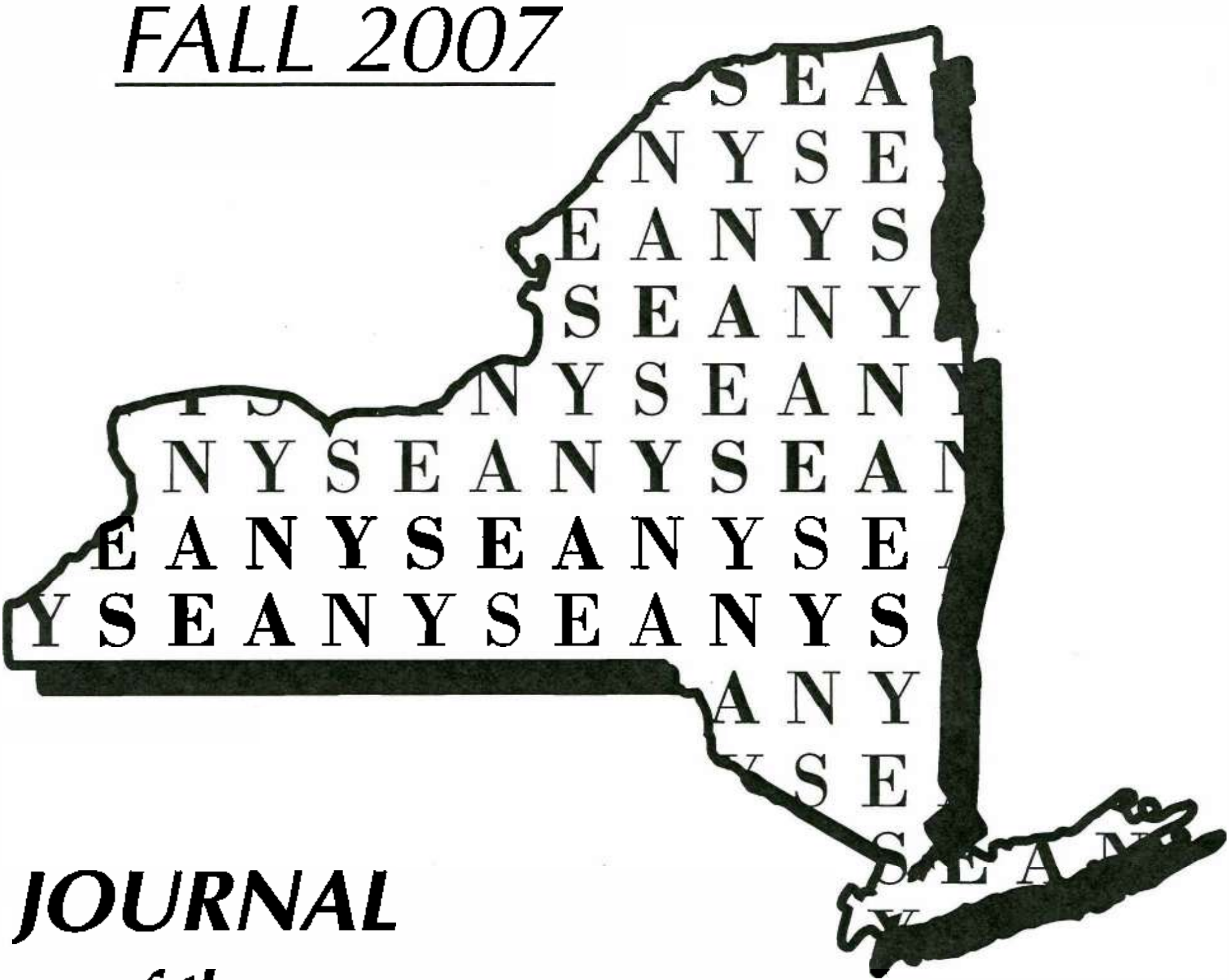
NEW YORK ECONOMIC REVIEW

FALL 2007



JOURNAL

of the

NEW YORK STATE

ECONOMICS ASSOCIATION

VOLUME XXXVIII

NEW YORK STATE ECONOMICS ASSOCIATION

≻ FOUNDED 1948 ≺

2006-2007 OFFICERS

President:
JAMES BOOKER • SIENA COLLEGE

Vice-President:
BRYAN McCANNON • ELMIRA COLLEGE

Secretary:
MICHAEL McAVOY • SUNY-ONEONTA

Treasurer:
ALFRED M. LUBELL • SUNY-ONEONTA

Editor:
WILLIAM P. O'DEA • SUNY-ONEONTA

Managing Editor:
JOHN W. PICCIONE • ROCHESTER INSTITUTE OF TECHNOLOGY

Web Editor:
WADE L. THOMAS • SUNY-ONEONTA

Board of Directors:
RICHARD DEITZ • FEDERAL RESERVE BANK-BUFFALO BRANCH
JOSEPH EISENHAUER • CANISIUS COLLEGE
WILLIAM KOLBERG • ITHACA COLLEGE
PATRICK MEISTER • ITHACA COLLEGE
JOHN PICCIONE • ROCHESTER INSTITUTE OF TECHNOLOGY
ELIA KACAPYR • ITHACA COLLEGE
F. PAN SHU • SUNY-POTSDAM
A. DALE TUSSING • SYRACUSE UNIVERSITY
MARTHA WOJTOWYCZ • SUNY UPSTATE MEDICAL UNIVERSITY

The views and opinions expressed in the Journal are those of the individual authors and do not necessarily represent those of individual staff members.

New York Economic Review
Vol. 38 Fall 2007

CONTENTS

ARTICLES

EDITORIAL

The *New York Economic Review* is an annual journal, published in the Fall. The *Review* publishes theoretical and empirical articles, and also interpretive reviews of the literature. We also encourage short articles. The *Review's* policy is to have less than a three month turnaround time for reviewing articles for publication.

MANUSCRIPT GUIDELINES

1. Please submit three copies of a manuscript.

2. All manuscripts are to be typed, double spaced and proofread. Prepared on a IBM PC/compatible computer in Microsoft Word format, the computer disk should be submitted in addition to the three hard copies.

3. All charts and graphs *must* be reproduction quality (Microsoft Word or Excel).

4. Footnotes should appear at the end of the article under the heading of "Endnotes."

5. Citations in the text should include the author and year of publication, as found in the references, in brackets. For instance (Marshall, 1980).

6. A compilation of bibliographic entries should appear at the very end of the manuscript under the heading "References."

Manuscript submissions should be sent to the editor, William O'Dea.

Division of Economics and Business
State University of New York College at Oneonta
Oneonta, New York 13820

The NYER is cited in:

- Ulrich's International Periodicals Directory
- Cabell's Directory of Publishing Opportunities
  in Business and Economics

ISSN NO: 1090-5693

Copyright © 2007 New York State Economics Association

Changes over Time in New York State's Responsiveness to Monetary Shocks

Todd Potts and David Yerger*

ABSTRACT

This paper extends the existing research on regional economic responses to Federal Reserve policy shocks along two dimensions. First, we focus on the evolution over time of a particular region's responsiveness to federal funds shocks. This differs from prior work that analyzed differences across regions in their responsiveness to a federal funds shock over a single sample period. For the state of New York, we track how the declining importance of interest-rate sensitive manufacturing sub-sectors and construction has altered the region's income response to federal funds rate shocks. The evolution of New York State's responses to fed funds shocks is contrasted with the changes in the Rest-of-Nation's responsiveness.

This paper's second extension of the literature is its use of sequential updating of the data set. Prior research utilized quarterly data sets starting in the late 1950's and ending in the early 1990's. We construct a parsimonious structural VAR model and first estimate the model over the 1958Q1 to 1992 Q4 period. Over this period our results are consistent with earlier findings. Next, we roll the sample period forward one year at a time, keeping the time period's length constant, up through 2004 Q2 and re-estimate the model after each resetting of the sample period.

Overall, our findings are consistent with the view that the declining importance of interest rate sensitive sectors will lead to a decline in the responsiveness of a region's income growth to federal funds rate shocks. In both New York State and the Rest-of-Nation, responsiveness to federal funds rate shocks declined in the more recent periods in a manner consistent with their declining shares of regional income coming from interest-sensitive sectors. Consequently, estimating the model using the entire available data set leads to an overestimation of the current impact on both regions from a federal funds rate shock.

I. INTRODUCTION

It has been established that U.S. regions historically have differed in their responsiveness to changes in Federal Reserve policy as manifested by variations in the federal funds rate. Three potential channels by which regional variation can arise have been identified. These channels involve differences across regions in: the percent of regional output coming from interest-rate sensitive sectors, the percent of regional output coming from small rather than large firms, and the percent of regional lending activity done by small rather than large banks. Empirically, however, for U.S. regions it is only variations in the share of regional output from interest rate-sensitive sectors that has been

*Indiana University of Pennsylvania

shown to correlate significantly, and in the expected manner, with variations in regional income sensitivity to Federal Reserve policy shocks.

Prior studies investigating variation across U.S. regions in income sensitivity to federal-reserve policy shocks utilize data beginning in the latter 1950's and ending no more recently than 1992. Since 1992, the ongoing evolution of the U.S. economy has resulted both in further declines in the relative importance of construction and interest-rate sensitive manufacturing sub-sectors, as measured by percentage of total output, and in considerable variation across U.S. states in the rate of decline in the importance of these sectors. In this paper, we investigate how the declining importance of construction and interest-rate sensitive manufacturing in total output since 1992 has altered the responsiveness of New York State's income to federal funds shocks.

Given changes in the region's economic structure, our results indicate that one should be cautious in using the previous cross-sectionally based research results linking a region's responsiveness to federal funds shocks to its economic structure to infer likely changes over time within a region in its responsiveness to federal funds shocks. The summative nature of the cross-sectional regressions across states can mask considerable variation across states in the evolution, if any, in a state's responsiveness to federal funds shocks. We show that regional analysts interested in the impact of Federal Reserve policy upon any particular state or region can gain useful insights by tracing changes over time in the region's responsiveness to federal funds shocks.

Our 'rolling regressions' approach also reveals that one should be cautious in assuming constant parameters if the entire sample of available data since the latter 1950's is used. When we compare estimation results over the entire sample period against results from our most recent sample period, we find material differences in the estimated impact of federal funds shocks upon both rest-of-nation and New York State real personal income growth. The large structural changes in the U.S. economy at both the national and regional levels over the past 25 years limit the usefulness of pre 1970 data in forecasting the responsiveness of national or regional economies to interest rate shocks.

II. LITERATURE REVIEW[1]

During the manufacturing boom years of the post-WWII U.S. economy, material variation existed across regions in their responsiveness to monetary policy. Regions for which manufacturing was more (less) important showed more (less) sensitivity to monetary policy shocks in several early studies. Toal (1977) examined regional responses to monetary policy shocks from 1952-1975. He found relatively large regional responses for the Great Lakes, Mideast, and Southeast regions, but small responses in the Rocky Mountain and Southeast regions. Similarly, Garrison and Chang (1979) analyze regional manufacturing earnings from 1969-76 and conclude that monetary policy had the largest impact in the Great Lakes region and the smallest in the Rocky Mountain region. Rather than using multi-state regions, Garrison and Kort (1983) investigate the impact of monetary policy at the state level from 1960 to 1978. Consistent with the earlier work, they found that states in the Great

Lakes region are the most sensitive to monetary shocks while states in the Rocky Mountains are the least sensitive.

These early studies used single-equation models of personal income, earnings, or employment as a function of national money supply (or other monetary policy variable) and other control variables. A potentially important shortcoming of this approach is the lack of accounting for possible feedback effects between the regional economy and the rest of the nation, as well as any feedback effects in the response of monetary policy to other control variables. Carlino and DeFina (1998) controlled for feedback effects by estimating a VAR model, which allows for feedback effects across regions as well as between the monetary policy variable and oil price shocks. This important methodological advance generated regional results consistent with the earlier studies.

Analyzing the 1958-1992 period using quarterly data, they identified five core regions whose responsiveness to monetary policy closely resembled the national response: New England, Mideast, Plains, Southeast, and Far West. Three non-core regions, however, differed significantly from the national response to monetary policy. The Great Lakes region was more responsive to monetary policy shocks than were the core regions while the Rocky Mountains and Southwest regions were less responsive. Over their sample period national manufacturing averaged 19.2 percent of output while it was 27.0 percent of gross state product (GSP) in the Great Lakes region, the highest of all regions, and only 12.6 percent and 15.2 percent for the Rocky Mountains and Southwest respectfully, the two lowest regional shares.

In an extension of their first paper, Carlino and DeFina (1999) examine the variation in regional responsiveness to monetary policy shocks at the U.S. state level over the same 1958-1992 period. They find substantial variation across states in the eight-quarter cumulative response of real personal income to a one-percentage point federal funds rate increase. Once more, Great Lakes region states are the most sensitive, led by Michigan's nation-leading 2.7 percent point drop in real personal income in response to a 1.0 percent point rise in the federal funds rate, while the Rocky Mountain and Southwest regions contain most of the states whose sensitivities are well below the average-across-all-regions response of a 1.16 percent point decline in real personal income per 1.0 percent point rise in the federal funds rate.

Carlino and DeFina (1999) also estimate a cross-section regression using the 48 states' cumulative response to a federal funds rate shock as the dependent variable. The explanatory variables include the percentage share of a state's GSP from manufacturing along with other variables meant to proxy alternative explanations of the variation across regions in monetary policy sensitivity. Specifications excluding BEA regional dummy variables find that every 1 percent point increase in manufacturing's share of GSP is associated with a 0.027 percent to 0.029 percent point increase (in absolute value) in the cumulative response of state real personal income to a 1 percent point rise in the federal funds rate. Adding BEA regional dummies lowers the impact to a 0.012 percent to 0.015

percent point increase in the cumulative state real personal income response to the rise in the federal funds rate. Since the average cumulative response to the federal funds shock was 1.16 percent points, this implies that those states most dependent on manufacturing over the sample period had sensitivities to federal funds shocks that were 10 to 25 percent larger than the national average.[2]

None of the other potential explanations of the variation across regions in their monetary policy sensitivity were well supported by Carlino and DeFina's (1999) cross-section regression results. The percent of small firms in a region had no discernable impact upon the region's sensitivity to federal funds shocks.[3] The percent of loans by small banks in the region had a negative effect which is the opposite of the sign implied by theory.[4] Several studies since the work by Carlino and DeFina have reinforced the existence of a linkage between a region's income sensitivity to monetary policy shocks and the share of regional output from interest-sensitive sectors.

