
Non-native	15.65 (23/147)	
Rank (Assoc., full)	24.5 (36/147)	
Semesters of teaching experience at Oneonta	11.52	11.56

The average of the student rating for the overall effectiveness item is 2.54 (halfway between very good and good) for all 100 course sections. For 200 level required courses, 200 level major courses and 300 level major course, the overall rating is much lower (better); the ratings are 2.17, 1.9 and 1.81 respectively. The difference between the overall rating in the 100 level sections and the upper division sections is statistically significant. The median grade for all course sections taught in EFA is a 2.82 which is a little over a B-. The median grade for all course sections taught in the College is in the A-/B+ range. It is important to note that the median grade is based on the entire population of students who complete a course. This population is larger than the population completing the SPI forms. Again, the issue is whether the group of students completing the form is representative. Table 3 also provides data on the

gender and the nationality of the course instructor. The percentage of courses taught by senior faculty members is also shown. In order to be promoted to the ranks of associate or full professor at our College a faculty member must demonstrate outstanding proficiency in the area of teaching. (Of course, there is also the possibility that once promoted the faculty member can rest on his or her laurels.) Finally, to account for the possibility that instructors with more experience are more effective, the table provides data on the number of semesters of prior teaching experience at Oneonta. In the next step of my research, I intend to collect data on the time of day at which courses are taught, the average GPA reported by the students in a course section, and instructor research productivity.

Table 4: OLS Regression Results for Course and Instructor Variables

Variable	Estimated coefficient	Standard Error	t-statistic
Constant	2.34	.6281	3.729**
Enrollment	.0067	.006	1.1169
Response rate	.0062	.0041	1.514
200 Required	-.3635	-.3635	2.6951**
200 Major	-.1231	.2437	.5052
300 major	-.03546	.2302	1.54
Median grade	-.266	.1462	1.932+
Male	.0397	.1158	.3426
For	.495	.1745	2.8371**
Rank	-.4575	.1292	3.5406**
Semesters of Exper.	.0108	.0056	1.9318+

Number of observations=147 R^2 = .3075 adjusted R^2 = .2566 F = 6.0388
** Significant at the 99 percent confidence level + Significant at the 90 percent confidence level

To gauge the impact of these variables on the student evaluation, I regressed the "overall evaluation of teaching effectiveness" on the course and instructor characteristics. The results are shown in Table 4. The coefficients for the dummy variables for the upper division courses (the dummy took on a value of 0 for 100 level course sections) indicate that the "overall evaluations of teaching effectiveness" ratings are lower (better) in the upper division courses than they are in the 100 level courses. The difference is statistically significant for the 200 level required courses. The rank dummy variable, which takes on a value of 1 for courses taught by faculty at the associate and full professor ranks, shows that student ratings of faculty effectiveness are significantly better in sections taught by senior faculty. On the other hand, the dummy variable for course taught by instructors born outside of the United States (FOR) shows that student evaluations of faculty effectiveness are significantly higher (worse) in sections taught by non-native instructors. The results indicate that the more semesters of teaching experience an instructor has the higher (worse) his or her ratings trend to be. The estimate is significant at the ten percent level. The median grade coefficient indicates that one point increase in the median grade in a course section lowers

(improves) the "overall evaluation of teaching effectiveness" by approximately .25 of a point. The coefficient is significant at the ten percent confidence level. Gender of the instructor, the number of students enrolled in the course and the response rate on the SPI forms do not appear to matter much. As a final exercise, I then added four of the course and instructor variables to the list of instructor qualities in Table two. The variables I selected are: whether the course section was upper division (a dummy variable which takes on a value of one for a course taught at the 200 or 300 level), the median grade, whether the course was taught by a non-native instructor, and whether the course was taught by a senior faculty member. The results are presented in Table 5. Adding the course and instructor variables contributes no explanatory power to the model. Indeed, adjusted R^2 declines slightly. None of the course or instructor variables are statistically significant. Changes in the median grade now have a negligible and insignificant impact on the "overall evaluation of teaching effectiveness" rating. The implication is that the main factors driving a faculty member's rating on the "overall effectiveness" summative item are the twelve characteristics of effective teaching. For example, the higher (lower) evaluation of teaching effectiveness in 100 level course sections can be traced back to the fact that the students in the sections gave their instructors higher (worse) ratings in each of the twelve characteristics.

Table 5: OLS Regression Results

Variable	Estimated coefficient	Standard Error	t-statistic
Constant	.0785	.1304	.6024
Rigor	-.0081	.0422	.1923
Organization	.0795	.0705	1.128
Teaching skill	.3168	.095	3.3362**
Coordination of materials	.1343	.0714	1.8811+
Poise	.0227	.0521	.4359
Planning and clarity of exam questions	.1621	.0529	3.0679**
Ability to answer questions	.1953	.0602	3.2457**
Tolerance	.0024	.0438	.0548
Maintain Control	-.0199	.0391	.511
Availability for consultation	.1348	.0452	2.9834**
Demanding grading	-.0562	.0554	1.0137
Fairness of grading	.1053	.0468	2.2532*
Upper Division	-.0209	.0314	.6638
Median grade	-.0015	.0346	.0424
For	.0146	.0457	.319
Rank	.0298	.0355	.8383

Number of observations=147 R^2 = .9661 adjusted R^2 = .9619 F =231.7745
** Significant at the 99 percent level * Significant at 90 percent level + Significant at the 90 percent level

IV. CONCLUSION

Overall, my results suggest that the component items in our evaluation form do a good job of accounting for the variation in the summative "overall evaluation" item. The qualities of good instruction that matter most to the students are qualities that should also matter to the institution. Importantly, the results indicate that there is no need for instructors to attempt to improve their ratings by lowering the rigor of their courses or by relaxing their grading standards. The results indicate that the basic "blocking and tackling" of teaching, delivering clear lectures, responding effectively to student questions and being available to the students outside of class are the best way to ensure good evaluations. To paraphrase Tip O'Neil, it is probable that all evaluations are local. I am not certain that my results would apply to other institutions or even to other departments at my own College.

The data on which my estimates are based cannot be viewed as the outcome of a controlled experiment. Indeed, I don't want to have to conduct such an experiment on a regular basis. Many studies in the literature have access to information on individual students. I don't have access to individual student data nor do I want access to such data. Like most department chairs in my position, I have access to an imperfect data set. The question is whether I have extracted the maximum amount of meaning out of it. This paper represents a beginning but there is much work yet to be done.

REFERENCES

Becker, William E., and Michael Watts. 1999. "How Departments of Economics Evaluate Teaching." *The American Economic Review,* 89(2): 344-349.

Boex, L. F. Jameson. 2000. "Attributes of Effective Economics Instructors: An Analysis of Student Evaluation." *The Journal of Economic Education,* 31(3): 211-227.

Bosshardt, William, and Michael Watts. 2001. "Comparing Student and Instructor Evaluations of Teaching." *The Journal of Economic Education,* 32(1): 3-17.

Carrell, Scott E., and James E. West. 2010. "Does Professor Quality matter? Evidence from Random Assignment of Students to Professors." *Journal of Political Economy,* 118(3):409-432.

Decanio, Stephen J. 1986. "Student Evaluation of teaching: A Multinomial Logit Approach." *The Journal of Economic Education,* 17(3): 165-176.

Gramlich, Edward, and Glen A. Greenlee. 1993. "Measuring Teaching Performance." *The Journal of Economic Education,* 24(1): 3-13.

Gujarati, Damodar N. 2003. *Basic Econometrics*, 4th edition. New York, NY:McGraw-Hill.

Mehdizadeh, Mostafa. 1990. "Loglinear Models and Student Course Evaluations." *The Journal of Economic Education,* 21(1): 7-21.

Needham, Douglas. 1978. "Student Effort, Learning and Course Evaluations." *The Journal of Economic Education,* 10(1): 35-43.

Nelson, Jon P., and Kathleen A. Lynch. "Grade Inflation, Real Income, Simultaneity, and Teaching Evaluations." *The Journal of Economic Education,* 15(1): 21-37.

Siegfried,Jjohn J., and Rendig Fels. 1979. “Teaching College Economics: A Survey.” *Journal of Economic Literature*, 17(3): 923-969.

Zagaenzadeh, Hamid. 1988. “Grade Inflation: A Way Out.” *The Journal of Economic Education*, 19(3): 217-230.

Determining Optimal Firm and Consumer Research and Development Spending in the Medical Technology Sector

Sameer J. Shah*

ABSTRACT

Many studies indicate that increasing expenditure on the development of medical technology is the main impetus behind rising health care costs. This study examines the relationship between medical technology companies' goal to maximize profit via Research and Development investments and R and D's effects on the competing factors of consumer's health benefit and health care costs. Through analysis of pharmaceutical company profits versus R and D costs, this project attempts to map out a level of optimum R and D investment spending for firms. On the consumer side, this project also attempts to map out R and D investment levels by firms that yield the best health benefit for patients of drug treatments. Finally, the project attempts to determine if an R and D optimum exists that can increase benefit to both firms and consumers.

INTRODUCTION

Health economics research indicates that a major force behind the rising costs of health care in the United States is the cost of medical technology. This is a complicated matter as these costs represent profits for medical technology companies who then proceed to reinvest these earnings into research and development spending. Thus, medical technology costs offer simultaneous costs and benefits for the consumer.

The calculation of medical technology's influence on increasing health care costs has relied on indirect measures. For instance, Cutler (2007) measured costs due to disease-specific medical technology to determine this effect. Other works, including Newhouse (1992) and Okunade and Murthy (2002) considered a residual approach (measuring other costs and declaring the residuals to be due to medical technology costs) and a proxy approach (using sample technologies to measure costs) respectively. However, while no direct measure exists to measure medical technology's impact, all reliably indicate that it one of the largest causes for the rise in health care costs.

Both the consumer and the firm seek to optimize efficiency by maximizing research and development output while minimizing costs. However, due to the high volatility of research and development's success in developing drugs and thus profits for the firm, there appears to be no clear standard across the industry in choosing an efficiency-maximizing research and development spending schedule. This study seeks to determine if firms currently operate inefficiently and to determine if an optimal research and development value can be determined.

For firms, this value must optimize profits, not maximizing drug output, but rather overall profit for

* New York University, Department of Economics. New York, NY. First place winner of the 4th Annual NYSEA Student Paper Competition.

minimal research and development input costs. However for consumers, an optimal value will maximize health benefits due to these drugs, while minimizing drug costs. Thus, this study also seeks to unite two potential optimization points, for the firm and consumer, to examine if both sides can benefit from changes in research and development spending.

THE FIRM

On the firm side, companies choose to maximize profits and are willing to set the price as high as their market power will allow. In the case of the medical technology sector, highly inelastic demand for health care represents a strong ability for companies to set their prices. However, companies have an interest in diverting a large percentage of profits towards research and development investment as this will yield future profits in the company's future. Thus the companies' profit-maximizing condition considers two effects: it must maximize profits in the current period while making an investment decision that will maximize future profits.

For example, consider a typical firm that must choose to invest in research and development spending (as an input of production). Greater investment in research and development increases the output of drugs, but may not necessarily be the most efficient, profit-maximizing solution. In a case where the company is past its drug production optimization point, there are diminishing returns to each additional level of research and development spending and the firm is barely increasing the number of drugs produced.

Additionally, consider two firms who choose two different levels of research and development spending levels. The budget constraint is the level of revenue earned from the prior period. The firms can choose to invest more in research and development spending as Firm A has done, or spend less and keep more profits from the prior period as Firm B. These indifference curves are a function of the company's willingness to take on risk by investing more heavily in research and development spending, or by choosing a more conservative route where prior revenue is more heavily kept as revenue.

Firms take on huge amounts of risk by investing in research and development and must find a suitable level of spending to maximize their profits. Also, this high level of risk leads to high prices to cover the cost of potential failures taken on by firms. Although medical technology companies have a large amount of market power, they must maintain balance to not set prices so high that they lose customers.

As further explained in the "Data" section below, this study will use pharmaceutical companies as a proxy for all medical technology companies with data from annual reports. This study will examine costs accrued by companies due to both successful and failed research and development products (listed as aggregate research and development spending in the base (current period) year's annual report. The future success and failure of spending is unobservable in the present period and thus by examining aggregate spending, "survivorship bias" of measuring only successful projects is eliminated. The study will then model medical technology costs against profits in the future period that drugs used in the model

were sold.