Gaudreault (2001) analyzes provincial responses to changes in Canadian monetary conditions. Ontario's output is the most sensitive to changing monetary conditions, followed by Quebec, while the Atlantic and Western provinces lag well behind the national level sensitivity to changing monetary conditions. These findings match up with the relative importance of manufacturing across the provincial economies. Arnold (2001) analyzes the effects of monetary policy across 58 regions within the five largest EU nations and finds that regional sensitivity to monetary policy is related to industry mix in the same manner as previously established for the U.S.A. Arnold and Vrugt (2002) analyze regional data for the Netherlands from 1973 to 1993 across 12 regions and 13 industry sectors. They conclude that industry sector variation accounts for most of the variation in interest sensitivity across regions. Arnold and Vrugt (2004) estimate the impact of interest rate shocks on regional output in Germany across ten provinces over the period 1970-2000. As for Carlino and DeFina (1999), they find that the differential effects of monetary policy are related to variations across provinces in industrial composition, but not to either variations in firm size mix or variations in bank size mix.

While the above regional studies typically take the entire manufacturing sector as the measure of a region's interest-sensitive output, more recent work using national data has identified considerable variation across manufacturing sub-sectors in their sensitivity to interest rate shocks. Irvine and Schuh (2005) estimate the interest-rate sensitivity of 27 different 2-digit SIC manufacturing, retail, and wholesale trade industries using quarterly data from 1959:1 to 2000:1. Using a variety of VAR models, they estimate several different measures of an industry's interest-rate sensitivity and then create a composite measure of interest-rate sensitivity from these results. They identify nine SIC sectors as being highly interest-rate sensitive. These nine sectors, in descending order of sensitivity are: Motor Vehicles, Retail Automotive, Transportation, Lumber, Stone, Clay, & Glass, Primary Metals, Rubber, Fabricated Metals, Textiles. This study examines the evolution over time in both the total manufacturing share of output and the share from these interest-rate sensitive sectors for both New York State and rest-of-nation.

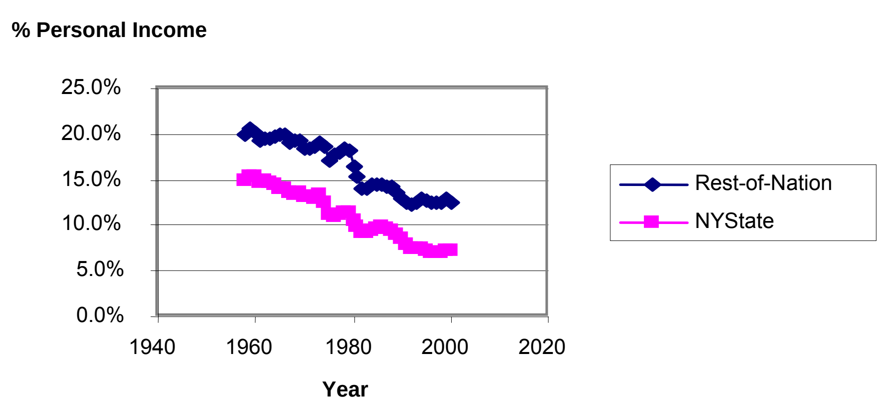
III. MOTIVATION FOR STUDY

The existing literature clearly has established that variations across regions within a monetary union in their responsiveness to monetary policy shocks are determined in part by the relative importance of interest rate sensitive sectors across regions. Regional policy makers and business decision makers, however, are likely to have different questions regarding the impact of monetary policy upon the regional economy. Is the region presently more, or less, susceptible to monetary policy shocks than the nation as a whole? If so, how large is this difference? Has the region's sensitivity to monetary shocks been changing in recent years, and if so, how?

We show that insights into these questions can be obtained from the results of a parsimonious structural VAR model that is re-estimated on an annual basis after rolling forward the sample period one year while holding the sample period length constant. By comparing the estimated impact of monetary policy shocks from the early estimating periods with those of the most recent estimating periods, the stability of the region's sensitivity to monetary shocks can be assessed. If these estimates differ materially across sample periods, then one should utilize the more recent sample period estimates for planning purposes and be cautious about using inferences based on econometric models utilizing all available data since the 1950's.

Our region of analysis in this paper is New York State. As seen in Figure 1, New York's share of personal income from construction and the nine interest-sensitive manufacturing & trade sectors of Irvine and Schuh (2005) has remained 5-6 percent points below the rest-of-nation's personal income share from these sectors. This suggests that New York State's personal income growth will be somewhat less sensitive to federal funds rate shocks than that of the rest-of-nation.

Figure 1: Percentage Personal Income From Construction + Interest Sensitive Sectors*



*Interest sensitive sectors defined by Irvine and Schuh (2005)

In this paper, we investigate how New York's responsiveness to monetary shocks has changed since the early 1990's given the underlying changes in its economic structure, and compare New York's evolving sensitivity to monetary shocks with the rest-of-nation's responsiveness to monetary shocks. Our findings are compared against the conclusions one would draw based upon the prior cross-sectional analysis literature on regional responses to monetary shocks, and key results are highlighted that illustrate the importance of specifically examining the region of interest rather than making inferences solely upon the region's relative ranking in terms of concentration of interest-sensitive sectors.

IV. MODEL

Model Specification

Economic activity in the state of New York and the rest-of-nation is modeled using a structural vector autoregression (SVAR) model. We analyze the dynamic behavior of the 5 x 1 covariance-stationary vector:

(1) $Z_t = [\Delta y_{st}, \Delta y_{Nt}, \Delta p_{ct}, \Delta p_t, \Delta m_t]'$

where Δy_{st} is the growth rate of state real personal income for New York state at time t, Δy_{Nt} is the growth rate of real personal income for the rest-of-nation at time t, Δp_{ct} is the growth rate in the core CPI, Δp_t is the growth rate of the relative price of oil, and Δm_t a measure of monetary policy actions.

The dynamics of Z_t are represented by

(2) $AZ_t = B(L) Z_{t-1} + e_t$

where A is a 5x5 matrix of contemporaneous correlation coefficients among the variables, B(L) is a 5x5 matrix of polynomials in the lag operator L, and e_t is 4x1 vector of structural disturbances, or primitive shocks, so $e_t = [\varepsilon_{st}, \varepsilon_{Nt}, \varepsilon_{pct}, \varepsilon_{pt}, \varepsilon_{mt}]$. Each variable in the model can be affected by its own idiosyncratic shock as well as by shocks to any of the other variables. The contemporaneous correlation coefficients in A, and the lag operators in B(L) will specify how shocks to any one variable are transmitted throughout the system of equations. If we rewrite equation (2) in its reduced form we see that:

(3) $Z_t = C(L) Z_{t-1} + \mu_t$

where $C(L) = A^{-1}B(L)$ is an infinite-order lag polynomial and $\mu_t = A^{-1} e_t$ provides the link between the model's structural residuals and its reduced form residuals.

We estimate the elements of A and B(L) using Bernanke's (1986) procedure. First, OLS estimates of the reduced form error terms $\mu_t = A^{-1} e_t$ are obtained from the estimation of equation (3). Next, the variance-covariance matrix for the structural errors e_t is restricted to be both orthogonal, zero contemporaneous covariance, and normalized to unity. This restricts the structural errors variance-covariance matrix to be an identity matrix. Lastly, we need to impose sufficient restrictions upon the A matrix to permit identification of A. Once A is identified, B(L) is estimated from $C(L) = A^{-1}B(L)$ where C(L) comes from the estimation of equation (3).

We place the following restrictions on the A matrix. Shocks to rest-of-nation real personal income do not affect New York State's real personal income contemporaneously and New York State's real personal income shocks do not affect rest-of-nation real personal income contemporaneously. Shocks to either region, however, can influence the other region with a lag of at least one quarter. Another restriction is that Federal Reserve policy shocks do not contemporaneously affect any of the other variables in the system, but can affect them with a lag of at least one quarter. We also restrict the impact of real oil price shocks and core inflation upon either New York State or rest-of-nation real personal income to have zero contemporaneous impact, but allow for impacts with a lag of one quarter or more. The final restriction is that neither New York State nor rest-of-nation real personal income can contemporaneously affect real oil prices or core inflation, but can affect them with a lag of at least one quarter. The federal funds rate variable can be affected contemporaneously, however, by shocks from any of the other variables in the system. With these restrictions in place, the elements of the model are identified.

Selection of Variables

The data in this study are quarterly and range from 1958 Q1 through 2003 Q4. New York State and rest-of-nation economic activity, the Δy_s and Δy_N variables, are measured using real personal income. This is computed using BEA data on nominal real personal income by state and nation and then deflating using the national CPI. State-level CPI are not available over the entire sample period so the national series must be used. Instead of personal income, employment growth could be used as an alternative measure of economic activity. Carlino and DeFina (1998,1999) show that the conclusions drawn regarding differences between a region's responsiveness to monetary shocks and the national average responsiveness to such shocks do not depend upon whether income or employment is used to measure economic activity. Consequently, we confine our analysis to the use of real personal income and focus upon analyzing how the relationships change over time.

The real oil price variable, Δp, is included to account for potential aggregate supply shocks and is computed as the PPI for fuels and related products divided by the total PPI. The core CPI variable, Δp_{ct}, is included as another macroeconomic control variable. It is included to account for changes in the nominal federal funds rate that may be intended to maintain a certain real federal funds rate. The monetary policy action variable, Δm, is the change in the federal funds rate as is used by Bernanke and Blinder (1992), among others. Leeper, Sims, and Zha (1996) and Bernanke and Mihov (1998) argue that this is the preferred measure. Other alternative measures of Federal Reserve policy would include changes in non-borrowed reserves and the Boschen and Mills (1995) narrative indicator of monetary policy. Once more, however, Carlino and DeFina (1998,1999) show that the analysis of differences between regional and national responsiveness to monetary shocks is robust across these three variables capturing federal-reserve policy actions. Hence, we restrict our analysis to using the more frequently utilized federal funds rate measure of federal-reserve policy actions.

Specification Issues

For inferences from this analysis to be valid, the variables in the SVAR need to be stationary. Results from standard Phillips-Perron (PP) unit-root tests are reported in Table 1. Phillips-Perron tests were used because they allow the disturbances to be weakly dependent and heterogeneously distributed. The critical values are the same as for the Dickey-Fuller battery of unit root tests (Fuller 1976). The real personal income and real price of oil data are reported in logs and log first-differences (growth rates) while the federal funds rate is in levels and first differences. Not surprisingly, the variables are non stationary in their level or log-level form, but are stationary when converted into first difference growth rates. So, the estimated SVAR model will use stationary first differences of real personal income for New York and rest-of-U.S., real personal income, real oil prices, and the federal funds rate.

For the actual estimation of the model, a four-lag structure was used on all variables. This is a sufficiently lengthy period to permit dynamics to work through the system. In addition, the Ljung-Box Q test statistics for the four-lag specification show that the null hypothesis of white noise error terms cannot be rejected at the .13 significance level for any of the system's equations.

Table 1: Phillips-Perron Unit-Root Test Results

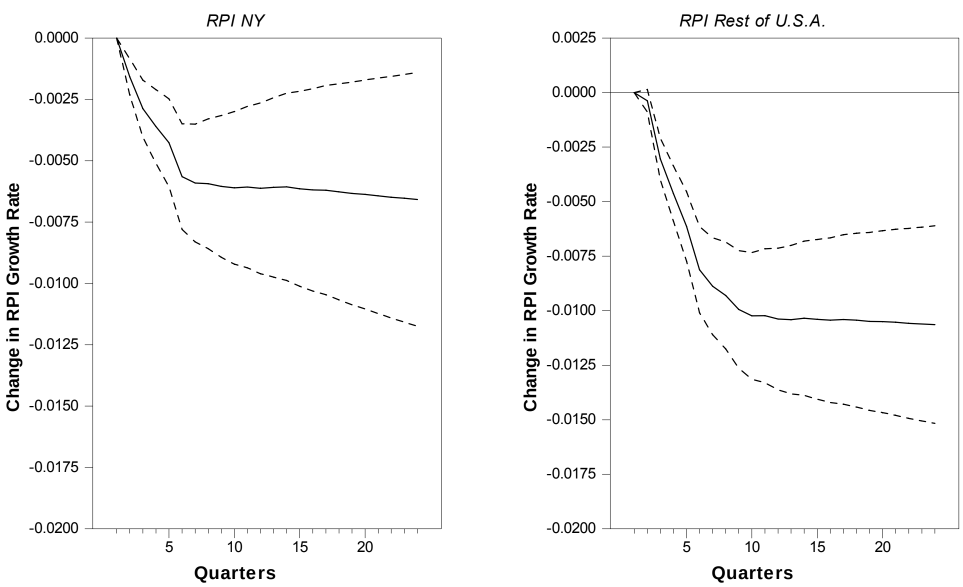
Levels	1st-Difference	
New York RPI	-1.91	-13.68*
Rest-of-Nation RPI	-1.82	-9.96*
Relative Core CPI	-1.11	-4.71*
Relative Price Oil	-1.22	-9.03*
Fed Funds Rate	-2.30	-10.27*

*significant at the 1 percent level, critical values in Phillips and Perron (1988)
RPI is real personal income
Two covariance lags used for all variables. Levels estimates had intercept & trend; 1st difference had intercept only.

V. EMPIRICAL RESULTS

We first estimate the model over the 1958 Q1 to 1992 Q4 sample period in order to match exactly the time period used by Carlino and Defina (1998, 1999). The cumulative impulse response functions for the impact of a twenty five basis point increase in the federal funds rate upon New York real personal income growth (RPI NY) and rest-of-nation real personal income growth are given in Figure 2. Consistent with New York state's lower share of income from interest-sensitive sectors as seen in Figure 1 above, the state's maximum response to a federal funds shock is considerably smaller than for the rest-of-nation. New York State's maximum RPI growth rate response to the shock is -.006577 (which implies that the growth rate in New York's real personal income fell by .6577 percent), a value only 61.8 percent as large as the rest-of-nation's maximum response of -.010638 (tables of these values not reported to save space). These results match almost perfectly with those of Carlino and Defina (1999) who estimate that New York State's output sensitivity to federal funds shocks is only 62 percent as large as that of the rest-of-nation (see results in Table 4 of that paper).

Figure 2: Real Personal Income Growth Response to Federal Funds Shock - 58:1-92:4

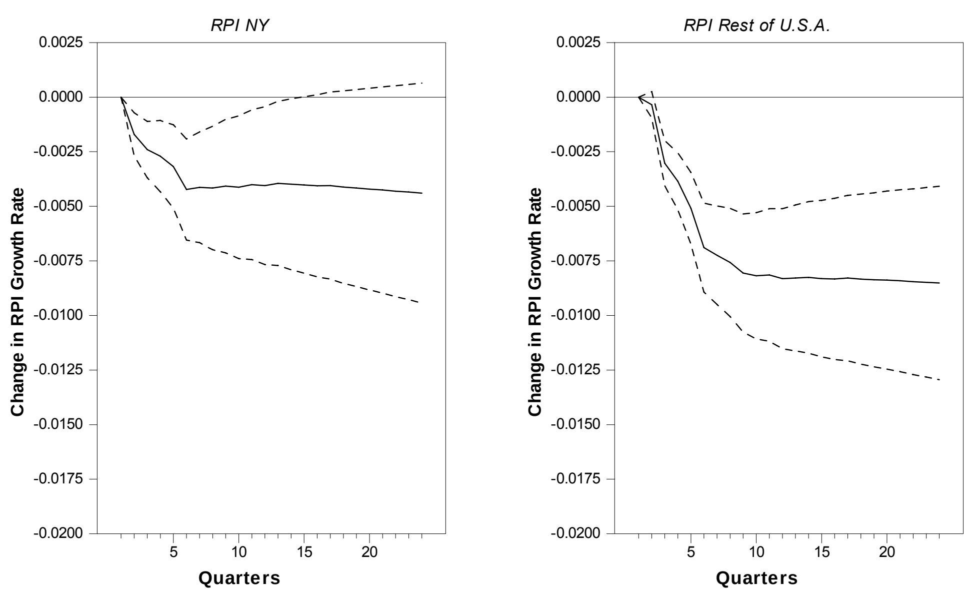


To investigate how this relationship may have changed over time, we next rolled the sample forward to cover 1969 Q1 to 2003 Q4. This time period uses the most recently available data at the time of analysis, and keeps the overall sample period length constant. Over this ending sample period, New York's share of personal income from construction and interest-sensitive manufacturing and trade sectors declined at approximately the same rate as for the rest-of-nation.

These changes raise several questions. Has New York's sensitivity to federal funds shocks declined as the importance of interest-sensitive sectors in its economy declined? How has New York sensitivity to federal funds shocks evolved relative to rest-of-nation's sensitivity given that both regions have exhibited similar trends for the importance of these sectors to total output?

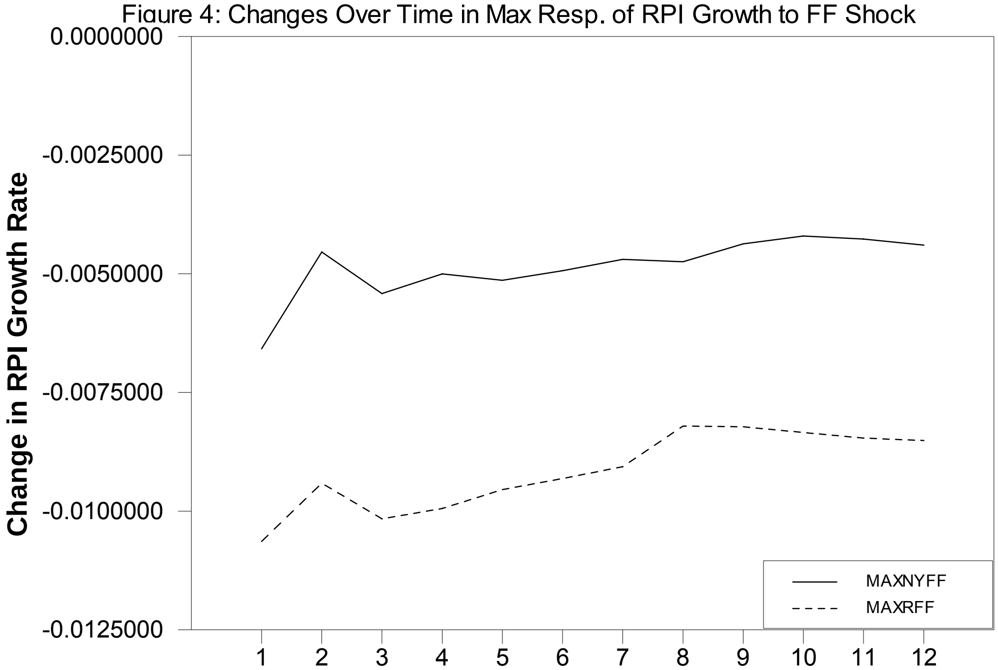
In Figure 3 we see the impulse response functions for RPI growth New York and RPI growth rest-of-nation in response to a twenty five basis point increase in the federal funds rate. For New York, the maximum response to the federal funds shock is -.004395 while for rest-of-nation the maximum response is -.008513. Comparing these results to those related to Figure 2, we find that both New York state and rest-of-nation's sensitivity to federal funds shocks declined considerably in the latter sample period, consistent with the declining importance of interest-sensitive sectors for both regions. New York State's personal income maximum response to federal funds shock in the 69:1-03:4 period is 33.2 percent smaller than it was for the 58:1-92:4 period. Similarly, the rest-of-nation's maximum response declines by 20 percent in the 69:1-03:4 period as compared to the 58:1-92:4 period. The slightly higher rate of declining interest rate sensitivity for New York state means that it has become even less susceptible than the rest-of-nation to interest shocks. For the last period, New York's maximum response of -.004395 is only 51.6 percent as large as the rest-of-nation's maximum response of -.008513, while for the 58:1-92:4 period New York's maximum response was 62 percent of the rest-of-nation's maximum response.

Figure 3: Real Personal Income Growth Response to Federal Funds Shock - 69:1-03:4



One possible cause of the differences in results over the two sample periods analyzed above is that the estimation results themselves are fragile and that parameter estimates fluctuate somewhat randomly as the sample period is adjusted. We check if the differences in results between the 58:1-92:4 sample period and the 69:1-03:4 sample period are due to a discernable ongoing trend in the estimated links between RPI growth and federal funds rate shocks, or if the results fluctuate randomly from sample period to sample period. To do this, we first run the model over the 58:1-92:4 time period and collect the maximum RPI growth rate response to a federal funds shock for both New York and rest-of-nation. Next, we rotate the sample period forward one year to 59:1-93:4 and repeat the same analysis. This procedure is done a total of 12 times until the final time period of 69:1- 03:4 is estimated. So, in Figure 4 Time 1 is 58:1-92:4, Time 2 is 59:1-93:4, …., and Time 12 is 69:1-03:4.

Figure 4: Changes Over Time in Max Resp. of RPI Growth to FF Shock



NOTE: Period 1 is 58:1-92:4, Period 2 is 59:1-93:4, ….Period 12 is 69:1-03:4

In Figure 4 the maximum response of New York RPI growth to a 25 basis point rise in the federal funds rate (MAXNYFF) for each sample period is reported along with the maximum response of rest-of-nation RPI growth to a 25 basis point rise in the federal funds rate (MAXRFF) where period 1 is the earliest sample and 12 is the latest sample. While the graph is not strictly monotonic, there is a clear trend over the rolling sample periods for a decline in the sensitivity of both New York's RPI and rest-of-nation's RPI to federal funds shocks as the magnitude of the negative effect from a rise in the federal funds rate is lower for more recent sample periods. Hence, our earlier findings of a 33.2 percent and 20 percent reduction over the earliest and latest sample periods in the sensitivity of RPI growth to federal funds shocks for both New York State and rest-of-nation respectively remain valid. Similarly, our earlier conclusion that New York State is less sensitive than the rest-of-nation to federal funds shocks and has become even less sensitive over time than has rest-of-nation also remains valid.

These findings are consistent with the decline over time in the importance of interest sensitive sectors for both New York State and the rest-of-nation as reported in Figure 1. Both regions had approximately a 7.5 percentage point decline in the share of personal income from interest sensitive sectors from 1958 to 2003. The downward trend in these shares matches well with the declining trend in the sensitivity of both regions to federal funds shocks as noted in Figure 4. Also, note that while both regions had approximately the same 7.5 percentage point decline in the share of personal income from interest sensitive sectors, since New York State started with a smaller share value than the rest-of-nation, the relative rate of decline in the importance of interest sensitive sectors was larger for New York State than the rest-of-nation. This is consistent with the earlier finding of a decline from 61.8 percent for the 1958:1—92:4 period to 51.6 percent for the 1969:1-2003:4 period in New York's maximum response to a federal funds shock relative to the rest-of-nation's maximum response to the same shock.

The trending values of the estimated impact from federal funds rate shocks seen in Figure 4 suggests that one might be better off confining their sample period to the more recent data so long as adequate degrees of freedom can be retained. If all available data is used, the resultant coefficient estimates might not best capture current dynamics for the purposes of inferring the likely impact today from a federal funds shock. To investigate this issue, we re-estimated the model using the entire 58:1-03:4 sample period. The maximum response of real personal income growth to a federal funds shock is compared against the results from the first and last sample periods of the rolling periods. The findings are summarized in Table 2 and indicate that using the entire sample period does in fact lead to an overestimate of both New York State and rest-of-nation's sensitivity to federal funds shocks. This is especially true for New York State. Comparing results for the 69:1-03:4 period against the 58:1-03:4 period, the estimated max response of real personal income to a federal funds shock is 29.9 percent higher for New York Sate and 6.1 percent higher for the rest-of-nation when the full sample is used.

Table 2: Estimated Maximum Responses of Real Personal Income Growth to Federal Funds Rate Shocks[5]

Period	RPI NY	RPI Rest-of Nation
58:1-92:4	-.006577	-.010638
69:1-03:4	-.004395	-.008513
58:1-03:4	-.005707	-009032

CONCLUSION

This study utilizes a structural VAR model to analyze the evolution since the early 1990's in the responsiveness of real personal income (RPI) growth to federal funds rate shocks for both New York and the rest-of-nation. We find that there has been a material reduction in the sensitivity of RPI growth to federal funds rate shocks for both New York and rest-of-nation. Comparing estimates over the 1958 Q1 to 1992 Q4 period with those from the latest period of 1969 Q1 to 2003 Q4, we find a 33.2 percent and a 20 percent decline in sensitivity of real personal income growth to federal funds rate shocks for New York and the rest-of-nation respectively. These declines in interest sensitivity are consistent with the declining trend for the share in personal income from interest-rate sensitive sectors for both New York state and rest-of-nation.

We completed a rolling regression VAR analysis that rotates the sample period forward one year at a time and shows that the RPI growth responses to federal funds shocks, for both New York and rest-of-nation, trended over time in a manner consistent with estimates from the first and last sample periods. These findings call into question the wisdom of estimating regional responses to federal funds shocks using all available data, and constant parameter estimate techniques, if one hopes to use the results to infer how the region's economy may respond in the near future to changes in the federal funds rate. Using the entire sample period, rather than the most recent 69:1-03:4 sample, overestimates the impact of a federal funds rate shock upon both New York and rest-of-nation RPI growth.

ENDNOTES

1. The material related to regional effects from monetary shocks in this section is drawn heavily from the excellent, and much more exhaustive, literature reviews found in Carlino and Defina (1998, 1999).
2. The basis for this estimate is that over their sample period, the average share of GSP from manufacturing was 20.1 percent while for several states it exceeded 30 percent. Using 10 percent points as the difference between the manufacturing intensive states and the national average, this adds between 10 * .012 = 0.12 percent points and 10 * .029 = 0.29 percent points to the cumulative response of real personal income. Starting from the national average base of 1.16 percent points, this is an increase of between 0.12 / 1.16 = 10.3 percent and 0.29 / 1.16 = 25 percent
3. See Bernanke and Blinder (1992) and Gertler and Gilchrist (1994) for an explanation of the 'credit view' of monetary policy transmission in which monetary policy shocks have a direct impact on banks' abilities to make loans. Since small firms are known to be more dependent than large firms on bank loans for financing, this suggests that regional variation in the small firms versus large firms mix may explain some of the regional variation in sensitivity to monetary shocks. Oliner and

Rudebusch (1995) provide another potential ‘credit channel’ that focuses upon the greater information asymmetry problems between firms and lenders for small firms rather than large firms. This implies that during tighter credit periods, a larger fraction of lending will go to larger firms.

4. Kashyap and Stein (1995) claim that since larger banks have more funding options than do smaller banks, lending by larger banks will be less sensitive to Federal Reserve policy changes than will be small bank lending. Hence, an increasing regional share of small banks should make the region more, not less, sensitive to federal funds rate shocks.
5. The summary of results in Table 2 highlights the three consistent findings in this paper. First, the responsiveness of RPI NY to a federal funds rate shock is smaller than the responsiveness of RPI Rest-of-Nation across all sample periods. Second, the responsiveness of RPI NY to a federal funds rate shock declines as the data sample period becomes more recent. Lastly, the responsiveness of RPI Rest-of-Nation also declines as the data sample period becomes more recent. It should be noted, however, that in this study, as often happens with VAR analysis utilizing a modest number of observations, the standard errors of the estimates are relatively large. Consequently, the 95 percent confidence band around both the RPI NY and RPI Rest-of-Nation parameter estimates encompass a wide range. Another consequence is that the hypothesis of no difference between RPI NY and RPI Rest-of-Nation cannot be rejected at the $p = .05$ level for any of the sample periods reported in Table 2. Nor can the hypothesis of no change in RPI NY (or of no change in RPI Rest-of-Nation) between the sample periods reported in Table 2 be rejected at the $p = .05$ level.