Implementation of a two-period model will be necessary to gauge the effects of research and development spending in terms of input costs and output profits. Research and development is unique in that the largest costs occur up-front and the benefits are not realized until subsequent periods. Thus, costs need to be measured in the period before the drug's release and as it is being produced (operational costs) in the period after its release along with profits post-release. According to Keyhani, Diener-West, and Powe (2006), average drug development time, post acceptance of the Investigational New Drug application (IND) is 6.3 years. During this time, many of the overall research and development costs are spent in trying to develop the new drug. Thus, this study uses this length of time to calculate aggregate research and development spending costs for the development of new drugs in a 2 year period that immediately follows.

Profit-cost analysis will lead to a cost-efficiency measure to determine how successful (in terms of maximizing profits) various research and development strategies have worked by analyzing profits in the subsequent period. The overall purpose of this part of the study is to determine a potential optimal value of research and development spending by firms that is the most efficient compared to current investment decisions. In addition, methods for companies to expand profits will also be examined to determine if a more-efficient, optimal decision exists.

THE CONSUMER

The study will then proceed to examine the consumer's side of the equation. The product of research and development is medical technology. Consumers desire to maximize the benefit of medical technology while at the same time minimize its costs that add to the increased spending on aggregate health care. Thus, it is important to find methods to measure both the health benefit and the cost associated with increased expenditure on health care for each level of research and development spending from different firms.

The problem of health benefit can be calculated by approximating the cost of an extended year of life by using a similar approach to the work of Cutler and McClellan (2001). In this study, various health conditions were examined by comparing the technology used to treat the condition and the average health benefit yielded by using the technology. The authors used $100,000 as the value for each year extended by the technology, using a quality-adjusted life year (QALY) method (see "Quality-Adjusted Life Year" section for a more in-depth explanation). The study will attempt to measure not only the years of life extended by the treatment, but also the quality of life rendered by the treatment as well (using any available data that can be found from other research studies of medical treatments, such as clinical trial studies) and create a series of QALY values for various drugs associated with different firm research and development spending schedules.

The quality of life/years of life extended variable will thus provide the backbone of the potential health benefits offered to consumers. It may be assumed that the more spending on research and development

that occurs, the higher the value of welfare benefit for the consumer. However, this study would like to test whether this is the case and hypothesizes that at a certain level, additional spending on research and development has a minimal health benefit and in fact increases associated costs for the consumer. For example, on a graph drawn with increasing return approaching a plateau, the point of the beginning of the plateau results in minimal benefits of additional research and development spending for the firm; likewise a similar graph can be drawn where a similar point will have minimal health benefits for the consumer with increased research and development spending.

Using the calculated values of technologies' benefits and costs, the study will attempt to find the most efficient technological output for cost (or the most cost-efficient solution). There is a possibility that the most cost efficient solution occurs far below current research and development expenditure levels, indicating that this efficient solution fails to provide enough utility for consumers to make up for the benefits of decreased costs. Thus, current data must be analyzed to determine if a more cost-efficient solution is also feasible when taking into account consumer utility for health benefits (that is, consumers demand a high amount of health benefit and low values will not occur due to the necessity of medical technologies).

The final goal of this project is to compare the optimal values facing the firm and facing the consumer of research and development spending. Firms may currently operate in an inefficient manner, spending far more than is necessary to maximize profits in the next period. In addition, the consumer health benefit may actually occur at a lower amount of research and development spending than would be predicted. Thus, this study will attempt to find such opportunities for Pareto Improvements in terms of an optimized research and development spending value for both sides, benefiting both consumers and producers.

UNIQUE FEATURES OF THIS STUDY

Many studies have attempted to map out optimality on both the firm and consumer sides. However, this study will attempt to match both sides to determine if there is a spending level that will benefit both groups. In addition, previous studies have not done a very thorough job in examining additional "quality of life" factors that are a byproduct of medical technology's health benefit, which this study will attempt to solve using clinical trial studies.

On the consumer side, Cutler and McClellan's (2001) work examining health benefits due to medical technology provides many key methodologies for this project. The study extended Newhouse's (1992) study examining the technological costs and health benefits for heart attacks. Their findings concluded that the between the years 1984 and 1998, heart attack technology provided a net social benefit of $70,000 for consumers. To reach this value, they took the average cost of treatments associated with heart attack patients and subtracted this value from the number of years that this treatment added to a patients' life on average. The conversion for a year of life was taken to be $100,000 per life year extended. This value was calculated from a quality-adjusted life year (QALY) approach that will be taken as a fact for the purposes of this project.

QUALITY-ADJUSTED LIFE YEAR APPROACH

Sassi explains the calculation of the QALY in his 2006 work. This paper will provide the conversions necessary to express the conditions stated in clinical trial studies of specific prescription drugs caused by their use to numerical values for easier drug health benefit evaluation. Quality of life extended is measured on a scale of 0-1 with 1 measuring someone in perfect health who has added a year to his or her original life expectancy. Those whose lives are extended, but live with some side-effect, have a positive QALY that is less than 1. Generally, this is determined through surveys involving people rating their quality of life with a certain illness (generating the Disability-Adjusted Life Year). This value is then subtracted from 1 to form the QALY. However, without the ability to survey thousands of participants with various conditions, this paper will generate a novel scale applying fixed values for certain problems associated with prescription drugs. This is an effective and simple way of measuring health benefit added by the prescription drug (the benefit – the side effects or risks). As with Cutler and McClellan (2001), these values can then be converted into dollar figures.

The Australian Productivity Commission (2005) performed an in-depth study to determine the benefits of medical technology innovations in addition to Cutler and McClellan (2001). The Commission broke down different ways in which technology affects consumers, although the material is rather dense. This paper will provide a resource to compare to Cutler's methodology. In addition Cutler (2007) performed a detailed examination of costs and benefits due to medical technology over an entire lifetime for the consumer following a myocardial infarction. While these and other studies approximate both benefits and costs well, they fail to examine an optimal value of research and development spending that would maximize consumer health benefits while at the same time minimize costs.

Cutler and McClellan's (2001) framework will be utilized for this study's methodology to calculate net social benefits for the consumer. However, as mentioned above I must add conditions that will take into account quality of life and other factors (using the QALY approach) that are not well-examined by current studies, which only indicate only how many years of life a technology adds on average for a patient of a disease. I will however follow their basic design on taking a certain market basket of drugs/conditions and analyze both the cost and benefit associated with treatments for patients with this specific disease.

FIRM RESEARCH AND DEVELOPMENT SPENDING

Kort's (1998) model takes the uncertainty of the success or failure into account of the firm's decision to invest in research and development. In addition, the model utilizes a two period system that takes into account the initial investment in period one and the rewards gained in the subsequent period. Firms must maximize their profits in both periods not based on research and development spending for the future and current profits from prior works. While this equation is rather complex, it nicely derives a value of K*, or optimal research and development investment. However, it seems firms do not seem to operate strictly to maximize profits and are constantly increasing their research and development budget due to competition

and the desire to capture as much market share as possible with new drug treatments. Thus, they operate to constantly increase production of new drugs and rarely consider other R and D spending schedules.

Data regarding pharmaceutical firm behavior will be taken from company annual reports and other information available on the internet. Other works, including Carlton and Perloff (2005) and Chakravarty (1995) helped with the methodology regarding the measurement of the net social benefit value for consumers and the basic implementation of models in calculating the firm research and development optimal condition. All of these works have helped to develop this paper's goal to compare both sides of the spending decision to try and develop an optimal condition that satisfies both sides.

DATA SOURCES

The study uses pharmaceutical companies as a proxy for all medical technology companies. Pharmaceutical companies publish a wealth of information through their annual reports, such as aggregate research and development spending as well as revenue information. Prescription drug costs and drug clinical trial reports can also be easily accessed over the internet. In addition, pharmaceutical companies' research and development decisions are often times quite dynamic and provide a good indicator of the effects influencing firm decision making. Finally, limiting the scope of technologies to just pharmaceuticals allows for greater simplicity and flexibility in data manipulation.

FIRM RESEARCH AND DEVELOPMENT SPENDING REVISTED

On the firm side, the key to this project is to accumulate the necessary data regarding firm research and development spending to test the model on optimal firm R and D spending. According to Keyhani, Diener-West, and Powe (2006), the average drug development time (for clinical trials and regulatory review) was 6.3 years between 1992-2002. Thus, this study uses R and D spending values from 2000-2006 and then examines what drugs entered the market between 2007-2008. This provides the success rate (number of drugs entering market/costs during drug development period) for firms to examine potential optimums.

Drug development also appears to be cyclical, thus the study also examines the number of drugs that were accepted by the FDA between 2000-2006 as well. This will be compared to the drugs approved between 2007-2008 to determine how realistic this window is compared to a longer period.

MEASURING HEALTH BENEFIT

On the consumer side, the most difficult step will be to analyze additional variables characterizing health benefit that go beyond the standard years of life extended value. All of these variables, currently grouped under the title "quality of life" will likely be difficult to observe. However, this project uses prescribing information on drugs to determine the how beneficial treatments are for patients compared to the risks from taking these drugs. This data is manipulated into a numeric value using the QALY that

provides an easier analysis for the quality of life under this treatment option.

PAIRING CONSUMER NET SOCIAL BENEFIT TO FIRM RESEARCH AND DEVELOPMENT

Finally, the study takes both the consumer optimal research and development value in terms on health benefit and compares it with the optimal research and development spending value (in terms of profit maximization) for the firm . The drugs which have been examined for consumer's health benefits will be the same drugs that entered the market between 2007-2008, thus pairing the firm's research and development strategy to the net social benefit of these same drugs for the consumer. It will be evident what research and development strategy taken on by the firm yields the highest value of net social benefit to yield an optimal research and development value for the consumer. The study will then examine if it is possible for both sides to choose a different R and D spending level that will provide benefits for both, or a Pareto Improvement.

FIRM RESEARCH AND DEMAND OPTIMUM

Using the FDA New and Original Drug Approvals search engine, I searched for new drugs approved between 2006-2007. I limited the drug companies I was searching for to the top 10 companies using PharmExec.com's list of the top 10 companies by U.S. market share. The data was puzzling because Pfizer, the leading pharmaceutical company in the world has only produced 1 drug for the market within this 2-year period compared to 8 each with Novartis and GlaxoSmithKline. This seemed to indicate that companies developed drugs in cycles. Thus, I decided to examine drugs produced within a wider window, from 2000-2008. This data is more representative of the current leaders by U.S. market share, although Novartis stands out as a clear outlier, producing far more drugs within this period than its current position in U.S. sales indicate.

Since 2000-2008 drug approval data seems to be more accurate, representing the perceived cyclical nation of drug development, I decided to graph this data versus aggregate research and development spending to determine if diminishing marginal returns to research and development spending occurs and if a possible optimum value can be determined. The data from this chart includes all research and development costs accrued by the company from the years 2000-2006, which represents the 6.3 year estimated average period to create a successful drug entrant into the market.

The data indicates that increasing spending in research and development leads to increased drug ouput. I removed the outliers of Novartis and GlaxoSmithKline and found that the R-squared of the linear trend line was more representative than the logarithmic R-squared with values of .8178 and .6952 respectively. Thus, it does not appear as though diminishing returns to R and D spending occur. Companies should strive to increase R and D spending and the yearly data from 2000-2006 indicated that almost always, companies increased their R and D expenditure.

Even though the data from 2007-2008 of new drugs approved by the FDA is not an accurate representation of R and D output for firms, I decided to examine the cost-profit ratio from R and D spending by firms from 2000-2006 versus the profit of drugs created in 2007-2008 using profit data from

2009 to see if I could still find an optimum spending level for firms. I subtracted the profit of 2006 from the profits of 2009 to get a rough estimate of profit from drugs entering the market in 2007 and 2008. I also added costs of advertising and selling the drug, as increases in expenditure in this category has the potential to increase sales of drugs for firms. Note that Wyeth was acquired by Pfizer in 2009 and thus 2009 Wyeth data is not available. I then determined the cost/revenue ratio. I sorted the data to reveal who had the lowest cost/revenue ratio and compared it to the companies who spent the least on research and development. Unsurprisingly, the data is inconclusive, likely representing the cyclical nature of drug development and the misleading number of drugs developed by companies in the period 2007-2008.