   Despite the inability to strictly reject these null hypotheses, the consistency of both the findings of New York having a lesser response than Rest-of-Nation, and of the sensitivity to federal funds shocks declining over time for both New York and Rest-of-Nation, indicate that the link between the importance of interest-sensitive sectors to a regional economy and that economy's responsiveness to federal funds shocks is operating in the expected manner.

REFERENCES

Arnold, Ivo J.M. 2001. “The Regional Effects of Monetary Policy in Europe,” *Journal of Economic Integration*, Vol. 16, No. 3, pp. 399-420.

Arnold, Ivo J.M. and Evert Vrugt. 2002. “Regional Effects of Monetary Policy in the Netherlands,” *International Journal of Business and Economics*, Vol. 1, No. 2, pp. 123-134.

Arnold, Ivo J.M. and Evert Vrugt. 2004. “Firm Size, Industry Mix, and the Regional Transmission of Monetary Policy in Germany,” *German Economic Review*, Vol. 5, No. 1, pp. 35-55.

Bernanke, Ben. 1986. “Alternative Explanations of Money-Income Correlation,” *Carnegie-Rochester Conference on Public Policy*, Vol. 25, pp. 49-100.

Bernanke, Ben and Alan Blinder. 1992. “The Federal Funds Rate and the Channels of Monetary Transmission,” *American Economic Review*, Vol. 82, pp. 901-921.

Bernanke, Ben and Mihov Ilian. 1998. “Measuring Monetary Policy,” *Quarterly Journal of Economics*, Vol. 113, pp. 869-902.

Boschen, John and Leonard Mills. 1995. “The Relation Between Narrative and Money Market Indicators of Monetary Policy,” *Economic Inquiry*, Vol. 33, pp. 24-44.

Carlino, Gerald and Robert Defina. 1998. “The Differential Regional Effects of Monetary Policy,” *The Review of Economics and Statistics*, Vol. 80, pp. 572-587.

Carlino, Gerald and Robert Defina. 1999. “The Differential Regional Effects of Monetary Policy: Evidence From the U.S. States,” *Journal of Regional Science*, Vol. 39, No. 2, pp. 339-358.

Enders, Walter. 1995. *Applied Econometric Time* Series. (New York: Wiley).

Fuller, W.A. 1976. *Time Series* (New York, Wiley)

Garrison, Charles and Hui S. Chang. 1979. “The Effects of Monetary Forces in Regional Economic Activity,” *Journal of Regional Science*, Vol. 19, pp. 15-29.

Garrison, Charles and John Kort. 1983. “Regional Impact of Monetary and Fiscal Policy: A Comment,” *Journal of Regional Science*, Vol. 23, pp. 249-261.

Gaudreault, Carl. 2001. “Impact economique regional d’un changement dans les conditions monetaires au Canada,” Working Paper Ministry of Finance, Government of Canada

Gertler, Mark and Simon Gilchrist. 1994. “Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms,” *Quarterly Journal of Economics*, Vol. 109, pp. 309-340.

Irvine, F. Owen and Scott Schuh. (2005). “Interest Sensitivity and Volatility Reductions: Cross-Section Evidence”, *Federal Reserve Bank of Boston Working Paper No. 05-4*, Federal Reserve Bank of Boston.

Kashyap, Anil and Jeremy Stein. 1995. “The Impact of Monetary Policy on Bank Balance Sheets,” *Carnegie-Rochester Conference Series on Public Policy*, Vol. 42, pp. 151-195.

Leeper, Eric, Christopher Sims, and Tan Zha. 1996. “What Does Monetary Policy Do?” *Brookings Papers on Economic Activity*, Vol. 2, pp. 1-78.

Oliner, Stephen and Glenn Rudebusch. 1995. “Is There a Bank Lending Channel for Monetary Policy?” *Federal Reserve Bank of San Francisco Economic Review*, pp. 2-21.

Toal, William. 1977. “Regional Impacts of Monetary and Fiscal Policies in the Postwar Period: Some Initial Tests,” *Federal Reserve Bank of Atlanta Working Paper*

An Economic Theory of College Alcohol and Drug Policies

Bryan C. McCannon[1*]

ABSTRACT

Colleges employ a wide variety of policies to regulate alcohol and drug use. A model where a regulator monitors the activity of a heterogeneous population of individuals is presented. Engaging in the activity is desired by each, but the aggregate activity exhibits a negative externality. The regulator is unable to observe the quantity or propensity for the activity of any individual, but can establish a maximum acceptable amount of activity and imperfectly monitor compliance with the standard. In this environment the choices made, the need for regulation, and imperfect monitoring are investigated to show that effective policy depends on the goal of the regulator.

I. INTRODUCTION

Institutions of higher education are charged with ensuring the quality of life of their students. Prominent among the factors that affect students' quality of life is the use of alcohol and drugs on campus. It is well documented that abuse of such substances can lead to individual harm such as a lack of academic success and health problems. Also, the use of alcohol and drugs by one student affects the quality of life of other students. Examples of such spillover effects include property damage, sexual assault, violence, and a diminished living environment (Wechsler *et al*, 2000b).

College policies attempting to regulate use vary significantly. Some schools relegate oversight to governments while others self-monitor activity. While some institutions standardize one policy for all substances others set separate policies for alcohol, marijuana, and other illegal drugs. Punishments differ substantially as well. Punishments for a violation of the drug policy range from immediate suspension from the college to only a reprimand. Furthermore, the amount of acceptable activity differs. For example, some schools prohibit all alcohol on campus.[2] Mitchell, Toomey, and Erickson (2005) survey college alcohol policies, Wechsler *et al* (2004) survey college alcohol prevention initiatives, and Wechsler *et al* (2000a) survey college administrators' responses to binge drinking. They find a wide variation of policies and initiatives, which is dependent on factors such as the size of the school, the proportion of students in residence, religious affiliation, and whether it is public or private. With such a wide variety of policies this paper attempts to explain the variation and determine the characteristics of an effective policy.

Alcohol and drug use on college campuses is an example of a negative externality. Students want to engage in the activity. For example, Teter *et al* (2005) document motivations for illicit use of

---

*Division of Business and Economics, Elmira College, One Park Place, Elmira, NY 14901, (607) 735-1923, mccannbc@wfu.edu.

prescription stimulants, which include euphoria, concentration, and alertness. The actions of one individual, while beneficial to that person, are harmful to others. Colleges recognize this spillover in their drug policies. As an example, one college states, "these drugs can...cause situations in the setting of a residential college in which individual actions affect all members" (Handbook, 2004) while another points out, "above all else...members of the [community] shall...acknowledge the impact of alcohol on communal living and work to limit its negative effects" (Communications, 2004).

The environment developed in this paper captures three important characteristics of alcohol and drug use on college campuses. First, there is a population of students each of whom decides how much of an activity, which they enjoy, to engage in. Secondly, the extent to which a student enjoys the activity varies across the population. Thus, some students put little weight on the activity while others gain significant happiness from it. Finally, the aggregate amount of activity is harmful to each student. For example, suppose this activity is interpreted as the amount of alcohol consumed. Each student enjoys drinking but each differs in how much they enjoy alcohol. As the total amount of drinking increases on campus the quality of life of each student diminishes.

In this environment the choices made are investigated and the need for regulation is illustrated. Comparing the choice of a student without regulation to the best outcome for the overall student body, regulation is shown to be needed. In practice, the regulator is neither able to observe a student's propensity for the activity nor the actual amount of the activity she selects. An extension to the model is considered where a regulator, unable to perfectly monitor the activity of students, sets a maximum acceptable amount of the activity and observes, with a probability less than one, whether a student has exceeded the standard. If a student is found to exceed the standard the regulator is able to enact a penalty. An effective alcohol and drug policy is shown to depend crucially on the goal of the regulator. Three separate goals are discussed: a reduction in the amount of the activity on campus, an increase in compliance with the policy, and a maximization of the total well-being of the population of students.

Other authors have noted that effective policy implementation is dependent on the goal of the regulator. Caulkins and Reuter (1995, 1997) differentiate between use reduction and harm reduction as the goal of a national drug policy. Which goal is chosen greatly determines how drug use is attacked. As an example they compare a pregnant, recovering addict shooting heroin with an HIV-infected needle to an employed, emotionally stable adult with no dependents using marijuana at home. While both offenses are Schedule I prohibited drugs, the harm of the former is greater. Furthermore, they differentiate between prevalence use reduction and quantity use reduction, which are the first two goals studied here.

The first two goals considered are, as stated, a reduction in aggregate activity and an increase in compliance with the policy. For both more effective monitoring and stiffer sanctions are useful since they increase the expected cost of engaging in an excessive amount of the activity. Policy recommendations differ with regard to the level of acceptable activity. A decrease in the amount tolerated induces some, previously complying, to no longer follow the policy and increase their activity. Others, who continue to follow the rule, reduce the activity to satisfy the lower amount of acceptable use. Thus, such a decrease reduces compliance, but has an ambiguous effect on the aggregate amount of the activity.

Both of these goals aim for improvements in a specific variable, either aggregate amount of the activity or number complying. The third goal for the regulator is to maximize the total well-being of the population of students. It is shown that the optimal sanction must be greater for larger populations of students and activities that have a stronger spillover onto other students. Both affect the size of the negative externality and, thus, must be accompanied by stiffer penalties. This is similar to the goal of harm reduction discussed by Castro and Foy (2002) and Caulkins and Reuter (1997). Since the expected cost of violating the rule is the product of the probability of apprehension and the sanction there is an inverse relationship between the probability of catching a student who has exceeded the tolerated amount of activity and the optimal sanction. The greater this probability the greater the marginal cost to the activity, thus, to remain at the level that maximizes the well-being of the population, the sanction must decrease. This result contrasts that of the previous two goals. Similarly, the maximum amount of acceptable activity must be lower for larger student populations and more detrimental activities. Furthermore, in this environment a policy of abstinence is never best for the student body.

Others have used economic theory to describe the use of alcohol and drugs and the enforcement of laws. Benson and Rasmussen (1998) use economic theory to explain the determination of drug laws and monitoring by law enforcement agencies. Krebs, Costelloe, and Jenks (2003) model a game between the government and drug smugglers. Miron and Zwiebel (1995) and Miron (1998) investigate the effects of prohibition of alcohol in the 1920s to discuss the impact of current policies of prohibition of drugs. Here, a standard model of negative externalities is applied to a new application, college alcohol and drug policies, to generate policy recommendations and explain observed differences in policies.

This work is also related to the literature on optimal deterrence in the field of Law and Economics. Following from Becker's (1968) seminal work it has been shown that since sanctions and enforcement both add to the expected cost of violating a policy they are substitutes. It is possible to decrease one and increase the other and maintain the same amount of compliance. Since enforcement tends to be more costly than implementing stiffer sanctions, his work predicts that full compliance with the law occurs using maximal sanctions. Neither of which occur in practice. A sizeable literature has developed to explain the lack of full compliance and the use of nonmaximal sanctions (see Garoupa (1997) and Polinsky and Shavell (2000) for more detailed literature reviews). The literature can be organized along three lines of explanation. First, some have argued that imposing more severe sanctions comes at a cost that balances the cost of enforcement. Examples include direct costs to the legal system (Polinsky and Shavell, 1992), the costly avoidance activities more severe sanctions encourage (Malik, 1990), and the imposition of socially inefficient nonmonetary sanctions (Shavell, 1987). A second line of explanations focuses on various forms of regulation. Examples include the *ex ante* vs. *ex post* regulation of the production of harm (Wittman, 1977; Kolstad, Ulen, and Johnson, 1990), general vs. specific enforcement (Shavell, 1991), and the tradeoff between monitoring and investigation (Mookherjee and Png, 1992). Finally, frictions in the enforcement activities may lead to reduced sanctions and incomplete enforcement. Examples include imperfect information about the magnitude of the sanctions (Craswell and Calfee, 1986; Bebchuk and Kaplow, 1992) or rate of apprehension (Bebchuk and Kaplow, 1993), heterogeneous wealth constraints of the population restricting the severity of the sanction (Polinsky and Shavell, 1991), and the problem of marginal deterrence where an increased sanction for one offense encourages violation of another (Shavell, 1992; Wilde, 1992; Friedman and Sjostrom, 1993). The model presented here adds to the literature of friction-based explanations. Individuals select an amount of activity and the regulator is only able to observe whether an activity has exceeded a set threshold. This is a new, specific informational friction in enforcement that is particularly applicable to the use of alcohol and drugs.

The work presented here is novel in two aspects. First, it applies a rather standard model of negative externalities to an important application generating policy recommendations. The model is used to rationalize the variation in policies observed among colleges and universities. There has been no prior theoretical, economic inquiry into this application. Second, it adds to the literature on optimal deterrence of law breaking and policy violation by considering an environment not before considered in the literature. The work provides new explanations for the fact that maximal sanctions are rarely used and full compliance is not achieved. Section II presents the base model and illustrates the need for regulation. Section III presents an extension to the model considering the imperfect monitoring of the regulator, which provides the results stated. Section IV concludes. Proofs of the results are given in the appendix.

II. THE NEED FOR REGULATION

Consider a population of N students. Each student selects an amount of activity in which to engage. Let y_i denote the amount of the activity student i selects where $0 \leq y_i \leq y^+ < \infty$ where y^+ is the upper bound to y_i. The activity can be interpreted as the consumption of a good such as a quantity of alcohol or another drug. Define Y to be the aggregate amount of activity of the students. Thus, $Y = \sum_{i=1}^{N} y_i$. Let $u_i(y_i, Y)$ denote the utility to student i engaging in y_i activity if the aggregate activity is Y. Each student benefits from the activity she does but is hurt by the total amount of activity, or rather, utility is increasing in y_i and decreasing in Y. To provide an analytical solution let

(1) $u_i(y_i, Y) = a_i \ln y_i - bY$

where $0 \leq a_i \leq a^+ < \infty$ and $0 < b < \infty$ where a^+ is the upper bound to a_i. The amount of satisfaction students receive from engaging in the activity varies across the population. For simplicity, the marginal impact the aggregate activity has on each student, b, is the same for each student.

Consider, first, the outcome with no regulation where every student selects the amount of the activity in which to engage. From standard optimizing behavior each student chooses y_i at the level where her marginal benefit exactly equals her marginal cost. In this model the marginal benefit is the direct gain to the activity, while the marginal cost is the lost utility due to the increase in Y. From (1),

(2) $a_i / y_i = b$.

Solving (2), student i selects a_i / b. Suppose a regulator is interested in the well-being of the entire population. Such a regulator wants to maximize the sum of the utilities of all students, $\sum_{i=1}^{N} u(y_i, Y)$. Selecting y_i to maximize this expression it follows that the amount of activity of student i solves[3]

(3) $a_i / y_i = Nb$.

Figure 1: The Need for Regulation

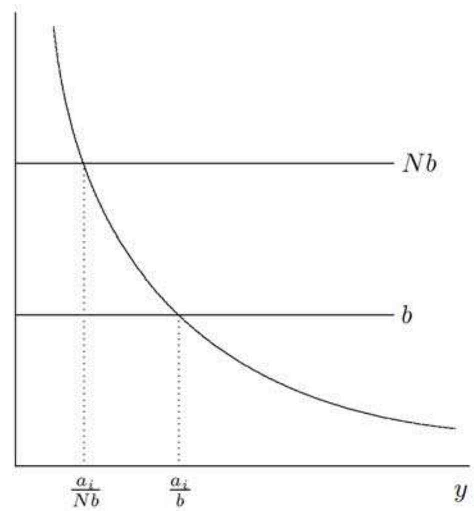


Figure 1 depicts the marginal benefit and marginal cost of both (2) and (3). Whereas student i would select a_i / b it would be best for the population if she only selected a_i / Nb. This is the standard negative externality result. Since the individual does not take into account the full cost of her activity too much is done. Notice that the amount of activity that is best for the population is less with a larger population or a more harmful activity because in either case the externality is greater. Finally, the amount of activity that achieves the greatest well-being for the population prescribes that each student does not engage in the same amount of activity. Those valuing the activity more do more of it. With full information the regulator can construct penalties that induce the socially optimal amount of activity.

Now consider an extension of the model where the regulator lacks the information needed to do so and can only imperfectly observe whether or not a violation of a set policy has occurred.

III. IMPERFECT MONITORING

The previous analysis studied the straightforward but unrealistic case where the regulator is able to observe both the amount of activity each student engages in and each student's propensity for the activity. In practice, neither is perfectly observable. How should a regulator act if it is unable to perfectly monitor the activity?

Suppose that the a_is are privately known. Since a_i represents student i's taste for the activity it is likely that this is not observable to the other students or the regulator. As a consequence, the regulator is unable to condition a penalty on a student's willingness to act. Also, assume that neither the set of y_is nor Y is observable to any student or the regulator. Furthermore, suppose the only available option for the regulator is to set a maximum acceptable amount of activity (i.e. "one size fits all"), y^*, which is referred to as the *standard*, and imperfectly observe whether or not a student complies with this standard. Assume that if a student exceeds y^* then with a probability less than one the regulator observes that $y_i > y^*$ and can levy a punishment. A student engaging in more of the activity has a greater probability of getting caught. Let the probability that a student selecting $y_i > y^*$ is caught be py_i where $p \geq 0$ and $py^+ \leq 1$. The regulator who has caught a student exceeding the standard knows that the policy has been violated, but is not able to determine the extent of the violation. In this model a student selecting less activity than the tolerated amount is never wrongfully punished. A penalty imposed on a student exceeding the standard with the informational constraints is a lump-sum sanction dependent only on being caught. Denote this punishment as f.

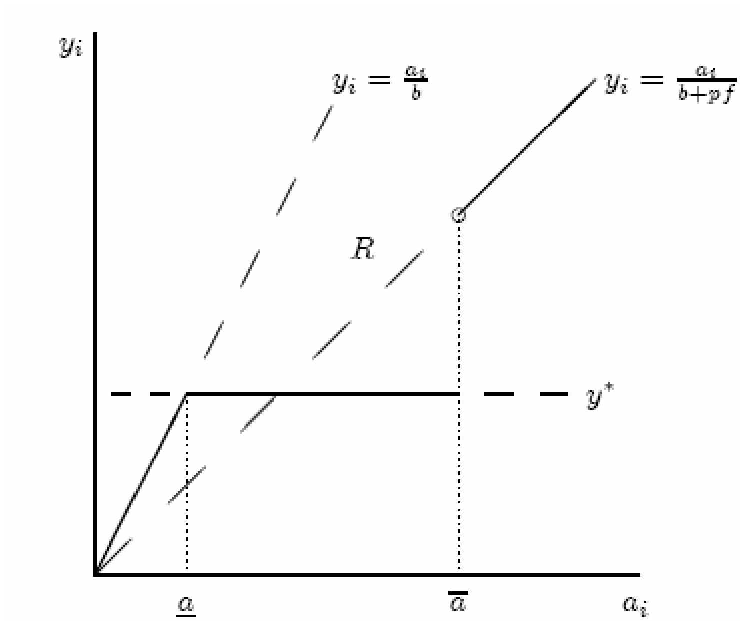
For many illegal activities this is a reasonable, accurate setup. A college administrator is unlikely to know exactly how much or how often a student engages in alcohol or drug misuse and is even less likely to know the aggregate amount of such activity on the campus. The institution is able to set limits to the activity and learn (imperfectly) whether or not this standard has been violated. Alternatively, if the regulator is able to observe either y_i or Y perfectly other enforcement mechanisms would be available.[4] Thus, this model focuses on activities, like alcohol and drug use on college campuses, that exhibit these characteristics and informational constraints.

How effective is the regulator? The expected utility of a student is

$$(4) \quad = \begin{cases} u_i(y_i, Y) - py_i f & \text{if } y_i > y^* \\ u_i(y_i, Y) & \text{if } y_i \leq y^* \end{cases}$$

First, a student who does not put much weight on the activity would select an amount less than the standard. Such a student is not affected by the regulation. Hence, for a given standard y^* there exists a value of a_i, denoted $\underline{a}(y^*)$, such that students with lower values of a_i select the same amount of activity as without regulation, a_i / b. A student with a large weight on the activity chooses to violate the policy. Again, such a student selects the amount that equates her marginal benefit with her marginal cost, which now includes the potential punishment. From (1) and (4) the optimal level of the activity is $y_i = a_i / (b + pf)$. Thus, for a given amount tolerated by the college administration there exists a second value of a_i, denoted $\overline{a}(y^*)$, and students with greater values select $a_i / (b + pf)$. Finally, a student with an intermediate value of a_i would exceed the standard if there were no regulation, but does not find it worth the potential cost to select that amount of activity. This student exactly complies with the amount tolerated selecting $y_i = y^*$. Figure 2 depicts the expected utility maximizing choice of y_i for students with various values of a_i. The derivation of $\overline{a}(y^*)$ and $\underline{a}(y^*)$ along with the proofs of the results are given in the appendix. The area above the line and below the $y_i = a_i / b$ dashed line, denoted R in Figure 2, is a measure of the reduction in activity due to the regulation.

Figure 2: Activity with Imperfect Monitoring



Consider, as an illustration, a policy of tolerating no more than $y^* = 4$ drinks, a fine of $f = 100$, a probability of $p = 0.01$ of catching a violation, and the aggregate activity has a $b = 1$ weight. Student 1 has a weight $a_1 = 20$ on the alcoholic drinks. For her breaking the rule with $a_1 / (b + pf) = 10$ generates a greater payoff, $20ln10 - bY - (.01x10x100) \cong 36 - bY$, than complying, $20ln4 - bY \cong 28 - bY$. Student 2 has a weight $a_2 = 10$. For him breaking the policy, with $a_2 / (b + pf) = 5$, results in a payoff of $10ln5 - bY - (.01x5x100) \cong 11 - bY$, which is less than the payoff from complying, $10ln4 - bY \cong 14 - bY$. Student 3 has a weight $a_3 = 1$ and, as a result, is only interested in having $a_3 / b = 1$ drink. Since this is an acceptable amount Student 3 consumes one drink. Hence, on Figure 2, Student 1 is beyond $\bar{a}$, Student 2 falls between $\underline{a}$ and $\bar{a}$, and Student 3 is less than $\underline{a}$.

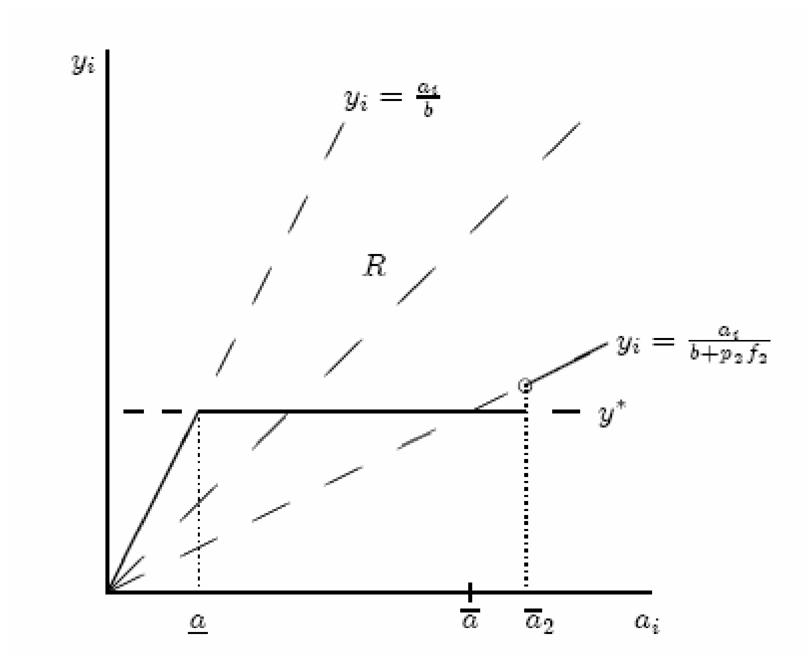
Result 1: *There is less activity with regulation than without.*

The regulator may be able to adjust its enforcement. Two possible goals of the regulator are to (1) reduce the aggregate amount of activity and (2) induce more compliance. The first would come from an increase in *R* while the second would be achieved by increasing $\bar{a}$.

EXPECTED COST

Suppose, first, that the regulator is able to stiffen its enforcement by either catching violations at a higher rate (increase *p*) or imposing a more severe sanction (increase *f*). Figure 3 illustrates the effect of an increase in either *p* to p_2 or *f* to f_2.

Figure 3: Activity with $p_2 > p$ or $f_3 > f$



Increasing *p* or *f* has the effect of increasing the expected cost when violating the standard. Such an increase would have two effects. First, it would reduce the activity of those continuing to violate the policy (rotation out of the a_i / $(b + pf)$ line). They reduce their activity to balance out the additional cost.

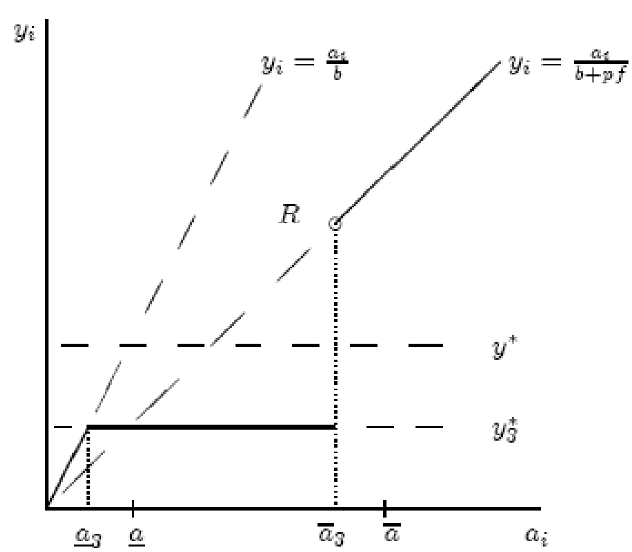
Second, it would encourage some students previously violating the policy to comply. The additional cost makes it no longer worthwhile to engage in the excessive activity (increase from $\bar{a}$ to $\bar{a}_2$).

Result 2: *An increase in enforcement (either p or f) reduces the amount of activity (increases R) and increases compliance (increases* $\bar{a}$*).*

STANDARD

Alternatively, the regulator may be able to adjust the standard, *y**. For simplicity, assume that the standard can be changed freely without affecting the ability to enforce it. In some situations technological constraints may preclude this. Alcohol and drug use can often be quantified; blood-alcohol levels, container size, and weights are examples of such measurements. Therefore, the assumption is used both because it seems realistic and it is useful to know what would be the preferred adjustments to the change in the standard. Consider a decrease in the amount tolerated by college officials from y^* to y_3^* so that less activity is acceptable as depicted in Figure 4. Again, there are two effects from such an adjustment. First, the amount of activity of those who continue to comply is reduced. Those previously choosing to exactly meet the standard are forced to reduce their activity, while some students, selecting an amount just less than the previous amount allowed (a_i in [$\underline{a}_3$, $\underline{a}$]), must now reduce their activity to comply with the new one. The second effect of the adjustment is that some students who were previously complying no longer find it worthwhile to do so because the marginal benefit at y_3^* now exceeds the marginal cost (a_i in [$\bar{a}_3$, $\bar{a}$]). Since their activity had been reduced to comply with the standard they now conduct more activity than before.