While the data is inconclusive, it suggests that medium spending on R and D tends to yield the best results. Even when the problem of cyclical drug development is considered, Novartis seems to have the most-efficient drug production. Novartis produced the most drugs between 2007-2007 and 2000-2008 and had only the third-highest aggregate R and D spending amount. The two ends of R and D spending, Pfizer and Amgen, had the highest cost/revenue ratio, although this is likely a function of cyclical drug production for Pfizer. In general, higher expenditure on R and D yields higher drug ouput, but medium spending appears to be the best and optimal strategy.

CONSUMER RESEARCH AND DEVELOPMENT OPTIMUM

To determine optimal consumer research and development spending, I decided to limit my data set to drugs produced between 2007-2008 for the 4 firms at both ends of the R and D spending spectrum. The highest spender was Pfizer, and the lowest spender was Amgen. I also included Novartis, which was the most efficient spender using the 2007-2008 data and 3rd highest spender overall, and Wyeth, which was the second lowest spender on R and D.

I used each of the government-mandated “Prescribing Information” to weight their adverse reactions. I also characterized the drug’s environment (what other drugs treat the same condition) and then developed my own QALY-like scale of values.

Amgen seems to have the most beneficial drug for consumers with a value of .85. It is possible tat Amgen’s research and development spending is aimed for blockbuster drugs that could arise from treatment of a previously untreated condition. However, it is likely this sort of targeted research and development approach is unlikely as companies do not have tight grip on when and how treatments can be developed. Novartis has the second most beneficial drug series, but also developed more drugs than any of its competitors. Thus, Novartis’s research and development spending yields the best package for consumers. It is important to note again that this two year window does not adequately represent the makeup of drug production as it appears as though some companies go through a cyclical wave of high and low development throughout various years. Nevertheless, Novartis seems to have the best package for consumers.

CONCLUSION

While overall the results of this study are inconclusive due to the perceived cyclical nature of drug development, it appears as though increased spending on research and development generally yields in higher drug output. Drug output may be the best way to judge R and D because firms may not be able to steer their development into higher-earning opportunities, such as developing treatments for currently poorly-treated conditions. Diminishing returns to R and D spending were noticed, but this could be because companies have yet not reached the plateau of spending which could eventually result in less drug production for increased expenditure.

In all data sets, Novartis appears to have the best results in terms of drugs developed. The company's spending is also 3rd highest; while on the higher side, it also shows that companies do not need to spend exquisite amounts to produce many drugs (as in the case of Pfizer). This efficient production could be due to a hidden variable in Novartis' operational management that cannot be elucidated here.

On the consumer side, Novartis again appears to be the leader in producing the optimum research and development spending leading to maximum number of useful drugs approved. Pfizer's high spending only resulted in 1 average-benefit drug. The low spending Amgen did better, producing 1 high-benefit drug. Wyeth did better, spending on the low side and producing 3 moderately beneficial drugs. However Novartis produced the highest amount of drugs with the second highest average benefit for consumers while being only the third highest spender on research and development.

In general, it appears that moderately high expenditure on research and development yields the most optimal solution in terms of profit for firms and health benefit for consumers. As a whole higher spending seems to be better than lower spending as it increase profit and benefit for both firms and consumers as well. Thus, it is no surprise that firms constantly strive to increase their research and development spending levels yearly as it benefits not only them, but the consumer as well.

ENDNOTES

1. The following websites provided brief summaries classifying the prescription drugs used in this study: www.associatedcontent.com (Exforge), www.centerwatch.com (Tekturna), www.drugs.com (Toviaz), www.guide4living.com (Exelon), www.healthcentral.com (Reclast)
2. Pharmaceutical company websites provided the "Prescribing Information" used to classify drugs
3. The Federal Drug Administration's Investigational New Drug Application website (http://www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm?fuseaction=Reports.ReportsMenu) provided the data for drugs approved by year
4. The Security and Exchange Commision's EDGAR Databse (http://www.sec.gov/edgar.shtml) provided annual report information when annual reports were not available on pharmaceutical company's websites.
5. The list for the top 10 drug companies by market share was provided by www.pharmexec.com
6. Data and graphs available on request.

Acknowledgement

I would like to thank Professor Daniel Xu and Professor James Ramsey for their guidance and support.

REFERENCES

Bessen, James E. 2007. “Estimates of Patent Rent from Firm Market Value.” Boston University School of Law Working Paper No. 06-14.

Carlton, Dennis W. and Jeffrey M. Perloff. 2005. *Modern Industrial Organization*. Pearson Addison-Wesley: Boston, MA.

Chakravarty, Satya R. (1995): *Issues in Industrial Economics*. Avebury: Aldershot, England.

Congressional Budget Office. 2006. “Research and Development in the Pharmaceutical Industry.”

Cutler, David M. and Mark McClellan. 2001. “Is Technological Change In Medicine Worth It?” *Health Affairs,* Volume 20:11-29.

Cutler, David M. 2007. “The Lifetime Costs and Benefits of Medical Technology.” NBER Working Paper Series.

DiMasi, Joeseph A., Ronald Hansen, and Henry Grabowski. 2003. “The price of innovation: new estimates of drug development costs.” *Journal of Health Economics*, Volume 22:151-185.

Eli Lilly and Company. 2008. “2008 Annual Report and Proxy Statement.”

Getzen, Thomas E. 2007. “Modelling Long Term Healthcare Cost Trends.” The Society of Actuaries.

Kaiser Family Foundation. 2007a. “Trends in Health Care Costs and Spending.” Menlo Park, CA.

Kaiser Family Foundation. 2007b. “How Changes in Medical Technology Affect Health Care Costs.” <http://www.kff.org/insurance/snapshot/chcm030807oth.cfm>.

Keyhani, Salomeh, Marie Diener-West, and Neil Powe. 2006. “Are Development Times For Pharmaceuticals Increasing or Decreasing?” *Health Affairs*, Volume 25:461-468.

Kort, Peter M. 1998. “Optimal R & D Investments of the Firm.” *Or Spektrum* (Netherlands), Volume 20: 155-164.

Newhouse, Joeseph P. 1992. “Medical Care Costs: How Much Welfare Loss?” *Journal of Economic Perspectives*, Volume 6(3): 3-21.

Okunde, Albert A. and Vasudeva N.R. Murthy. 2002. “Technology as a “Major Driver” of Health Care Costs: a Cointegration Analysis of the Newhouse Conjecture.” *Journal of Health Economics,* Volume 21: 147-159.

Perloff, Jeffrey M. 2008: *Microeconomics: Theory & Applications with Calculus*. Pearson Education: Boston, MA.

PriceWaterhouseCoopers. 2006. “The Factors Fueling Rising Healthcare Costs.”

Productivity Commission of Australia. 2005. “Impacts of Advances in Medical Technology in Australia.” Research Report, Melbourne.

Ringel, Jeanne S., Susan Hosek, Ben Vollard, and Sergej Mahnovski. 2001. “The Elasticity of Demand for Health Care.” National Defense Research Institute (RAND).

Sassi, Franco. 2006. “Calculating QALYs, comparing QALY and DALY calculations.” Oxford University Press: London, UK.

The Far Reaching Effect of Ground Water Reductions

Anthony Signorelli*

ABSTRACT

As our greatest natural resource, water is a variable in future economic growth and the public health of our nation and the world. Policy making, related to global health and food production is influenced by changes in water quality and availability. The United States must understand the limitations of our own water resources. An increased awareness of the problems created by ground water reduction around the world can increase our diligence and attention to conservation policy.

INTRODUCTION

Natural disasters affecting billions of people around the world have to a large extent been related to floods and droughts and in developing countries this carries a death toll. Populations without adequate water supplies cannot sustain or recover from these natural disasters. There are 1.1 billion people, or 18% of the world's population, who lack access to safe drinking water; 2.6 billion or 42%, who lack access to basic sanitation. Water related diseases, diarrhea and malaria, rank 3rd and 4th in cause of death among children under 5 years old. At any given time, half of the world's hospital beds are occupied by patients suffering from a water-related disease (Water 2002, Water 2006). The water resources in many areas of the world are only marginal at best to maintain normal health even without added natural disasters.

There are many examples of ground water reduction and increased usage throughout the world. The health issues related to lack of fresh drinking water and water needed for agriculture will become a determining factor in survival of numerous populations which are currently teetering on the edge. As a world producer of technology, we must support world health. We must evaluate our own policy on many fronts to be available for programs which will help maintain and develop water supplies and support world health. One disturbing relation is the lack of grain production as a food source as we squander this in the quest for more fuel sources. Ethanol production utilizes a key food supply with very inefficient use of water and energy resources.

As a professor in the Health Care Management program at the State University of New York at Canton I have been teaching students the need for attacking health issues with insight into the future. We must bring world attention to the economic stress related to this growing problem, creating an understanding of how our water resources, even in the United States are relevant to world health. New initiative and organizations could include education within the United States along with programs that address the problem in countries with urgent need. The Royal Bank of Canada has acted by creating the

* Business Department, State University of New York at Canton, Canton NY 13617.

RBC Blue Water Project, a 10 year commitment to provide $50 million toward global fresh water initiatives to foster a culture of water stewardship.

The value of ground water will be determined by how industry, agriculture and domestic use control ownership. The domestic use of water has an impact on public health. The availability of water for agriculture and industry will also affect public health by stabilizing food supply and maintaining adequate energy supplies in a viable economy. The issues of maintaining industry and all of its far reaching effects on the economy and energy must now be addressed from the view point of public health.

On a global level the effects of ground water reduction on public health results in political strife and impacts the growth of nations. National leaders with policy planning toward water sustainability can fit into a world economy more favorably. One does not need to look to global conflict or natural disasters to find examples of political discord related to reduction in ground water. However it seems the public health issues in countries which are already struggling with inadequate water supply feel the most pain. In a prosperous industrialized nation the water infrastructure can buffer the effects of growing industry and energy needs on the availability of ground water.

As the reduction in ground water coincides with the exponential growth to our world population, the need to value our water supply with the hope of directing action to maintain sustainability of the resource becomes evident. The growth of industrial energy use and water requirements will continue to be exponential as is population growth. In a developing country, water requirements correspond to agriculture and domestic need. The growth of industry will force a reduction in water available for agriculture. The natural progression of domestic use will increase with the population. Thus, clean drinking water and adequate water for sanitation are two serious public health issues that need to remain in balance with population rise, industrial development and increase need for agriculture efficiency.

VALUE OF WATER

The United Nations report on industrial development estimates the rise in water use by industry from 1995-2025 will be 35%. An expected world population growth of 6 billion to 9 billion will clearly stress water availability for industry. The high income countries now using 59% of their resources for industry will also need to be leaders in water management. The United States will need to evaluate where population growth is stressing water supply and how industry is using and polluting water (UNSummit 2002).

A discussion of world energy policy and the quest for energy independence in the United States is an extension of water as a sustainable resource. The correlation between hydropower and electricity, water and energy will also control ground water depletion and overall public health. As we link public health with industry, agriculture and energy, with an understanding of water availability, the value of water becomes more evident.

Urban areas across the United States, those in areas of minimal rainfall and those in the Great Lakes region have been studied along with the 20 largest cities to evaluate residential water use and cost. It is interesting to note, in Boston, Massachusetts, the average daily use per person is 41 gallons at a cost of

$67/100 gallons and in Fresno, California, where the cost is $34/ 100 gallons the use is over 200 gallons per person per day. In another example, Milwaukee, with an abundant supply of lake water and easy access for its population; prices are on the rise as usage drops. Breweries and tanneries have left town causing a 41% drop in water revenue from 1976 to 2008 (Walton,B 2010).

Buying water on the open market and desalination are costly alternatives for many southwestern US cities. This has promoted many conservation minded activities with good results. Since 1995, Santa Fe, New Mexico, water use has dropped 42% and Phoenix has not increased domestic use over 10 years even though they have an additional 400,000 residents.