Figure 4: Activity with $y_3^* < y^*$



The two effects from the lowering of the acceptable level of activity (from y^* to y_3^*) have opposing impacts on the change in the aggregate amount of activity. Those complying with a_i between $\underline{a}_3$ and $\overline{a}_3$ reduce the amount of activity. Students with a_i between $\overline{a}_3$ and $\overline{a}$ no longer abide by the rule. A number of factors determine which of the two effects is stronger. One is the number of individuals complying. If many students are complying then a lowering of the tolerated amount would act to reduce the activity of much of the population. Thus, R would increase. The high rate of compliance could be due to a lenient, high standard (large y^*) or from a population that consists of many students with low values of a_i. Another factor that would effect which of the two effects dominate is the stringency of the enforcement. If the enforcement is tough (large p and/or f) then the increase in activity of those that now choose to comply is mitigated. Thus, the gain in activity would be less. Furthermore, a high level of enforcement induces more to comply, which enhances the reduction in the activity.

Result 3: *A decrease in y^* decreases compliance, but has an ambiguous effect of the reduction of the activity. If compliance and/or enforcement is high then a decrease in y^* diminishes the aggregate activity.*

As a consequence of Results 2 and 3 it is important to identify the goal of the regulator. If the regulator is interested in reducing the aggregate activity lowering the acceptable amount of activity is effective only when it is coupled with strict enforcement. If few are complying then it is likely that the number of students who refuse to comply and increase their activity will exceed the number reducing their activity. In contrast, if the goal of the regulator is to increase compliance the lowering of the amount tolerated works against this goal.

WELFARE

As discussed in the previous section the regulator may instead be interested in the total well-being of the population of students. With the imperfect monitoring of students' activity the socially optimal outcome cannot be achieved, but given the constraints on the regulator's information, what policies can the regulator use to most enhance the well-being of the students?

There are two variables which the regulator might have control over: the penalty and the acceptable level of activity. For example, consider a college attempting to regulate alcohol and drug use on campus. It is able to set a policy that outlines the acceptable and unacceptable amounts of each activity. Violations of the standard may be difficult to observe but once the violation is caught assessing whether or not it exceeded the acceptable level should be straightforward. The size of the sanction for the violation is also within the control of the college.

Consider, first, a change in the size of the punishment, *f*, handed down to a student caught exceeding the standard. An increase in *f* affects the well-being of the population in the same two ways as mentioned before. First, the increase causes some students to alter their activity, which reduces their utility. Students continuing to exceed the tolerated amount reduce their activity due to the greater potential cost while others now decide to comply with the standard. Secondly, from Result 2, the increase in *f* reduces *Y*. This increases the well-being of all students.

What, then, can be said about the optimal penalty that maximizes the total well-being of the population? First, the optimal penalty is increasing in both the size of the population, *N*, and the magnitude of the disutility from aggregate activity, *b*. Both *N* and *b* affect the magnitude of the negative externality; either by affecting more individuals or by having a more severe impact on each person. By controlling the penalty the regulator is able to control the expected cost to a student engaging in excess activity. An increase in the probability of being caught also increases the expected cost. Since the optimal penalty sets the marginal benefit equal to the marginal cost the penalty is less when the probability of apprehension is greater.

Result 4: *The larger the population of students or the more harmful the activity (greater N or b) the higher the optimal penalty, while a greater probability of apprehension (p) reduces it.*

Next, consider a decrease in the amount allowed, y^*. The optimal standard balances the benefit from the reduced aggregate activity felt by the entire student population with the lost utility of those meeting or exceeding the standard. As before, a greater value of *N* or *b* increases the benefit of a lower acceptable amount of activity.

Result 5: *The larger the population of students or the more harmful the activity (greater N or b) the lower the optimal standard.*

Furthermore, the optimal level of acceptable activity is always greater than zero. Because of the inability to perfectly monitor the students' activities every student, who generates a positive utility from the activity, would violate the policy and be subject to the potential punishment. An increase in the acceptable level, which from Result 3 would increase the amount of activity, would allow students to engage in small amounts of it without punishment. Such an increase would also increase compliance and would result in the students caught exceeding the rule being the ones who engage in the most activity.

Result 6: *A policy of abstinence is never best.*

It should be pointed out that by using the functional form given in (1) it is assumed that for every student with $a_i > 0$ the marginal benefit of engaging in the activity is infinitely large if she does not engage in the activity. If this assumption is relaxed so that the marginal benefit is finite, for large populations and sufficiently harmful substances it may then be best to prohibit the activity (see Miron and Zwiebel (1995) and Miron (1998) for detailed discussions of prohibition of alcohol and drugs).[5] Furthermore, the only harm the activity is assumed to cause is the externality generated by the aggregated activity. This setup does not include direct harm that the activity might cause, such as sexual assault.

IV. CONCLUSION

The goal of this paper is to address effective college alcohol and drug policy. A theoretical model of the effects and use of alcohol and drugs as an activity where there is a personal preference for the activity, there is variation in the intensity of this preference across the student body, and the aggregate quantity of the activity has harmful externalities is developed. It should be pointed out that the activity is interpreted as the use of alcohol and drugs on college campuses, but the analysis could be applied to any activity that has these three characteristics.

Effective policy is shown to be determined by the goal of the regulator. Three goals were analyzed. The following table summarizes the results.

Table 1: Policy Recommendations

Goal 1: Reduce Aggregate Activity	
1. Regulate activity	Result 1
2. Increase sanction	Result 2
3. Catch violations with a greater probability	Result 2
4. Decrease acceptable amount of activity if compliance is high and enforcement is effective, otherwise increase the tolerated level	Result 3
Goal 2: Increase Compliance	
1. Regulate activity	Result 1
2. Increase sanction	Result 2
3. Catch violations with a greater probability	Result 2
4. Increase acceptable amount of activity	Result 3
Goal 3: Maximize Student Well-Being	
1. Regulate activity	Result 1
2. Larger populations need a stiffer sanction and a lower acceptable standard	Results 4 & 5
3. Detrimental activities need a stiffer sanction and a lower acceptable standard	Results 4 & 5
4. Stiffer sanction with a lower probability of catching violations	Result 4
5. No abstinence policy	Result 6

This paper can explain much of the variation in college alcohol and drug policies. Larger schools need stronger sanctions and lower acceptable levels of activity. More harmful activities require lower amounts of acceptable activity and stiffer penalties. Thus, separate policies for alcohol, possession of illegal drugs, and distribution of illegal drugs are in order. It also calls for policies to discriminate based on the size of the externality. For example, off-campus activity might be punished differently than on-campus violations. Furthermore, the abandonment of policies of abstinence is in accordance with this goal. Consequently, colleges with abstinence policies or those that do not differentiate between activities must have some other goal besides improving the well-being of the student population.

These results fit reasonably well with the limited empirical evidence. Mitchell, Toomey, and Erickson (2005) survey college alcohol policies. They report that larger schools more frequently prohibit alcohol use in Greek houses, ban beer kegs, and provide on-campus, alcohol-free activities than smaller schools. Furthermore, colleges with a higher percentage of students living on campus more frequently prohibit the use, possession, advertising, and sale of alcohol. Thus, their results support the prediction that larger populations with spillovers affecting more people need lower levels of acceptable activity

As with any theoretical model assumptions are made to simplify the analysis and focus the discussion on the important aspects of the interaction. The work could be expanded to take into consideration many factors not included in the analysis. First, student preferences are assumed to be static. Many initiatives in place on college campuses specifically attempt to influence student's preferences. For example, marketing campaigns are used to affect the social norms of drinking and drug use. (Mitchell, Toomey, and Erickson, 2005). Kacapyr and Choudhury (2006) illustrate, though, that student perception of drinking on a college campus does not significantly affect the quantity of drinks consumed or the prevalence of binge drinking. Another important aspect to regulating alcohol and drug use on campus is the costs of regulation. Monitoring student activity is time consuming and costly. Also, overregulation by the college reduces the students' quality of life, which may have detrimental impacts on students not violating the policy. For example, room searches affect not only the accused, but also those living with the accused. Thirdly, the model looks at three potential goals of the regulator. Many more goals are possible. For example, a regulator interested in the well-being of the student body prior to penalties is considered, but not the well-being net of the penalties. Presumably, optimal penalties would be lessened and the optimal level of acceptable activity would be higher. Other policy goals include harm reduction, reducing legal liability, and expenditure reduction. None of these are studied in this model. Additionally, enforcement is modeled simply as a sanction and a probability of apprehension. The work does not differentiate between various forms of sanctions (monetary fines and restrictions on student life activities are two examples of potential punishments) or modes of enforcement (investigations based on reports and the monitoring of activities are two examples). A useful avenue would be to identify the effective forms of enforcement. Finally, the environment might be expanded to allow for interaction over time. The punishment of multiple offenses is not modeled and distinguishing between users and experimenters is not done. Also, the probability of being caught is assumed to be a function of the quantity of activity. A more elaborate model explicitly modeling the oversight by the college might add some important results.

ENDNOTES

1. I would like to thank Jonathan Caulkins, Maureen Donohue-Smith, Amihai Glazer, Patrick Meister, and Jim Mullen for their helpful comments. Suggestions made by William P. O'Dea and an anonymous referee also proved quite valuable.
2. An example of relegating oversight to legal authorities see the stated policy of the University of Kansas (www.vpss.ku.edu/alcoholbrochure.pdf). Ripon College is an example of a school laying out separate policies for many substances (www.ripon.edu/administration/Plant/ drugs.htm). Elmira College suspends students for a first time violation of its drug policy (www.elmira.edu/pdfs/campuslife/0405handbook.pdf) while Monmouth University uses fines and probation for a first offense (www.monmouth.edu/student/ policies/student.asp). Finally, Houghton College bans alcohol on campus (campus.houghton.edu/orgs/student_life/ comm_res_policies.htm#Drugs%20and%20Alcohol). It should be emphasized that these schools are just examples; for each there exists multiple examples.
3. From (1) $\sum_{i=1}^{N} u(y_i, Y) = \sum_{i=1}^{N} a_i \ln y_i - \sum_{i=1}^{N} bY = \sum_{i=1}^{N} a_i \ln y_i - NbY$. Since $Y = \sum_{i=1}^{N} y_i$ this simplifies to $\sum_{i=1}^{N} (a_i \ln y_i - Nby_i)$. Therefore, setting the derivative of this expression equal to zero yields equation (3).
4. If Y is observable an example of another mechanism would be to target an optimal level of Y and indiscriminately punish all students if it is not achieved.
5. If the functional form is altered to have a finite marginal cost a policy mandating y^* = 0 would be socially optimal if $\partial u_i / \partial y_i \big|_{y_{8i}=0} \in (b, Nb)$.

REFERENCES

Bebchuk, Lucian Arye, and Louis Kaplow. 1992. "Optimal Sanctions When Individuals are Imperfectly Informed About the Probability of Apprehension." *Journal of Legal Studies* 21 (June): 265-70.

Bebchuk, Lucian Arye, and Louis Kaplow. 1993. "Optimal Sanctions and Differences in Individuals' Likelihood of Avoiding Detection." *International Review of Law and Economics* 13 (March): 217-24.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (March): 169-217.

Benson, Bruce L., and David W. Rasmussen. 1998. "The Context of Drug Policy: An Economic Interpretation." *Journal of Drug Issues* 25 (Summer): 681-700.

Castro, Ralph J., and Betsy D. Foy. 2002. "Harm Reduction: A Promising Approach." *Journal of American College Health* 51 (September): 89-91.

Caulkins, Jonathan P., and Peter Reuter. 1995. "Redefining the Goals of National Drug Policy: Recommendations From a Working Group." *American Journal of Public Health* 85 (August): 1059-62.

Caulkins, Jonathan P., and Peter Reuter. 1997. "Setting Goals for Drug Policy: Harm Reduction or Use Reduction?" *Addiction* 92 (September): 1143-50.

Communications from the Dean. 17 March 2004. Carleton College. webapps.acs.carleton.edu/campus/dos/communications/?story_id=44480.

Craswell, Richard, and John E. Calfee. 1986. "Deterrence and Uncertain Legal Standards." *Journal of Law, Economics, and Organization* 2 (Fall): 279-303.

Friedman, David, and William Sjostrom. 1993. "Hanged for a Sheep: The Economics of Marginal Deterrence." *Journal of Legal Studies* 22 (June): 345-66.

Garoupa, Nuno. 1997. "The Theory of Optimal Law Enforcement." *Journal of Economic Surveys* 11 September): 267-95.

Handbook of Community Standards. 2004. Hobart and William Smith Colleges. www.hws.edu/pdf/communitystandards.pdf.

Kacapyr, Elia, and Samira Choudhury. 2006. “Determinants of Alcohol Consumption By College Students.” *New York Economic Review* 37 (Fall): 3-19.

Kolstad, Charles D., Thomas S. Ulen, and Gary V. Johnson. 1990. "Ex Post Liability for Harm vs. Ex Ante Safety Regulation: Substitutes or Complements?" *American Economic Review* 80 (September): 888-901.

Krebs, Christopher P., Michael Costelloe, and David Jenks. 2003. "Drug Control Policy and Smuggling Innovation: A Game-Theoretic Analysis." *Journal of Drug Issues* 33 (Winter): 134-60.

Malik, Arun S. 1990. "Avoidance, Screening and Optimal Enforcement." *RAND Journal of Economics* 21 (Autumn): 341-53.

Miron, Jeffrey A. 1998. "An Economic Analysis of Alcohol Prohibition." *Journal of Drug Issues* 28 (Summer): 741-62.

Miron, Jeffrey A., and Jeffrey Zwiebel. 1995. "The Economic Case Against Drug Prohibition." *Journal of Economic Perspectives* 9 (Fall): 175-92.

Mitchell, Rebecca J., Traci L. Toomey, and Darin Erickson. 2005. "Alcohol Policy on College Campuses." *Journal of American College Health* 53 (January/February): 149-57.

Mookherjee, Dilip, and I. P. L. Png. 1992. "Monitoring vis-à-vis Investigation in Enforcement of Law." *American Economic Review* 82 (June): 556-65.

Polinsky, A. Mitchell, and Steven Shavell. 1991. "A Note on Optimal Fines When Wealth Varies Among Individuals." *American Economic Review* 81 (June): 618-21.

Polinsky, A. Mitchell, and Steven Shavell. 1992. "Enforcement Costs and the Optimal Magnitude and Probability of Fines." *Journal of Law and Economics* 35 (April): 133-48.

Polinsky, A. Mitchell, and Steven Shavell. 2000. "The Economic Theory of Public Enforcement of Law." *Journal of Economic Literature* 38 (March): 45-76.

Shavell, Steven. 1987. "The Optimal Use of Nonmonetary Sanctions as a Deterrent." *American Economic Review* 77 (September): 584-92.

Shavell, Steven. 1991. "Specific versus General Enforcement of Law." *Journal of Political Economy* 99 (October): 1088-108.

Shavell, Steven. 1992. "A Note on Marginal Deterrence." *International Review of Law and Economics* 12 (September): 345-55.

Teter, Christian J., Sean Esteban McCabe, James A. Cranford, Carol J. Boyd, and Sally K. Guthrie. 2005. "Prevalence and Motives for Illicit Use of Prescription Stimulants in an Undergraduate Student Sample." *American Journal of College Health* 53 (May/June): 253-62.

Wechsler, Henry, Kathleen Kelley, Elissa R. Weitzman, John Paul San Giovanni, and Mark Seibring. 2000a. "What Colleges are Doing About Student Binge Drinking: A Survey of College Administrators." *Journal of American College Health* 48 (March): 219-26.

Wechsler, Henry, Jae Eun Lee, Meichun Kuo, Hang Lee. 2000b. "College Binge Drinking in the 1990s: A Continuing Problem; Results of the Harvard School of Public Health 1999 College Alcohol Survey." *Journal of American College Health* 48 (March): 199-210.

Wechsler, Henry, Mark Seibring, I-Chao Liu, and Marilyn Ahl. 2004. "Colleges Respond to Student Binge Drinking: Reducing Student Demand or Limiting Access. *Journal of American College Health* 52 (January/February): 159-68.

Wilde, Louis. 1992. "Criminal Choice, Nonmonetary Sanctions, and Marginal Deterrence." *International Review of Law and Economics* 12 (September): 333-44.

Wittman, Donald. 1977. "Prior Regulation versus Post liability: The Choice Between Input and Output Monitoring." *Journal of Legal Studies* 6 (January): 193-211.

APPENDIX

The goal of this appendix is to formally prove the results presented in the text. First is the derivation of $\overline{a}(y*)$ and $\underline{a}(y*)$. Without regulation a student selects $y_i = a_i / b$. When this is less than $y*$ it is preferred if

(5) $a_i \ln(a_i / b) - b(Y_{-i} + a_i / b) \geq a_i \ln y* - b(Y_{-i} + y*)$

where Y_{-i} is the total activity of all students except i. Define $\underline{a}(y*)$ as the value of a_i that equates (5). It follows immediately that

(6) $\underline{a}(y*) = y*/b$.

If a student violates the policy and is subject to a potential punishment a selection of $y_i = a_i / (b + pf)$ maximizes her expected utility. This is preferred to exactly complying with the policy if

(7) $a_i \ln(a_i /(b + pf)) - b[Y_{-i} + a_i /(b + pf)] - pf[a_i /(b + pf)] \geq a_i \ln y* - b(Y_{-i} + y*)$, or rather, if

$$a_i \ln(a_i /(b + pf)) - a_i \geq a_i \ln y* - by* \; .$$

Define $\overline{a}(y*)$ as the value of a_i that equates (7). Notice that the left-hand-side of (7) increases at a rate of ln (a_i / (b + pf)) + 1 when a_i increases while the right-hand-side increases at a rate of ln $y*$. Since the threshold a_i / (b + pf) is of interest only when it is greater than $y*$ the left-hand-side of (7) increases at a faster rate. Thus, if either p or f increases then $\overline{a}$ must increase as well to maintain the equality. Also, if $y*$ increases then the right-hand-side of (7) increases as well since the condition matters only if $y*$ is below a / b. Since, as established, the left-hand-side increases with a_i at a faster rate, if $y*$ increases so to does $\overline{a}$ to maintain the equality.

Furthermore, $\overline{a}$ requires that $a_i / (b + pf) \geq y*$, or rather, that $a_i \geq (b + pf)y*$, while $\underline{a}$ requires that $a_i / b \leq y*$, or rather, $a_i \leq by*$. Therefore, $\overline{a} \geq (b + pf)y* > by* \geq \underline{a}$ so that an interval where $y_i = y*$ is the expected utility maximizing choice exists. It follows immediately that since both $y*$ and b are nonnegative there exists an interval [0, $\underline{a}(y*)$] where $y_i = a_i / b$ is the expected utility maximizing selection. Finally, there exist students with values of a_i where the choice $y_i = a_i / (b + pf)$ is the optimal selection when $\overline{a}(y^+) \leq a^+$. If the upper bounds satisfy (7) when $a_i = a^+$ and $y_i = y^+$ this inequality holds and students make such a choice.

Now consider the results presented in the text.

Proof of Result 1: R is the sum of the difference between a_i / b and $y*$ for students with a_i in [$\underline{a}(y*), \overline{a}(y*)$] and $y*$ and a_i /(b + pf)) for students with $a_i > \overline{a}(y*)$. Since $a_i / b \geq y*$ for $a_i \geq \underline{a}(y*)$ and $pf > 0$, $R > 0$.

Proof of Result 2: As shown previously, $d\overline{a}/dp > 0$ and $d\overline{a}/df > 0$. Since students with $a_i \leq \overline{a}(y*)$ are complying with the policy then an increase in p or f increases compliance. For students with $a_i > \overline{a}(y*)$ an increase in p or f increases the difference between a_i / (b + pf)) and a_i / b.

Furthermore, the increase in $\overline{a}(y*)$ shifts some from $y_i = a_i/(b + pf)$ to $y_i = y*$, which is a reduction in their activity. Thus, R increases.

Proof of Result 3: As previously shown, $d\overline{a}/dy* > 0$. Hence, a decrease in $y*$ decreases compliance. A decrease in $y*$ has three effects on R. It increases the difference between a_i / b and $y*$ for students with a_i in $[\underline{a}(y*), \overline{a}(y*)]$, increasing R. Second, since $d\overline{a}/dy* > 0$, some students shift from $y_i = y*$ to $y_i = a_i / (b + pf)$, which decreases R. Finally, since $d\underline{a}/dy* > 0$, other students shift from $y_i = a_i / b$ to $y_i = y*$, which increases R. Thus, a decrease in $y*$ reduces activity if the number of students between $\overline{a}(y*)$ and $\underline{a}(y*)$ is large. Also, such an adjustment reduces activity if the jump from $y*$ to $a_i / (b + pf)$ for those near $\overline{a}(y*)$ is small, which occurs when pf is larger.

Proof of Result 4: Consider an increase in the penalty f. Students with $a_i > \overline{a}(y*)$ decrease their activity. Since these students select $a_i / (b + pf)$ the rate of decrease is $pa_i / (b + pf)^2$. The regulator selects f at the point where the marginal benefit equals the marginal cost. The benefit is the gain in utility to all students when the aggregate activity is reduced. Since the change in Y with an increase in f is $\sum_{i \in E} pa_i /(b + pf)^2 df$ where E is the set of students exceeding $y*$, the marginal benefit is $Nb\sum_{i \in E} pa_i /(b + pf)^2 df$. The cost is the lost utility to those who reduce their activity. Therefore, the marginal cost is $\sum_{i \in E}(a_i / y_i)pa_i /(b + pf)^2 df = \sum_{i \in E} pa_i /(b + pf)df$. The optimal penalty, $f*$, is the one that satisfies

(8) $Nb\sum_{i \in E} pa_i /(b + pf)^2 df = \sum_{i \in E} pa_i /(b + pf)df$.

This simplifies to

(9) $f* = b(N - 1) / p$.

Consequently, $df* / dN = b / p > 0$, $df* / db = (N - 1)/ p > 0$, and $df* / dp = -b(N - 1) / p^2 < 0$. Hence, Result 4 holds.

Proof of Results 5 and 6: For a small decrease in $y*$ the students exactly complying with the standard reduce their activity. The optimal penalty sets the marginal benefit equal to the marginal cost. The benefit is the gain in utility to all students when the aggregate activity is reduced. Thus, the marginal benefit is $Nb\sum_{i \in C} dy^*$ where C is the set of students exactly complying with the standard. The cost is the lost utility to those in C. The marginal cost is $\sum_{i \in C} a_i / y^* dy^*$. The optimal level of acceptable activity, y**, is the one that equates the two. This simplifies to

(11) $y^{**} = \sum_{i \in C} a_i \;/ Nb$.

Since $y^{**} > 0$, Results 5 and 6 hold.

The Impact of Welfare Reform on the Employment and Labor Supply of Female High School Dropouts

Jeffrey T. Lewis*

I. INTRODUCTION

Welfare participation began to decline dramatically and low-skill female employment began to rise substantially in the United States during the 1990's. Two competing explanations for these developments are the strong economy and welfare reform. The U.S. experienced a remarkable economic boom during the 1990's. From 1992 to 2000, the unemployment rate fell steadily from 7.5 percent to 4.0 percent, and total non-farm employment grew by over 20 percent from 109 million to 132 million.[1]

The 1990's were also a period of unprecedented welfare reform in the U.S. The policy of liberally granting waivers to the states so that they could experiment with reform was adopted by the first Bush administration and later continued by the Clinton administration. From 1992-1996, 29 states ultimately implemented waivers. States that were granted waivers typically pursued a bundle of reform policies. A 1999 Council of Economic Advisors study classifies six types of waivers as "major": termination time limits, work requirement time limits, family caps, strict work exemptions, sanctions, and liberalized earnings disregards. Under termination time limits, after the time limit is reached, welfare recipients' benefits are cut off. Under work requirement time limits, on the other hand, after the time limit is reached, welfare recipients must work in order to continue receiving benefits. Under family caps, benefits are not increased when a women has an additional child while on welfare. Strict work exemptions exempt only single mothers with very young children from having to participate in work activities.[2] Sanctions are levied when welfare recipients violate their work requirements. Liberalized earnings disregards allow women to retain more of their earnings while on welfare.

Time limits, work requirements, strict work exemptions, and sanctions are expected to decrease welfare use and increase employment. Although the primary aim of family caps is to discourage welfare beneficiaries from having more children while receiving public assistance, family caps could reduce welfare participation by making life on welfare less attractive. Liberalized earnings disregards, by increasing the return to working while on welfare, encourage both employment and welfare participation.

*Cornell University, Department of Policy Analysis and Management, Ithaca, NY 14853, 607-254-5496, jtl28@cornell.edu

President Clinton signed into law the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) in August 1996. The federal guarantee of cash assistance to the poor was abolished and replaced by a program called Temporary Assistance for Needy Families (TANF), which empowered states to run their own welfare programs with block grants from the federal government. Under TANF, a five-year lifetime time limit and a two-year continuous-use time limit on benefits were established, and states were allowed to implement shorter versions of each type of time limit if they so desired. States were, however, also permitted to exempt up to 20 percent of their caseloads from the lifetime time limit. Work participation requirements were established by the federal government, and states that failed to meet their targets faced a reduction in block grant funds. The federal government also gave states the power to impose family caps and regulate earnings disregards and asset limits. Like the state waiver programs, TANF was expected to reduce caseloads and boost employment.

Numerous papers that attempt to econometrically disentangle the effects of the strong economy and welfare reform on increases in female employment have been written. These papers are marked somewhat by a lack of consensus. Summarizing the findings in the literature, Blank (2001) notes, “Both the economy and waivers appear to have raised employment in the early 1990s; studies that look at the effect of TANF on employment in the late 1990s show more mixed results” (32).

Using March Current Population Survey (CPS) and CPS Outgoing Rotation Group (CPS-ORG) data, Meyer and Rosenbaum (2001) examine the impact of numerous policy variables on the employment of single women over the 1984-1996 period. The authors construct two welfare reform variables. One dummy is turned on when a time limit waiver is implemented, and a second is turned on once any welfare case has been terminated in a state under a waiver. Meyer and Rosenbaum find that both the fraction of single female high school dropouts who were employed last week, and the fraction who were employed last year, rose due to both time limit waivers and the onset of terminations. The authors do not examine the effects of TANF on employment.

Theorizing that families with younger youngest children should be less likely to use welfare than families with older youngest children under time limits since families with younger children face a longer time horizon over which they might experience an adverse economic event and, therefore, have a greater need to stockpile their benefits, Grogger (2003) investigates how the effects of welfare reform on outcomes for female heads of family vary depending on the age of the youngest child. Using March CPS data and examining the 1978-1999 period, Grogger focuses on the effects of time limits. He uses one dummy that is turned on if a time limit is in effect under either waivers or TANF, and another dummy that is turned on if any reform is in effect under either waivers or TANF. Grogger finds that the employment of female family heads rose because of both time limits and other reforms, and that the employment gains were largest for families with younger youngest children in both cases. He also finds that reforms other than time limits resulted in an increase in weeks worked by female family heads.

Using March CPS data, Fang and Keane (2004) examine the impact of reform on the employment of single mothers over the 1993-2002 period by, similar to Grogger (2003), exploiting the demographic dimensions of policy variation from different welfare program rules. Commenting on their paper, Grogger (2004a) contends that "Fang and Keane push [the] demographic dimensions of policy variation further than any previous study" (104). The authors find that both time limits and work requirements contributed to an increase in the employment of single mothers, and, consistent with Grogger (2003), they estimate larger impacts for families with younger children.

Examining the 1995-1999 period with the March CPS surveys, Kaushal and Kaestner (2001) employ a difference-in-difference strategy to determine the effects of time limits and family caps on work. They use one dummy that is turned on if a time limit is in effect under either waivers or TANF, and one dummy that is turned on if a family cap is in effect under either waivers or TANF. The authors find that both time limits and family caps caused the employment and labor supply of unmarried female high school dropouts with children to rise relative to married female high school dropouts with children.