What is the most important determining factor to value water correctly? Is it availability, cost or even more important do we drive the price up if we conserve?

In many arid western cities federal investment to water infrastructure has reduced the capitol cost to cities and states. These subsidies create an artificial price. Considering the California Central Valley Project, Central Arizona Project, Hoover Dam, Colorado's Big Thompson Project, federal funding of 3.6 billion with only 45 percent funded by residents has off set the real cost of water. In Santa Fe, New Mexico where ground water use is unsustainable, the city is building a $217 million water diversion project without federal money. City and county taxes will fund 75% and 25% will be grant funded (Gleick 2010).

In numerous cases around the country infrastructure was developed 50 years ago. In 2009 the EPA has estimated $335 billion will be needed to fix the countries aging water supply. Will these funds be generated by tax or tariff? Will water supply infrastructure to areas of ground water reduction even be realistic? Ground water sustainability seems to be a more important question than infrastructure cost.

In other countries the ground water reduction is so severe that bulk water transport is planned. Tankers holding 80 million gallons will take water from the Blue Lake Reservoir of Sitka, Alaska to India in 2010. Additional plans to transport to East China and the Caribbean will follow. Water from Sitka can unload at $0.07 per gallon. Australia is a penny per a gallon for desalination (10). Keep in mind that shipments can be turned on and off avoiding the capitol burden as multimillion dollar desalination plant remains dormant during a wet season. Every year the allocation water rights of farmers and cities in Australia can be resold. In 2000 one million liters traded for AU$2. In 2010 the same volume sold for AU$1,300 to 2,400. Even with market value drops of 40% in 2009, AU$3 billion in water rights was traded (Walton,Brett 2010).

It has been estimated that Quebec, Canada could see $6.5 billion annually through water export. If the resource is used with sustainability in mind it could enhance the economy and encourage conservation at the same time (Katz 2010,Olson 2010). Does the government or private interest have a right to privatize a public resource?

Since the 1800's through a US Supreme Court decision the Great Lakes and all navigable waters are subject to public ownership in a public trust. It demands water remain available for all citizens. However, a

recent Michigan law is permitting private use to deplete up to 25% of lakes and streams, ultimately means ground water depletion (Olson 2010).

COLLISION BETWEEN ENERGY DEMAND AND WATER SUPPLY

The next valuation of water occurs with a relationship of ground water conservation and energy production. In 1973 the government began talks of energy independence and over more than 40 years has neglected the significance of a water energy link. Each process has an environmental impact along with inequitable water use.

- *Hydro fracturing for Natural Gas:* The Pennsylvania – New York Marcellus Shale formation is 95,000 square miles of dense shale, which could provide enough natural gas to supply the entire east coast for half a century; a $2 trillion per year business. The projection for 2011 is 109 million gallons of waste water per day. Seventeen million people from Pennsylvania and New York acquire their drinking water from here, including a 2,000 square mile water shed supplying pure unfiltered water to New York City. There are 400,000 hydro fractured gas wells in the United States; each has created 4 million gallons of polluted waste water (over 1.5 billion gallons of polluted water) (AP Report 2010, Fox 2010, Rousseau 2010).
- *Ethanol production:* It takes more energy to produce ethanol than we obtain from the same unit of ethanol. It takes 1000 gallons of water to produce 1 gallon of ethanol. Annually 40 million bushels of corn are processed into 115 million gallons of ethanol in each plant (8). There are hundreds of these plants in the United States which waste water and energy, not to mention depletion of an exportable food supply.
- *Oil producers:* Oil producers injected 1.3 billion barrels of water into the ground to obtain 162 million barrels of oil. In August 2010 amidst severe drought the oil industry received 8.4 billion gallons of water (Schneider 2010).
- *Chip manufacturing:* A 2 gram, 32 megabyte memory chip needs 10 gallons of water and 80 pounds of chemicals to etch the chip. Intel employs 10,000 people in Chandler, Arizona investing 9 billion dollars into the economy. Chip manufacturing uses 2 million gallons of water per day, in a region of water scarcity.
- *Thermo electric power:* Coal fired and nuclear powered plants turn 3 billion gallons of water per day into steam. Thermo electric power generation accounts for half of the 400 billion gallons of water withdrawn daily from our nation’s rivers and lakes. The Energy Information Administration of the Department of Energy forecast a nearly 50% increase in electricity demand from 2005 to 2030.

CONCLUSION

The value of water must be measured against securing world food supply, protecting the environment, balancing withdrawal with recharge, while meeting energy needs with future growth in mind.

> "Water sits at the nexus of so many global challenges, including health, hunger and economic growth. And sadly, water scarcity takes its greatest toll on society's least fortunate. I am absolutely convinced that the only way to measurably and sustainably improve this dire situation is through broad-scale collaborative efforts between governments, industry, academia, and other stakeholders around the world" Indra Nooyi Chairman & CEO, PepsiCo, Inc.

The water debate has been about utilizing water in different ways. The debate now needs to consider risks related to water usage. Forward thinking countries are now importing water intense products rather than developing domestic production. The purchase of water rich farm land in Africa is becoming common place for countries to improve their food supply. It seems the most water intensive crops have been produced in countries with water scarcity.

Continued political achievements to promote ground water sustainability could (Katz 2010):

- improve public understanding of water issues
- conduct ground water and fresh water inventories
- determine sustainable water levels; determining use and recharge rates
- reform public subsidies for water use, artificially low residential and industrial water rates should be phased out and agricultural use must adopt conserving technologies
- repeal bulk water export prohibitions
- build future thermo electric plants in areas where cooling water can utilize waste water, sea water, and mining/ drilling water

At the same time we need to make critically needed observations and plan to deal with sanitation systems and their costly infrastructure upgrades.

REFERENCES

AP Report. 2010. "EPA Announces study to Re-Examine the Risk of Hydrofracking." www.circleofblue.org, July 29, 2010.

Fox, Josh. 2010. "Toxification of U.S. Water Sources in the Dig for Natural Gas." HBO Documentary "Gasland."

Gleick, Peter. 2010. "California's Next One Million Acre-Feet of Water." www.unwater.org/statistics/html.

Katz, Diane. 2010. *Making Waves: Examining the Case for Sustainable Water Exports from Canada.* Fraser Institute.

Olson, James. 2008. "Great Lakes Compact, Water Privatization." Circle of Blue/2010 Report. www.circleofblue.org.

Rousseau, Heather. 2010. “Where energy development puts rivers at risk.” Circle of Blue. www.circleofblue.org.

Schneider, Keith. 2010. “Energy Department Blocks Disclosure of Road Map to Relieve Critical U.S. Energy-Water Choke Points.” www.circleofblue.org.

United Nations Summit on Sustainable Development. 2002. World Water Assessment and World Energy Council, www.unesco.org/water

Walton, B. 2010. “The Price of Water: A Comparison of Water Rates.” www.circleof blue.org.

Walton, Brett. 2010. “S2CGlobal Report on Water Export.” Circle of Blue.

United Nations. 2002. “Water: A Matter of Life and Death.” www.un.org/water/factsheet.pdf.

Water Partners International. 2006. “Water Facts, 2006.” www.waterpartners.aspx.

Bankruptcy Prediction Using Neuro Fuzzy: An Application in Turkish Banks

Birol Yildiz[*] and Soner Akkoc[**]

ABSTRACT

With the global crisis, bankruptcies have increased and bankruptcy prediction models have become more important. The purpose of this study is to actualize the prediction of Turkish bank bankruptcies with Neuro Fuzzy (NF). NF does not have the problems, which come from the assumptions of statistical methods, and, as in Artificial Neural Network (ANN), it has learning ability. At the same time, the model does not stay in a black box. The proposed NF bank bankruptcy prediction model's performance was compared with Multivariate Discriminant Analysis (MDA) and ANN. Besides getting the best prediction of accuracy from Neuro Fuzzy, in this study, the addition of the forerunner indicators on the decision making process can also be interpreted.

I. INTRODUCTION

Monitoring and controlling the banking sector is important for maintaining confidence in a financial system. As a result, the banking sector is one of the most tightly regulated sectors in modern economies. This is especially important in transition economies, as a healthy banking sector is a prerequisite for increasing private savings and allocating loans to their most productive use (Lanine and Vennet, 2006). The banking sector is especially important for Turkey because banks constitute a major part of the finance sector, and companies such as asset investment funds, leasing companies, insurance companies, and factoring companies are generally subsidiaries of banks. The first major bankruptcies in Turkey appeared during the 1994 crisis. After that, as a result of reflecting upon the Asia crisis, which happened in 1997 and 1998, bankruptcies came into question. The biggest bankruptcies to date in Turkey occurred in 2000 and 2001. Bank bankruptcies have a greater effect on an economy than bankruptcies of any other type of company. One reason for this is because bank bankruptcies affect more than just the banks' shareholders: borrowers, depositors and other institutions that lend funds to the banks are also negatively affected. According to the bank size, the economy of the country can also be affected negatively in this case. Bank bankruptcies sometimes can prompt other bank bankruptcies successively.

When the studies on bankruptcy predictions are examined, firstly, it is seen that statistical models have been used in this area. However, the assumptions within the statistical models reveal some

[*] Assistant Professor, Business Administration Department, FEAS, Eskisehir Osmangazi University, Eskisehir, Turkey E-mail: birol.yildiz@gmail.com

[**] Visiting Scholar, School of Business, State University of New York - Oswego
Assistant Professor, School of Applied Sciences, Department of Banking and Finance, Dumlupinar University, Kutahya, Turkey E-mail: akkocsoner@hotmail.com

objections about the subject of generalizing the success of these models. ANN has been used since 1990s. But there is an important disadvantage of ANN. The coefficients regarding the ANN model cannot be interpreted. So, it cannot be known how the independent variables are used in the model.

Neuro Fuzzy (NF), which is one of the Artificial Intelligence (AI) technologies, is a hybrid technology obtained by using ANN and Fuzzy Logic (FL) simultaneously. NF has learning ability. Moreover the most important feature that separates NF models from ANN is the built model that does not stay in a black box. The purpose of this study is to investigate the bank bankruptcy prediction ability of the NF model by using the Turkish banks' data. The performance of the NF model is also compared with ANN and MDA. The remainder of the paper is organized as follows. Section 2 reviews the literature. Section 3 gives a brief outline of NF in building bank bankruptcy prediction model. Section 4 presents the research design and methodology. Section 5 presents the empirical results of the bank bankruptcy prediction models. In the last section concluding remarks are given.

II. PREVIOUS RESEARCH

Bankruptcy prediction studies go back to 1960s. Beaver (1966) pointed out that early warning signals could be taken from the financial ratios. Altman (1968) used MDA successfully in predicting bankruptcies for the first time. In predicting bankruptcies, MDA was used by Sinkey (1975), Multivariate Regression Analysis (MRA) was used by Meyer and Pifer (1970), Logistic Regression Analysis (LRA) was used by Martin (1977) successfully. West (1985) pointed out that using Factor Analysis and LRA together in evaluating bankruptcies produced hopeful results.

ANN models have been used for predicting firm bankruptcies since 1990s. In the studies, results showed that ANN models are more successful than the statistical models such as MDA, MRA, LRA and Probit Analysis (PA) (Davalos et al, 1999; Han and Lee 1997; Lee et al, 2005; Leshno and Spector 1996; Salchenberger et al, 1992; Sharda and Wilson 1993; Tan and Dihardjo 2001; Tsukuda and Baba 1994; Wilson and Sharda 1994; Yang et al, 1999; Zhang et al, 1999). ANN models have been also used in bank bankruptcy predictions. Tam (1991) and Tam and Kiang (1992) pointed out that in predicting Texas bank bankruptcies ANN was more successful than the statistical techniques. While Bell (1997) pointed out that ANN is more successful than LRA, Swicegood and Clark (2001) pointed out that ANN is more successful than MDA in prediction of bank bankruptcies. Alam et al, (2000) pointed out that in predicting bankruptcies, Fuzzy Clustering and Self-Organizing Maps would be able to be done successfully. Ravi and Pramodh (2008) have actualized prediction of the bankruptcies with the principal component of the ANN. Data Envelopment Analysis on bank bankruptcy prediction has been used successfully in the studies which were done by Cielen et al, (2004) and Kao and Liu (2004). Kolari et al, (2002) and Lanine and Vennet (2006) have found that the Trait Recognition model is more successful than LRA in bankruptcy predictions. With the significant nine variables, the prediction of bank bankruptcies was actualized in a 21-year-process in the USA between 1980-2000 by using the Generic Self-Organizing Fuzzy Neural Network (GenSoFNN) model by Tung et al (2004). According to the findings GenSoFNN

has been found more successful than Cox's Proportional Hazards model on a total of 3,635 banks. However, Ng and Jiang (2008) found that Fuzzy Cerebellar Model Articulation Controller (FCMAC) model more successful than GenSoFNN and Cox's proportional hazards model with the same data set. Nguyen and Quek (2008) obtained prediction accuracy up to 95% by using the Ying-Yang FCMAC model. With wavelet ANN models, Chauhan et al, (2009) have obtained, 100%, 89.99%, 93.33% prediction accuracy on the data sets of Turkey, Spain and USA respectively.

Also in Turkey, some researchers have built various bankruptcy prediction models, in which bank bankruptcies that happened in the period between 1994 and 2003. Canbas et al (2005) got 90%, 87.5% and 87.5% prediction performance from MDA, LRA and PA, respectively. Doganay et al (2006) obtained prediction success which changed between 81% and 95% by using MRA, MDA, LRA and PA. Karacabey (2007) obtained 92.3% prediction accuracy from mathematical based MDA. In the study done by Boyacioglu et al (2009), the best bank bankruptcy prediction performance was obtained from Multilayer Perception and Learning Vector Quantization. Çelikyılmaz et al (2009) got 98%, 94% and 97% prediction performance from the Improved Fuzzy Classifier Functions model.

III. NEURO FUZZY

Zadeh (1965) introduced FL, a mathematical system which deals with modeling imprecise information in the form of linguistic terms. The point where FL emerges becomes the crisis about the classic set theory. One member definitely belongs to the set or does not belong to the set in the classic set theory. FL makes possible that one individual can be a member of more than one set in a certain degree by means of membership functions. A is defined as a fuzzy set below.

$$A = \{(x, \mu A(x)) | \; x \in X\}, \qquad (1)$$

This equation $\mu A(x)$ shows membership function which gets value between 0 and 1, x shows a member of A set. FL represents models using if-then rules. For example;

If *Liquidity* is **average** and *Capital Ratio* is **high,** then *Bankruptcy Risk* is **low,** where *Liquidity, Capital Ratio* and *Bankruptcy Risk* are linguistic variables; **low, average** and **high** are linguistic values that are identified by membership functions. Each AI technology has a unique ability. ANN carries out machine learning by stimulating the human being's neural system. However, FL is very similar to a human being's reasoning. But these technologies have some unique disadvantages. The information that stays in the black box is an important disadvantage for ANN. An important disadvantage of FL is not having the ability to learn. In parallel to AI technology development, combinations of these technologies have come into question. To take advantage of the learning capability of ANN and the modeling superiority of FL, these technologies are used simultaneously and that is called NF. NF systems have the ability to apply human experience on the problem area by means of fuzzy rules (Jang et al., 1997). Learning capability by using verbal variables is the most important advantage for NF models, compared to other non-linear AI technologies (Abonyi, 2003). Another advantage of NF is being able to make a comment on how the

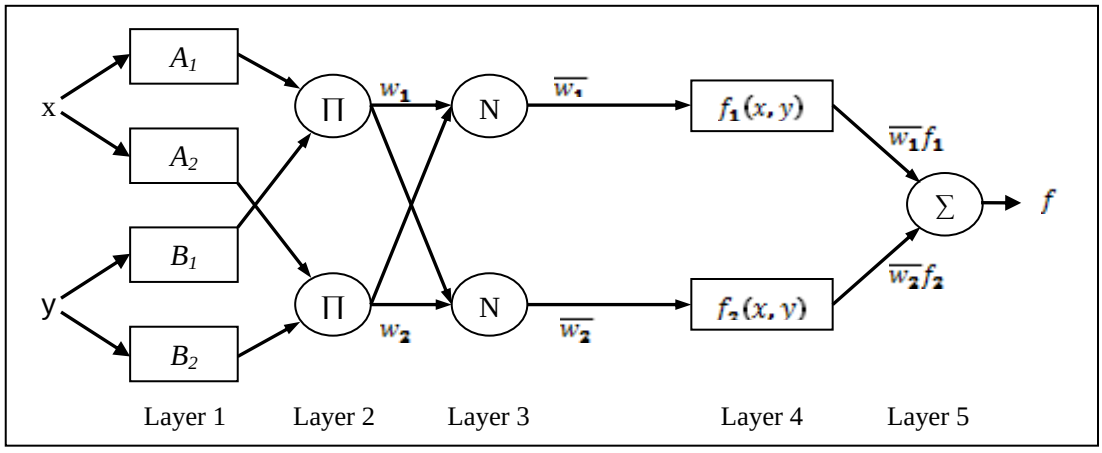
model produced the output value. NF has been applied to few researches for financial prediction (Akkoc, 2007; Chen et al., 2009; Malhotra and Malhotra, 2002; Piramuthu, 1999; Yildiz and Akkoc, 2009).

Adaptive Neuro Fuzzy Inference System (ANFIS), which has been developed by Jang (1993) and used in the application part of this study, is a kind of NF systems. ANFIS utilizes human expertise in the form of fuzzy if-then rules. ANFIS has the ability to construct models only with target sample data and exhibits fault tolerance. ANFIS determines itself appropriate parameters to provide the best learning by describing membership functions (Jang et al., 1997). The most important feature that separates ANFIS from ANN is the obtained model that can be presented with rules like "If…then…" To give two fuzzy if-then rules example, for a first order Sugeno model the two rules will be below:

Rule-1: If $x = A_1$ and $y = B_1$ then $f_1 = p_1 x + q_1 y + r_1$

Rule-2: If $x = A_2$ and $y = B_2$ then $f_2 = p_2 x + q_2 y + r_2$ (2)

where x and y are independent variables, A_i and B_i are fuzzy sets (linguistic labels like; low and high), p_i, q_i, r_i are the parameters of dependent variable. The corresponding equivalent ANFIS architecture which consists of five layers is shown in Figure 1.



Layer 1: Every node i in this layer, is an adaptive node with a node function described by;

$O_{1,i} = \mu A_i(x)$, for i=1,2 and $O_{1,i-2} = \mu B_i(y)$, for i=3,4, (3)

where x is the input node i, A_i and B_i are the linguistic label (low, medium, high, etc.) associated with this node function $O_{1,i}$ and $O_{1,i-2}$ are the membership function of A_i and B_i respectively. This research utilized $\mu A_i(x)$ and $\mu B_i(y)$ to be bell-shaped membership function with maximum and minimum equal to 1 and 0 respectively, such as: $\mu A_i(x) = 1 / 1 + [(x-c_i / a_i)^2]^{bi}$, where a_i, b_i and c_i are the premise parameters of the membership function.

Layer 2: Every node in this layer is a fixed node labeled Π which multiplies the incoming signals and sends the product out. The outputs of this layer which represents firing strength of the rules can be represents as: $O_{2,i} = w_i = \mu A_i(x) \times \mu_{Bi}(y)$, i = 1,2. (4)

Layer 3: Every node in this layer is a fixed node. Every node is labeled N. The ith node calculates the ratio of the ith rules firing strength to the sum of all rule's firing strength. In other words, this layer normalizes firing strength of the node i.

$$O_{3,i} = \overline{w}_i = \frac{w_i}{w_1 + w_2}, i = 1,2. \quad (5)$$

Layer 4: This layer calculates the consequent. Every node in this layer is an adaptive node, with a node function, where $\overline{w}_i$ is the output of layer 3 and p_i, q_i and r_i are the parameter set.

$$o_{4,i} = \overline{w}_i f_i = \overline{w}(p_i x + q_i y + r_i), \quad (6)$$

Layer 5: The single node in this layer is a fixed node labeled Σ that computes the overall output as the summation of all incoming signals.

$$O_{5,i} = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (7)$$

IV. RESEARCH DESİGN AND METHODOLOGY

Data

In the sample of the study, there are 55 banks and 19 of these banks were bankrupt on different dates. As the bankruptcy prediction was actualized one year before, the data set has been constituted by taking into consideration the financial ratios of the previous year. Because 2001 was the year of crisis, in which most of the bankruptcies occurred, the financial ratios of the non-bankrupt banks in 2000, have taken place on the data set, too. The data set, constituted of 55 banks, has been divided into two groups as training and validation. The training set consisted of 11 bankrupt banks and 22 non-bankrupt banks, whereas the validation set consisted of 8 bankrupt banks and 14 non-bankrupt banks. The data have been taken from the Bank Associations of Turkey (BAT) (http://www.tbb.org.tr). Table 1 shows the descriptive statistics regarding financial ratios. On the data set, there are 36 financial ratios under 6 titles of Capital Ratios, Assets Quality, Liquidity, Profitability, Income-Expenditure Structure and Activity Ratios. As a result of the “t test” on bankrupt and non-bankrupt banks, the differences among 24 financial ratios have been found significant at the 5% level and stated in italics.

Selection of variables and factor analysis

In this study, some data reduction techniques have been used to increase prediction performance. The ratios which do not show a significant difference are taken out of the first data set and a second data set is generated. The second data set has 24 financial ratios. The third data set is taken out by applying the Factor Analysis method on the second data set. Firstly, factor analysis is taken out after the financial ratios, which affect KMO negatively, are extracted from the data set, and reliability analyses are done. As an extraction method, the Principal Component Analysis is applied to the data set. KMO value of 0.63, given at Table 2, shows that data set is suitable for factor analysis. As a result of the Factor Analysis, there are 4 factors which are above the Eigenvalue 1. All these 4 factors explain 78.7% of total variance. The first factor explains 45.7% of total variance. This factor includes P1, P4, C2, C4, C5, IE4 and C3

ratios. So, it can be said that this factor represents profit and capital ratios. The second factor explains 16% of total variance. This factor includes L1, A3, A2 and L2 ratios. This factor is a mixed factor which is constituted with liquidity and activity ratios. The third factor explains 8.9% of total variance. This factor includes IE11, IE2 and C1 ratios. It can be said that this factor reflects the form of income and expense ratios. The fourth factor explains 8% of total variance. The fourth factor includes only IE3. The fourth data set is generated with the variables that are used in discriminant function. ANN and MDA are applied on all data sets, but NF is applied on the third and the fourth data sets because it is capable of building models with a few variables. The best classification results, obtained from the models we used, are presented.

Table 1: Statistical Information about Financial Ratios

Financial Ratios	Non-Bankrupt		Bankrupt		Test Values	
1. Capital Ratios	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**t**	**P**
(C1) Standard Capital Ratio	*34*	*38.4*	*5.84*	*11.09*	*-3.116*	*0.003*
(C2) (Shareholders' Equity+T.Income)/Total Assets	*16.25*	*10.01*	*-1.94*	*30.72*	*-3.263*	*0.002*
(C3) (Shareholders' Equity+T.Income)/(Deposits+Non-depodit Funds)	*28.96*	*27.9*	*2.81*	*18.5*	*-3.674*	*0.001*
(C4) Net Working Capital/Total Assets	*10.53*	*9.49*	*-12*	*36.79*	*-3.487*	*0.001*
(C5) (Shareholders' Equity+ T.Income)/(T.Assets+Contin.and Com.)	*7.53*	*7*	*0.17*	*10.25*	*-3.149*	*0.003*
(C6) Fx Position/Shareholders' Equity	226.93	270.39	205.21	293.47	-0.275	0.784
2. Assets Quality						
(AQ1) Total Loans/Total Assets	*24.43*	*15.72*	*33.54*	*12.24*	*2.197*	*0.032*
(AQ2) Non Performing Loans/Total Loans	*4.55*	*6.59*	*42.24*	*102.97*	*2.206*	*0.032*
(AQ3) Permanent Assets/Total Assets	13.31	13.42	15.5	14.35	0.561	0.577
(AQ4) Fx Assets/Fx Liabilities	70.6	24.75	62.97	22.77	-1.117	0.269
3. Liquidity						
(L1) Liquid Assets/Total Assets	*55.53*	*22.97*	*32.7*	*16.25*	*-3.846*	*0.000*
(L2) Liquity Assets/(Deposits + Non-deposit Funds)	*96.46*	*91.85*	*37.36*	*19.32*	*-2.761*	*0.008*
(L3) Fx Liquid Assets/Fx Liabilities	45.68	26.48	33	16.23	-1.903	0.062
4. Profitability						
(P1) Net Income(Loss)/Average T.Assets	*3.32*	*3.24*	*-11.44*	*29.68*	*-2.977*	*0.004*
(P2) Net Income(Loss)/Average Equity	42	46.82	-221.67	932.43	-1.707	0.094
(P3) Net Income(Loss)/Average Share-in Capital	*74.7*	*118.15*	*-216.81*	*698.46*	*-2.458*	*0.017*
(P4) Income Before Tax / Average Total Assets	*5.18*	*5.21*	*-10.9*	*30.02*	*-3.151*	*0.003*
(P5) Provision for Loan Losses/Total Loans	*1.52*	*2.58*	*21.17*	*58.43*	*2.031*	*0.047*
(P6) Provision for Loan Losses/Total Assets	0.36	0.6	5.35	15.3	1.972	0.054
5. Income-Expenditure Structure						
(IE1) Net Interest Income After Provision/Average T. Assets	12.28	7.88	5.43	21.29	-1.729	0.090
(IE2) Interest Income/Interest Expenses	*221.42*	*81.6*	*146.23*	*43.2*	*-3.738*	*0.000*
(IE3) Non-Interest Income/Non-Interest Expenses	*16.18*	*79.9*	*-42.61*	*87.62*	*-2.51*	*0.015*
(IE4) Total Income/Total Expenditure	*130.09*	*29.11*	*92.3*	*31.82*	*-4.434*	*0.000*
(IE5) Interest Income/Average Profitable Assets	*33.72*	*12.7*	*62.37*	*36.78*	*4.247*	*0.000*
(IE6) Interest Expenses/Average Non-Profitable Assets	*18.46*	*10.88*	*30.34*	*13.89*	*3.496*	*0.001*
(IE7) Interest Expenses/Average Profitable Assets	*16.74*	*7.67*	*45.7*	*30.85*	*5.368*	*0.000*
(IE8) Interest Income/Total Income	96.22	23.72	32.07	338.02	-1.143	0.258
(IE9) Non-Interest Income/Total Income	3.78	23.72	67.93	338.02	1.143	0.258
(IE10) Interest Expenses/Total Expenses	*59.29*	*13.87*	*69.56*	*11.22*	*2.779*	*0.008*
(IE11) Non-Interest Expenses/Total Expenses	*40.71*	*13.87*	*30.44*	*11.22*	*-2.779*	*0.008*
6. Activity Ratios						
(A1) (Salaries and Emp'ee Benefits+Reserve for Retirement)/T.Assets	2.73	1.83	2.97	1.42	0.494	0.623
(A2) (Salary and Emp'ee Bene.+Res. for Retire.)/No.of Pers.(Billion TL)	*25.63*	*15.64*	*10.61*	*5.88*	*-4.023*	*0.000*
(A3) Reserve for Seniority Pay/No.of Personnel (Billion TL)	*0.62*	*0.58*	*0.21*	*0.22*	*-2.915*	*0.005*
(A4) Operational Expenses/Total Assets	3.5	2.53	3.93	1.79	0.668	0.507
(A5) Provisions except Provisons for Income Tax/Total Income	*1.84*	*1.64*	*6.39*	*13.33*	*2.038*	*0.047*
(A6) Provisions including Provisons for Income Tax/Total Income	5.13	5.6	7.34	13.11	0.876	0.385

Table 2. KMO and Bartlett's Test

KMO		0.63
Bartlett's Test	Approx. Chi-Square	1135.949
	Df	105
	Sig.	0.000

The multivariate discriminant analysis

MDA, one of the frequently used statistical techniques on bankruptcy prediction studies, puts forward whether there is an explicit difference or not among two or more groups depending on a group of variables. The model that MDA has is like stated below.

$$Z_i = B_0 + B_1 X_{i1} + B_2 X_{i2} + \ldots\ldots\ldots + B_m X_{im,}$$

where Z_i is a discriminant score, B_0 is the intercept term, B_m are the estimated coefficient and X_{im} are the independent variables. MDA has been actualized on the training data and the prediction performance of the model has been evaluated by applying the obtained function on the validation data.

Artificial neural network model

This study uses the feedforward multilayer perceptrons with the BPN training algorithm. In this study, forming optimal ANN model comes with the development of many unsuccessful models in the process of trial and error. It has been observed that some of the developed unsuccessful ANN models cannot develop learning and cannot minimize error. This trial and error process has been kept on until finding an optimum architecture. Finally, it has reached an optimum architecture when it learns the training data and provides high prediction success on the validation data. The optimum architecture's learning rate, momentum, hidden nodes and training epochs are 0.4, 0.6, 6, and 2,650, respectively.

Neuro fuzzy model

This process starts with entering variables to ANFIS and assigning membership functions. Hereby, ANFIS has determined each parameter about membership functions by considering the inputs and the outputs. ANFIS has set up rules in order to make the best learning by considering independent variables and membership functions. When the training of the network is completed in the frame of developed rules, decisions on the banks are made. Table 3 shows the parameter of the optimum NF model.

Table 3. Parameters about developed NF

The number of the processing elements in the input layer	3
Membership Function Number	C4 2; IE 2; A5 3
MF Type	Gaussmf
Epochs Number	3,322
Rule Number	12

V. EMPIRICAL RESULTS

The empirical results of the models

Whether or not the prediction performance of MDA is found as successful can be understood from Canonical Correlation, Eigen value and Wilk's Lambda values given in Table 4. Because the value of Eigen value is higher than 0.40, the discrimination power of the model can be considered as pretty good. Canonical Correlation (0.798) measures the discriminant score and the relationship among the groups, and the square of this value shows the explained total variance. As for Wilk's Lambda statistic, it shows

the part of the total variance that is not explained by the discrimination among the groups. Discrimination functions developed on training data will be used in predicting financial situations of the banks on the validation data. Developed discrimination function is as stated below.

$$Z36 = 1.303 + 0.038 * C4 – 0.097 * IE7 + 0.232 * A5$$

Table 4. Statistics relating to MDA

Function	Eigen value	Can. Cor.	Wilks' Lambda	Sig.
36 financial ratios*	1.751	0.798	0.364	0.000

* Same results have been obtained on the second data set which has 24 financial ratios.

Prediction performance of the developed models on training and validation data sets is given in Table 5. MDA has predicted correctly 29 of 33 banks on training set and 18 of 22 banks on validation set. ANN has predicted correctly 32 of 33 banks on training set and 19 of 22 banks on validation set. NF has predicted correctly 30 of 33 banks on training set and 20 of 22 banks on validation set. It can be said that the high prediction success has been provided from each methods. But it is seen that the best result has been obtained with the NF model, when the validation set is taken into consideration.

Table 5. The Empirical Results of Training and Validation Data Set

Prediction Model	Variables	Training performance %	Validation performance %
MDA	36 financial ratios	87.9	81.82
ANN	3 financial ratios	96.97	86.36
NF	3 financial ratios	90.91	90.91

The 3D graphics of NF model

The most important feature that separates NF from ANN is the built model that does not stay in a black box. Non-Linear functions, which NF produces, take place in Figure 2a, 2b, and 2c. Non-Linear functions reflect just two dimensions of mixed relationships. So, significant relations should not be expected at each point of the functions. In Figure 2a, functions, which are developed relating to ratios of C4 and IE7, take place. Figure 2a indicates that the probability of going bankrupt increases dramatically when C4 is lower than -0.2 and IE7 higher than 50. On the Figure 2a, this point is shown in relatively dark places. Non-Linear functions developed relating to C4 and A5 ratios take place in Figure 2b. Figure 2b indicates that the probability of going bankrupt increases dramatically when C4 is lower than -0.2 and A5 is between 5 and 30. Non-Linear functions developed relating to IE7 and A5 ratios take place in Figure 2c. Figure 2c indicates that the probability of going bankrupt increases dramatically when IE7 higher than 50 and A5 is between 5 and 30. As for the positive values of C4 and the low values of IE7 and A5, banks were evaluated as non-bankrupt.

Figure 2a: The 3D graph for C4 and IE7

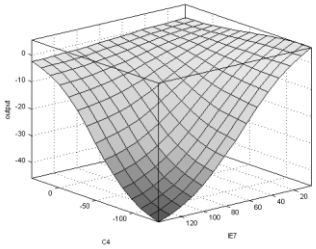


Figure 2b: The 3D graph for C4 and A5

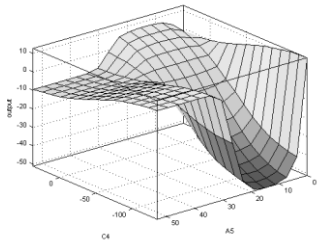
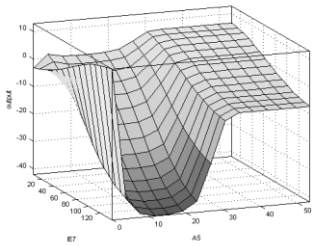


Figure 2c: The 3D graph for IE7 and A5



VI. CONCLUSİON

In this study, bank bankruptcies prediction was actualized with NF, with the aim of being an early warning system. Prediction accuracy of NF, ANN and MDA models were found at 90.91%, 86.36% and 81.82% respectively. Although there were not great differences between the performances of the models, it can be said that NF was slightly more successful. Models, which are built in the studies of bankruptcy prediction, generally focus on high prediction success. In addition to this success, how variables are used in the decision process by the models is also important. This study affirmed that in bank bankruptcy predictions, high prediction accuracy can be taken with NF. In addition to this, NF model presented how it makes its prediction. So, it is also possible for people to understand the nature of the bankruptcy model. This is NF's important superiority in comparison with ANN.

REFERENCES

Abonyi, J. (2003). *Fuzzy model identification for control*. Birkhauser

Akkoc, S. (2007). *Bankruptcy prediction using neuro fuzzy modeling and an empirical analysis.* Doctoral Dissertation. Kutahya, Turkey: The Dumlupinar Universty.

Alam P., D. Booth, K. Lee and T. Thordarson, 2000. "The use of fuzzy clustering algorithm and self-organizing neural network for identifying potentially failing banks: An experiment study."*Expert Systems with Applications* 18: 185–199.

Altman, E.I., 1968. "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *The Journal of Finance* 23 (4): 589-609.

Banking Regulation and Supervision Agency. 2001. The program of Restructuring in Banking Sector, www.bddk.org.tr

Banking Regulation and Supervision Agency. 2003. The program of Restructuring in Banking Sector, Progress Support - (VII), www.bddk.org.tr

Banking Regulation and Supervision Agency. 2009. "Financial Markets Report." March 2009, www.bddk.org.tr

Beaver, W.H. 1966. "Financial ratios as predictors of failure." *Journal of Accounting Research*, 4 (1): 71–111.

Bell, T.B., 1997. “Neural nets or the logit model: A comparison of each model’s ability to predict commercial bank failures, International Journal of Intelligent Systems in Accounting.” *Finance and Management* 6: 249–264.

Boyacioglu, M. A., Y. Kaya and O.K. Baykan. 2009. “Predicting bank financial failures using neural networks, support vector machines and multivariate statistical methods: A comparative analysis in the sample of savings deposit insurance fund (SDIF) transferred banks in Turkey.” *Expert Systems with Applications* 36 (2): 3355-3366.

Canbas, S., A. Cabuk and S.B. Kilic. 2005. “Prediction of commercial bank failure via multivariate statistical analysis of financial structure: The Turkish case.” *European Journal of Operational Research:*. 528–546.

Celikyilmaz, A., B.I. Türksen, R. Aktas, M.M. Doganay, and B.N. Ceylan. 2009. “Increasing accuracy of two-class pattern recognition with enhanced fuzzy functions.” *Expert Systems with Applications* 36: 1337-1354.

Chauhan, N., V. Ravi, and K.D. Chandra. 2009. “Differential evolution trained wavelet neural networks: Application to bankruptcy prediction in banks” *Expert Systems with Applications 36:* 7659-7665.

Chen, H. J., Huang, S.Y., and Lin, C. S. (2009). Alternative diagnosis of corporate bankruptcy: A neuro fuzzy approach.” *Expert Systems with Applications,* 36: 7710-7720.

Cielen, A., L. Peters and K. Vanhoof. 2004. “Bankruptcy prediction using a data envelopment analysis.” *European Journal of Operational Research* 154: 526–532.

Davalos, S., R.D. Gritta and G. Chow. 1999. “The application of a neural network approach to predicting bankruptcy risks facing the major US air carriers: 1979–1996.” *Journal of Air Transport Managements* 5 (2): 81-86.

Doganay, M.M., Ceylan N.B. and R. Aktas. 2006. “Predicting financial failure of the Turkish Banks.” *Annals of Financial Economics* 1: 97-117.

Jang J.S.R. 1993. “ANFIS: Adaptive Network Based Fuzzy Inference System.” *IEEE Transactions on Systems, Man and Cybernetics* 23 (3): 665-685

Jang, S.R., C.T. Sun, and E. Mizutani. 1997. *Neuro-Fuzzy and Soft Computing.* New Jersey: Prentice-Hall,Inc.

Jo, H., I. Han and H. Lee. 1997. “Bankruptcy prediction using case-based reasoning, neural networks, and discriminant analysis.” *Expert Systems with Applications* 13 (2): 97-108.

Kao, C.S. and T. Liu. 2004. “Prediction bank performance with financial forecasts: A case of Taiwan commercial banks.” *Journal of Banking & Finance* 28: 2353–2368.

Karacabey, A.A. 2007. “Bank failure prediction using modified minimum deviation model.” *International Research Journal of Finance and Economics* 12: 147-159.

Kolari, J., D. Glennon, H. Shin and M. Caputo. 2002. “Predicting large US commercial bank failures.” *Journal of Economics and Business* 54 (32): 361–387.

Lanine, G. and R.V. Vennet. 2006. “Failure prediction in the Russian bank sector with logit and trait recognition models.” *Expert Systems with Applications* 30: 463-478.

Lee, K., D. Booth and P. Alam. 2005. “A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms.” *Expert Systems with Applications* 29 (1): 1-16.

Leshno, M. and Y. Spector. 1996. “Neural network prediction analysis: The bankruptcy case.” *Neurocomputing* 10: 125-147.

Malhotra, R., and Malhotra, D. K. 2002. “Differentiating between good credits and bad credits using neuro-fuzzy systems.” *European Journal of Operational Research,* 136: 190-211.

Martin, D. 1977. “Early warning of bank failure.” *Journal of Banking and Finance* 1: 249-276.

Meyer, A. Paul and H.W. Pifer. 1970. “Prediction of bank failures.” *The Journal of Finance*, 25: 853-858.

Ng, G.S., C. Quek and H. Jiang. 2008. “FCMAC-EWS: A bank failure early warning system based on a novel localized pattern learning and semantically associative fuzzy neural network.” *Expert Systems with Applications* 34: 989–1003.

Nguyen, M.N., D. Shi and C. Quek. 2008. “A nature inspired Ying–Yang approach for intelligent decision support in bank solvency analysis.” *Expert Systems with Applications* 34: 2576-2587.

Kumar, P.R. and V. Ravi. 2007. “Bankruptcy prediction in banks and firms via statistical and intelligent techniques - A review.” European *Journal of Operational Research* 180: 1–28.

Piramuthu, S. 1999. “Financial credit-risk evaluation with neural and neurofuzzy systems.” *European Journal of Operational Research,* 112: 310–321.

Ravi, V. and C. Pramodh. 2008. “Threshold accepting trained principal component neural network and feature subset selection: Application to bankruptcy prediction in banks.” *Applied Soft Computing* 8 (4): 1539-1548.

Salchenberger, L., C. Mine and N. Lash. 1992. “Neural networks: A tool for predicting thrift failures.” *Decision Sciences* 23: 899–916.

Sharda, R. and R.L. Wilson. 1993. “Performance comparison issues in neural network experiments for classification problems.” *Proceedings of the 26th Hawai International Conference on System Scientists*.

Sinkey, J.F. 1975. “A multivariate statistical analysis of the characteristics of problem banks.*” Journal of Finance* 30: 21-36.

Swicegood, P. and J.A. Clark. 2001. “Off-site monitoring for predicting, bank under performance: A comparison of neural networks, discriminant analysis and professional human judgment.” *International Journal of Intelligent Systems in Accounting, Finance and Management* 10: 169–186.

Tam, K.Y. 1991. “Neural network models and the prediction of bank bankruptcy.” *Omega* 19 (5): 429–445.

Tam, K.Y. and M. Kiang. 1992. “Predicting bank failures: A neural network approach.” *Decision Sciences* 23: 926–947.

Tan, N.W. and H. Dihardjo. 2001. “A study on using artificial neural networks to develop an early warning predictor for credit union financial distress with comparison to the probit model.” *Managerial Finance* 27 (4): 56-77.

Tsukuda, J. S. and I. Baba. 1994. “Predicting Japanese corporate bankruptcy in terms of finance data using neural network.” *Computers and Industrial Engineering* 27 (1-4): 445-448.

Tung, W.L., C. Quek and P. Cheng. 2004. “GenSo-EWS: A novel neural-fuzzy based early warning system for predicting bank failures.” *Neural Networks* 17: 567–587.

West, R.C. 1985. “A factor analytic approach to bank condition.” *Journal of Banking and Finance* 9: 253–266.

Wilson, R.L. and R. Sharda. 1994. “Bankruptcy prediction using neural networks.” *Decision Support Systems* 11: 545–557.

Yang, Z.R., M.B. Platt and H.D. Platt. 1999. “Probabilistic neural networks in bankruptcy prediction.” *Journal of Business Research* 44 (2): 67 - 74.

Yildiz, B. and Akkoc, S. 2009. “Predicting bank bankruptcies with neuro fuzzy method.” *Journal of BRSA Banking and Financial Markets,* 3 (1): 9-35.

Zadeh, L. 1965. “Fuzzy Sets.” *Information and Control* 8 (3): 338-353.

Zhang, G., M.Y. Hu, B.E. Patuwo and D.C. Indro. 1999. “Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis.” *European Journal of Operational Research* 116: 16-32.

The Effects of Structural Deregulation on State Banks in New York During the 1970s

Michelle Zagardo*

ABSTRACT

New York eliminated restrictive branching and merging laws in 1976. A study of national- and state-chartered banks in New York found that the number of commercial banks as well as the number of mergers decreased during the 1970s (Kunreuther, 1976). In a study of state- chartered banks I find no change in merger activity as well as an increase in the number of state-chartered commercial banks during this period. Additionally, I categorize banks by size to examine their activity around the period of deregulation and find differences in behavior when comparing banks with total assets above the state median to banks with total assets above the state average.

I. INTRODUCTION

In 1971 the state of New York passed banking legislation that would eliminate restrictive branching and merging laws within the state in 1976. I compiled a data set to explore the effects of structural deregulation on New York state-chartered commercial banks during this time period and I found that there was an increase in bank mergers in 1976. In addition, the number of state-chartered commercial banks increased between 1972 and 1975 which is opposite what Kunreuther (1976) found when she studied this deregulation.

Restrictive structural regulations were pervasive in the United States banking system until the late 1990s. When structural deregulation began, it took place in a piecemeal fashion which made banking the ideal industry in which to evaluate the response of market structure to distinct policy changes. A review of the literature revealed several nationwide studies of structural deregulation but only one study done at the state level (Carlson & Mitchener, 2009).

The unique structure of the New York banking system in the 1970s and the legislative reforms combined with the concentration of many of the nation's largest banks operating in New York provided a rich environment to analyze that state's banking market. The 1971 law provided for an intermediate period between 1972 and 1976 that let small banks prepare for the likely increased competition they would face from bank holding companies and larger banks from urban markets.

* Mount Holyoke College, class of 2010, 50 College Street, South Hadley, MA 01075, zagar20m@mtholyoke.edu.
Paper from the 4th annual NYSEA Undergraduate Paper Competition.

The customary categorization of banks by size in the literature is median asset size but this practice does not capture the control of very large bank holding companies in the market. Therefore, I categorized banks by average asset size which preserved the dominance of the bank holding companies in the analysis and revealed stark differences between the two classifications.

II. A BRIEF HISTORY

In 1934 the Stephens Act was passed in NY and divided the state into nine banking districts. Banks were permitted to branch and merge within their district. This was the beginning of a period of continuous regulatory change that affected the activities of banks and bank holding companies (BHCs). Regulators struggled between allowing increased competition in the evolving industry while protecting the interests of smaller banks. In 1967 regulators and law makers began reexamining the Stephens Act and, after much debate, proposed legislation to eliminate the nine banking districts (Robards, 1975). William T. Dentzer was the superintendent of banks in NY at the time and said that:

> "the problem in all of this for any public official is to be sure that in the end we come up with a system which is more competitive than it would otherwise be if it's left to go its own normal, course, without any inhibitions" (Heinemann, 1970, September 8, 59).

The legislation passed in June 1971 and would become effective in 1976 (Heinemann, 1971, June 5, 37). Banks would be allowed to branch and merge throughout the state beginning in 1976.

III. THE LITERATURE

Regulators speculated about the effects that the elimination of banking districts in NY would have on entry, exit, and mergers in the market. In a nation wide study to identify which banks engage in mergers, Wheelock and Wilson (2004) found that the largest banks in a BHC are more likely to be the acquiring banks and that banks from urban areas make more acquisitions than those in rural areas. Carlson and Mitchener (2009) studied branching in California during the 1920s and found that the Bank of America was more likely to acquire state-chartered banks than national banks. McLaughlin (1995) found that when states lifted the ban on intrastate banking, BHCs merged their bank subsidiaries.

Kroszner & Strahan (1997) study the time until a state deregulates and they use the median asset size of banks as the cutoff between large and small banks. When using this classification with my data I found that it did not represent BHC activity and so I classified banks using the average asset size.

According to a 1976 study by Kunreuther the number of bank mergers in New York decreased throughout the 1970s. There were 130 bank mergers between 1961 and 1970 and 46 bank mergers between 1971 and 1975. The study included both state and federal commercial banks in New York but does not include a breakdown of mergers by year or type of charter. Kunreuther's finding drives her conclusion that the number of mergers that regulators were willing to approve decreased. She also found that the total number (federally- and state-chartered banks) of commercial banks in New York decreased from 296 in 1970 to 276 in 1975.

IV. DATA

I construct a data set of state-chartered banks in New York between the years 1965 and 1979. The bank level, balance sheet data come from the "Consolidated Statement of Conditions of Banks, Trust companies, and Private Bankers," compiled by the New York State Banking Department. I obtained this historical data from the New York State Library. The bank balance sheet information is as it was reported on December 31 of each year between 1965 and 1979 except the year 1966 in which banks reported on June 30.[2]

The literature reveals that there are important differences between big banks and small banks so I use two measures of "bigness" to examine the banking activity. First I define "big" banks to be those with total assets above the average in each year. I chose this measure because it is skewed toward the biggest banks and therefore provides a distinction between small banks and those banks that are among the largest in the country (as some New York banks were). Additionally, if a bank had assets above the average in each year and was a member of a BHC then it was one of the biggest banks in its BHC. Wheelock and Wilson (2004) differentiate the biggest bank in a BHC from the rest of the banks in the BHC and find that they are more likely to engage in mergers. This is because a BHC is likely to use the big bank to acquire the other BHC subsidiaries once merging is no longer limited geographically.

The second measure of "bigness" that I use defines "big" banks to be those with assets at or above the median in each year. Whenever there was an even number of banks, I chose the big banks to be those with assets strictly greater than the median. This is a measure used by Kroszner and Strahan (1997) as well as Stiroh and Strahan (2003). Both measures reveal different characteristics of banking activity before and after the policy change.

V. RESULTS

Banks were broken up into categories using two definitions of "big". Regardless of the definition of "big" Tables 1 and 2 show that there is a market consolidation in 1976 that coincides with the elimination of banking districts.

When "big" is those banks with above average assets, the market consolidation of 1976 includes the acquisition of 3 big banks, 6 national banks, and 12 small banks. There were 3 big banks that were each acquired by a member of their own BHC during the consolidation of BHC subsidiaries. When "big" banks are those with assets above the median, the 1976 consolidation is comprised of the acquisition of 14 big banks, 6 national banks, and 4 small banks. After 1976, with the exception of the Chemical Bank Corporation merging its subsidiaries in 1978, overall merger activity decreased following the removal of branching and merger restrictions. This could be because banks were able to open branches in order to expand their operations instead of expansion through mergers. When "big" is above the median, all of the acquiring banks except 3 are big.

Table 1: Big Bank Behavior
(those with assets above the average in each year)

	1972	1973	1974	1975	1976	1977	1978	1979	Total
Total number of banks	123	124	125	125	109	112	109	112	
Number of big banks	17	15	13	14	12	12	12	12	
Percent big	13.8	12.1	10.4	11.2	11.0	10.7	11.0	10.7	
Number of big banks that made acquisitions	1	0	1	1	5	1	1	0	**
Number of banks acquired by big banks	1	0	1	2	21	1	5	0	31
Number of big banks that were acquired	0	0	0	0	3*	0	0	0	3
Number of national banks acquired	1	0	1	2	6	1	1	0	12

* Two acquisitions by Marine Midland bank and 1 acquisition by Bank of New York in the consolidation of their respective BHC subsidiaries.
**This value is not equal to the sum of the row values because some banks made acquisitions is more than 1 year.

Table 2: Big Bank Behavior
(those with assets above the median in each year)

	1972	1973	1974	1975	1976	1977	1978	1979	Total
Total number of banks	123	124	125	125	109	112	109	112	
Number of big banks	62	62	63	63	55	56	55	57	
Number of big banks that made acquisitions	3	0	6	3	8	2	2	1	**
Number of banks acquired by big banks	3	0	6	4	24	2*	6	1	46
Number of big banks that were acquired	0	0	0	0	14	0	0	0	14
Number of national banks acquired	2	0	5	2	6	1	1	1	18

*One of the banks was chartered and merged in the same year so it does not report balance sheet data
**This value is not equal to the sum of the row values because some banks made acquisitions in more than 1 year.

Tables 3 and 4 display small bank activity. There are no banks below the median asset that acquire national banks as they do when “small” is defined to be below the average. Banks below the average asset engaged in mergers steadily throughout the period with no bank making more than one acquisition at a time. The differences in the activity that is seen in the different threshold levels that define the size of the banks reveals that there is a cutoff between the biggest banks and those banks that acquire one bank at a time. Banks below the average do not appear to be affected by the policy change because they do not change their behavior in 1976 by acquiring more banks. Big banks, on the other hand, acquire several banks as soon as statewide mergers are permitted. There is a stark difference between above median asset banks and their smaller counterparts. There were no banks below the median asset that acquired a national bank and only three banks with assets below the median that made any acquisitions at all.

Table 3: Small Bank Behavior
(those with assets below the average in each year)

	1972	1973	1974	1975	1976	1977	1978	1979	Total
Total number of banks	123	124	125	125	109	112	109	112	
Number of small banks	106	109	112	111	97	100	97	100	
Percent small	86.2	87.9	89.6	88.8	89.0	89.3	89.0	89.3	
Number of small banks that made acquisitions	2	0	7	2	3	1	2	1	**
Number of banks acquired by small banks	2	0	7	2	3	1*	2	1	18
Number of big banks that were acquired	0	0	0	0	0	0	0	0	0
Number of national banks acquired	1	0	4	0	0	0	0	1	6

*Chartered and merged in the same year so it does not report balance sheet data
**This value is not equal to the sum of the row values because some banks made acquisitions in more than 1 year

Table 4: Small Bank Behavior
(those with assets below the median in each year)

	1972	1973	1974	1975	1976	1977	1978	1979	Total
Total number of banks	123	124	125	125	109	112	109	112	
Number of small banks	61	62	62	62	54	56	54	55	
Number of small banks that made acquisitions	0	0	2	0	0	0	1	0	*
Number of banks acquired by small banks	0	0	2	0	0	0	1	0	3
Number of big banks that were acquired	0	0	0	0	0	0	0	0	0
Number of national banks acquired	0	0	0	0	0	0	0	0	0

*This value is not equal to the sum of the row values because some banks made acquisitions in more than 1 year.

The New York policy change that I am analyzing by looking at state-chartered bank activity was passed in 1971 and became effective in 1972. Banking districts were eliminated in 1976. Kunreuther (1976) groups her data into the period between 1971 and 1975. While my functional data set does not include the years prior to 1972, it is likely that my data captures the trend because the policy was passed in 1972. If my results are biased from the year omitted in my data then it is probable that it would increase the number of banks that above median, non-BHC banks acquire. The acquirers are likely to have assets above the median because there is a very low incidence of below median asset banks making acquisitions throughout the entire period. They are also not likely to be the biggest members of BHCs because BHCs were interested in entering new markets and were therefore seeking acquisitions in other

districts. Since they were not the biggest banks in BHCs they were most likely banks with assets below the average in each year. BHCs were not able to merge their BHC subsidiaries until statewide mergers were permitted in 1976. Since BHC acquisitions are legally different than acquisitions by banks in BHCs, the BHCs themselves were making acquisitions during this period in different banking districts but waited to consolidate their subsidiaries through mergers until later.

In my data set there are 14 state-chartered banks acquired between 1972 and 1975. This accounts for one third of the acquisitions cited by Kunreuther (1976) during the period from 1971 to 1975. Small banks below the average asset acquired 11 banks during this period while above average banks made 3 acquisitions. Defining “big” through median assets changed this so that below median banks only made 2 acquisitions during this period while above median banks made 12 acquisitions. The state banks below the average asset in my data engaged in steady merger activity between 1972 and 1979. Above average banks also saw steady mergers but the large banks waited to make their acquisitions in 1976 while small banks below the average did not change their merger behavior around the policy change. This is likely due to the fact that below average banks are not the lead banks in their BHC.

Using the definition of big banks as those with assets above the median we can see that mergers are relatively constant throughout the period with a cluster of acquisitions in 1976. Kunreuther (1976) concluded that the decrease in mergers between the 1960s and 1970s in her data was due to a decrease in regulators willingness to approve mergers. However the relatively constant behavior regardless of the definition of “big” suggests that this may not have been the case for state-chartered banks.

There are possible reasons for the differences between Kunreuther’s and my findings. Wheelock and Wilson (2004) found that national banks were more likely to make acquisitions than state banks and banks in statewide branching states were more likely to engage in mergers than banks in limited branching and unit banking states. New York was a limited branching state before 1976. National and big banks may have changed their acquisition strategy once the policy was passed in 1971. They may have geared up for the mergers they would face in 1976 by reducing merger activity relative to their behavior prior to 1970 because they would no longer have to use mergers as their primary expansion method.

The year of second greatest activity for acquiring state-chartered banks was 1974. In fact, the same number of state-chartered banks acquired banks in 1974, before the policy change, as in 1976 (8 institutions), but the number of acquisitions was higher in 1976. If Kunreuther (1976) is correct that there was a decrease in mergers that the authorities would approve then the merger activity of state-chartered banks suggests that they may have been held to different regulatory standards. State bank merger activity appears differently than Kunreuther’s data that includes both national and state banks. The rate at which state banks merge appears relatively constant prior to 1975 so the decrease in mergers was on the end of national banks.

Kunreuther found that the total number of commercial banks, including both national and state banks, decreased from 296 in 1970 to 276 in 1975. In my data, state-chartered banks increase from 123 banks in 1972 to 125 banks in 1975. While I do not have the number of state-chartered commercial banks in

1970, the nature of the policy change suggests that the market consolidation did not begin until 1972. If there is a bias then the rate of mergers (about 3.5 per year between 1972 and 1975) by above median banks allows me to estimate that there would be 7 more banks in 1970 than in 1972. Therefore, state-chartered banks would have decreased from 130 banks in 1972 to 125 banks in 1975 indicating that state banks have a different experience than national banks which would have decreased by 15 in number during the same period.

Between 1976 and 1979, the number of acquiring banks decreased and the number of acquisitions more than doubled from 14 to 34 compared to the period between 1971 and 1975. Above average banks were primarily acquiring while the acquisitions by below average banks stayed constant. The notable difference between above average, big banks and small banks is that big banks acquired many banks at a time while small banks only acquired one bank at a time. Notice that big banks acquire 21 banks in 1976. There are 2 banks that each acquire 1 bank, 1 bank that acquires 2 banks, and together the other 2 banks acquire 17 banks. The Bank of New York (New York City) acquired 7 banks when it consolidated its BHC subsidiaries and Marine Midland Bank (Buffalo) acquired the other 10 banks while also consolidating its BHC subsidiaries. Acquisitions by big banks were clustered around 1976 while acquisitions by small banks were clustered around 1974 when 7 banks acquired one bank each. Notably, 4 of the 7 acquired banks were national banks. When looking at above median banks, 6 of the 8 that made acquisitions in 1976 acquired banks outside of their initial banking district.

Following the increase in mergers up to 1976 there was a decrease in merger activity (with the exception of the Chemical Bank that acquired member banks in its BHC in 1978). Given the policy change that allowed banks to expand through branches, it appears that banks decreased expansion through merger activity.

VI. CONCLUSION

The goal of my study was to examine the activity of state-chartered banks following structural deregulation. My results are consistent with the literature that the majority of acquiring banks are large, BHC members. Acquired banks primarily had assets below the state average. Banks below the average asset engaged in mergers steadily throughout the period with no bank making more than one acquisition at a time. After 1976, with the exception of the Chemical Bank Corporation merging its BHC subsidiaries in 1978, overall merger activity decreased following the removal of branching and merger restrictions. This could be because banks were able to open branches in order to expand their operations instead of expansion through mergers.

Unlike Kunreuther's (1976) study, there is no evidence in my data that mergers are decreasing throughout the time period. The difference between our data is that her data includes national banks. The push for structural deregulation began in 1967 so national banks and big banks may have changed their expansion strategies as they anticipated statewide branching and merging. Since national banks are more likely to engage in mergers a change in policy, or even regulatory focus, would affect them more than the state banks. I suggest this alternative to her theory that regulators stopped approving mergers.

Lastly, Kunreuther (1976) found that the total number of banks in New York decreased from 1970 to 1975. In my data, state banks had a net increase in numbers from 1972 to 1975.

ENDNOTES

1. I would like to thank Professors Melanie Guldi and James Hartley for their guidance during the research process.
2. The December 31 filing could not be located by the reference librarian with whom I corresponded.

REFERENCES

Carlson, M. & Mitchener, K. 2009. “Branch banking as a device for discipline: competition and bank survivorship during the Great Depression.” *Journal of Political Economy*, 111: 165-210.

Heinemann, H. E. 1970 (September 8). *New York Times* [national ed.], p. 59.

---------- 1971 (June 5). *New York Times* [national ed.], p. 37.

Kidder, K. 1971. “Bank Expansion in New York State: The 1971 Statewide Branching Law.” *Federal Reserve Bank of New York, Monthly Review* (November).

Kroszner, R. & Strahan, P.1997. “The political economy of deregulation: evidence from the relaxation of bank branching restrictions in the United States.” Federal Reserve Bank of New York, Research Paper 9720.

Kunreuther, J. 1976. “Banking structure in New York State: progress and prospects.” *Federal Reserve Bank of New York, Monthly Review* (April).

McLaughlin, S. 1995. “The impact of interstate banking and branching reform: evidence for the states.” *Current Issues in Economics and Finance*, Federal Reserve Bank of New York..

Robards, T. 1975 (November 2). *New York Times.* [national ed.], p. 181.

Wheelock, D. & Wilson, P. 2004. “Consolidation in U.S. banking: which banks engage in mergers?” *Review of Financial Economics*, 13: 7-39.