Kaestner and Kaushal (2005) also employ a difference-in-difference strategy using 1994-1999 CPS-ORG data to study the effects of reform on work. In this paper, the authors use separate dummies for waivers and TANF, although they only report their TANF results. They find that the employment of unmarried female high school dropouts rose relative to married female high school dropouts under TANF. Their results for usual hours worked are statistically insignificant.

Using the March CPS surveys to examine the 1983-2000 period, O'Neill and Hill (2001) find that the employment of single mothers rose under both waivers and TANF. Blank (2002) points out, however, that the "lack of time fixed effects in the O'Neill and Hill study almost surely results in a larger coefficient on the TANF dummy variable than in other studies" (1140).

Moffitt (1999) uses the March CPS surveys from 1977-1995 to determine the effects of waivers on labor supply. Unlike the above-mentioned authors, Moffitt does not restrict his sample to some group of unmarried women or single mothers. He finds that under waivers there was an increase in annual weeks and hours worked by female high school dropouts, and an increase in annual hours worked by female high school graduates. Moffitt does not investigate the impact of TANF on labor supply.

Also not restricting their sample to some group of unmarried women or single mothers, Schoeni and Blank (2000) examine the 1976-1998 period with the March CPS surveys to determine the effects of reform on employment and labor supply. Because they have more years of data, the authors, unlike Moffitt, analyze the effects of both waivers and TANF. They find that, under waivers, employment and labor supply increased for female high school dropouts, but not for women with either a high school degree or some college experience. Interestingly, Schoeni and Blank find no statistically significant employment or labor supply increases for any group under TANF.[3] Since marital status is likely endogenous with welfare reform, it is noteworthy that the only paper using post-1996 data that does not estimate TANF policy effects is the only one that does not restrict its sample to some group of unmarried women or single mothers.

While welfare reform could have been expected to affect employment, it also could have been expected to affect marriage. Indeed, using vital statistics data, which contain a near-universe of new marriages and divorces, Bitler et al. (2004) estimate that reform had a sizable impact on marriage and divorce. The authors find that marriage rates fell by about 5 percent under waivers and 20 percent under TANF, and that divorce rates fell by about 5 percent under waivers and 10 percent under TANF. Further, it should be noted that welfare reform plausibly could have affected the marital status of women of different skill levels differently. Less-skilled women, particularly burdened by the stringent elements of reform, might have become more likely to marry or remain married under reform. On the other hand, more-skilled women, better able to take advantage of the elements of reform that increased the return to working, might have become less likely to marry or remain married under reform.

Since welfare reform could have caused the pool of unmarried women to become more skilled, relying on a sample of some group of unmarried women or single mothers to determine the effects of reform on employment could produce upward-biased estimates. This is why it is a concern that, as noted above, the only paper using post-1996 data that does not find that TANF caused women to work more is the only one that does not restrict its sample to some group of unmarried women or single mothers. In this paper, however, using pooled cross-sectional data from the 1989-2004 CPS-ORG surveys and employing a sample of female high school dropouts that is not restricted to unmarried women or single mothers, I do find that TANF is associated with an increase in both employment and labor supply. The results in this paper, then, strengthen the case in the literature that it was not just the strong economy but also federal welfare reform that contributed to the work gains of low-skill women in the post-1996 period.

Here is how this paper proceeds. In Section II, I describe my data. In Section III, I discuss my model specifications. In Section IV, I present my results first for regressions run on all women, pooling over race and ethnicity, and then for regressions run separately by race or ethnicity. Lastly, I conclude in Section V.

II. DATA

Numerous papers that examine the impact of welfare reform on a variety of different outcomes use March CPS data. When examining the impact of reform on employment, however, using CPS-ORG data might be preferable. The CPS is a nationally representative survey of approximately 60,000 households that is administered every month. The appeal of the March CPS surveys is that individuals are asked detailed questions about their employment, labor supply, and earnings during the previous calendar year. In the CPS-ORG surveys, individuals are asked questions about their employment, labor supply, and earnings only during the previous week.

The principal advantage of using CPS-ORG data instead of March CPS data is sample size. In the CPS, a household is interviewed for four consecutive months, not interviewed for eight months, and then interviewed for another four consecutive months. The CPS-ORG consists of observations from households in either their fourth or eighth month of being interviewed. Every month, then, one-fourth of the CPS observations are included in the CPS-ORG. The March CPS includes observations from every household that is interviewed in March. Hence, because CPS-ORG data are collected every month of the year, while March CPS data are collected only once a year, the CPS-ORG is approximately three times the size of the March CPS.

Sample size is particularly important when examining the impact of welfare reform on employment because the groups that are most likely to be affected by reform are low-education minority groups with the highest welfare participation rates. Table 1 displays, for various groups of women, the percentage that lived in a household that received any AFDC income during the year.[4] The table demonstrates that there is much heterogeneity in welfare use among racial and ethnic groups with different levels of educational attainment. Since black dropouts are highly likely to be affected by reform since such a large percentage of them live in households receiving welfare income (31.4 percent), one would want to look at a sample of black dropouts when investigating the impact of reform on employment. A concern is that the sample mean for black dropouts in a given state-year cell will be a good estimate of the true mean for black dropouts in that state-year only if the sample size is sufficiently large. Since the CPS-ORG is three times as large as the March CPS, the sample mean for black dropouts is more likely to be representative of the group's true mean when CPS-ORG data are used.

Many researchers have pointed out that because all states implemented TANF within a seventeen-month period, identifying the effects of TANF is difficult.[5] Respondents are asked about their employment status last week in both the March CPS surveys and the CPS-ORG surveys. Using the March CPS surveys, in a given state-year in which reform is implemented, all women will have been interviewed in that state-year either before or after the onset of reform. Using the CPS-ORG surveys, on the other hand, in a given state-year in which reform is implemented, some women will have been interviewed before the onset of reform, and some afterwards. One, then, might be better able to identify the effects of TANF using CPS-ORG data as opposed to March CPS data.

III. MODEL SPECIFICATIONS

Using the CPS-ORG surveys from 1989-2004, I examine the impact of welfare reform on the employment and labor supply of women.[6] My empirical approach is similar to that of Schoeni and Blank (2000). The first model specification I use is:

(1) $y_{ist} = W_{st}*\beta_W + T_{st}*\beta_T + X_{ist}*\beta_X + L_{st}*\beta_L + \gamma_s + v_t + \delta_m + trend*\gamma_s + \varepsilon_{ist}$

Table 1- Sample sizes and percentages receiving AFDC income

	Sample size for employment regressions (1989-2004)	Sample size for hours worked regressions (1989-2004)	Percentage who had AFDC income during the year (1989-1992)
All women, <HS	217,099	217,097	15.7%
All women, =HS	576,580	576,560	5.2%
All women, >HS	867,678	867,612	1.9%
Black women, <HS	34,109	34,109	31.4%
Black women, =HS	70,519	70,517	15.9%
Black women, >HS	81,146	81,144	6.5%
Hispanic women, <HS	60,284	60,284	13.9%
Hispanic women, =HS	46,246	46,246	7.5%
Hispanic women, >HS	44,495	44,495	3.6%
White women, <HS	110,072	110,070	10.8%
White women, =HS	435,426	435,408	3.1%
White women, >HS	693,584	693,522	1.2%

The black samples consist of non-Hispanic black women and the white samples consist of non-Hispanic white women. The employment and hours worked samples include women aged 16-54 who are not enrolled in school. Women whose CPS-ORG weight or employment status is missing are excluded from both samples. Women who were working last week whose hours worked last week are missing are also excluded from the hours worked samples. The last column gives the percentage of non-enrolled women aged 16-54 who lived in a household that received AFDC income during the previous year. The data, which come from the 1990-1993 March CPS surveys, refer to calendar years 1989-1992. The percentages are weighted by the March CPS supplemental weight.

The variables are defined for women *i* who lives in state *s* in year *t*:

y: outcome of interest (employed last week[7] or hours worked last week).

W: a dummy that equals one if a waiver is in effect.

T: a dummy that equals one if TANF is in effect.[8]

X: demographic variables (age, age-squared, non-Hispanic black dummy, Hispanic dummy).

L: state-level labor market variables (current and lagged average annual state unemployment rate, current and lagged annual state employment growth rate, log real maximum AFDC benefit for a family of three).

γ_s: state fixed effects.

v_t: year fixed effects.

δ_m: month fixed effects.

$trend*\gamma_s$: linear state-specific time trends.

I run separate regressions for three different education groups (less than 12 years, 12 years, more than 12 years), thereby allowing each right-hand-side variable to vary by education group.[9] Robust standard errors corrected for clustering in state-year cells are used. Regressions and means are weighted by the CPS-ORG weight variable.

Welfare reform was implemented while the U.S. economy was booming. The numerous state-level labor market variables that are used (the unemployment rate and its lag and the employment growth rate and its lag) are designed to control for the effects of the strong economy on female employment. The inclusion of year effects, which control for unobservable factors that affect employment that vary over time, but not across states, is important for a number of reasons. First, if the state-level labor market variables that are used inadequately describe the economic expansion of the 1990's, then the year effects can help to soak up some of the effects of the nationwide economic boom on female employment. Second, year effects can also absorb the effects of policy changes besides welfare reform that were implemented nationally in the 1990's that could also be expected to boost female employment, such as the expansions of the EITC and the increases in the federal minimum wage. For both of these reasons, the exclusion of year effects from Model (1) would likely cause the effect of welfare reform on female employment to be significantly overstated. Finally, state effects control for unobservable factors affecting employment that vary across states but not over time, linear state-specific time trends control for unobservable factors that affect employment that vary over time within states, and month effects control for seasonal factors that could affect employment.

With the inclusion of all of the above-mentioned controls, the coefficients on the waiver and TANF dummies should estimate the causal impact of each reform regime on employment. In a given state, the waiver dummy is turned off once TANF is implemented. Thus, the coefficient on the waiver dummy estimates the effect of waivers relative to the old AFDC system, and the coefficient on the TANF dummy also estimates the effect of TANF relative to the old AFDC system.

Table 1 indicates that, before the onset of welfare reform, 15.7 percent of female high school dropouts lived in a household that received AFDC income during the year, compared to 5.2 percent of female high school graduates, and 1.9 percent of females with some college experience. Since the low-education group is the one most likely to be affected by welfare legislation, finding the effects of reform to be concentrated among high school dropouts would lend credibility to the notion that policy effects are truly captured. As noted previously, Table 1 also suggests that running regressions separately by race or ethnicity is important. Hence, besides examining the effects of reform on all female high school dropouts, all high school graduates, and all females with some college experience, I also examine the effects of reform on black, Hispanic, and white women of the three different education levels.[10]

Bitler et al. (2002) note that the effects of TANF in waiver states and non-waiver states might have been different. States that had already implemented a waiver might have made few adjustments to their welfare system under TANF, whereas states that had never implemented a waiver might have made significant changes under TANF, which might have resulted in a bigger employment increase under TANF in non-waiver states than in waiver states. On the other hand, states that had already had experience with waivers might have implemented more far-reaching reform measures under TANF, which might have caused the employment increase under TANF to be larger in waiver states than in non-waiver states.

Accordingly, the second model specification I use is:

(2) $y_{ist} = W_{st}*\beta_W + TNOW_{st}*\beta_{TNOW} + THADW_{st}*\beta_{THADW} + X_{ist}*\beta_X + L_{st}*\beta_L + \gamma_s + v_t + \delta_m + trend*\gamma_s + \varepsilon_{ist}$

The new reform variables are:

TNOW: a dummy that equals one if TANF is in effect in a state that had never implemented a waiver.

THADW: a dummy that equals one if TANF is in effect in a state that had previously implemented a waiver.

Again, the waiver dummy is turned off once a state implements TANF. Hence, β_{TNOW} measures the impact of TANF in states that had never implemented a waiver relative to the old AFDC system, and β_{THADW} measures the impact of TANF in states that had previously implemented a waiver relative to the old AFDC system.

While many papers have attempted to determine the impact of specific reform policies on outcomes of interest, numerous researchers have pointed out the difficulty of such an undertaking (Bell, 2001; Blank, 2002; Moffitt, 2002; Bitler et al., 2004; Bitler et al., 2006). First, because data on the details of state policies are limited, researchers typically code only major reform policies. Since states implemented many other reforms, however, the effects of the coded reform policies might be overstated.[11] Second, since there is no way to measure how different states implemented and enforced the same reform policies, this makes it hard to attribute changes in outcomes of interest to specific reform policies.[12]

A research strategy that presents fewer difficulties is to code the intensity of work incentives in each state, and then compare the effects of reform in states that implemented policies with strong work incentives to the effects of reform in states that implemented policies with weaker work incentives. Basing their categorization on information about benefit levels, earnings disregards, time limits, and sanctions, Blank and Schmidt (2001) characterize the work incentives of each state's TANF program as either "strong," "mixed," or "weak."[13] This coding scheme is used by Blank and Schoeni (2003) to examine the effects of TANF on the distribution of children's family income, by Bitler et al. (2004) to examine the effects of TANF on marriage and divorce, and by Bitler et al. (2006) to examine the effects of TANF on children's living arrangements.[14]

Also adopting the coding scheme developed by Blank and Schmidt (2001), I use the following as my third model specification:

(3) $y_{ist} = W_{st}{*}\beta_W + TSTRONG_{st}{*}\beta_{TSTRONG} + TMIXED_{st}{*}\beta_{TMIXED} + TWEAK_{st}{*}\beta_{TWEAK} + X_{ist}{*}\beta_X + L_{st}{*}\beta_L + \gamma_s + v_t + \delta_m + trend{*}\gamma_s + \varepsilon_{ist}$

The new reform variables are:

TSTRONG: a dummy that equals one if a TANF program that is characterized as having strong work incentives is in effect.

TMIXED: a dummy that equals one if a TANF program that is characterized as having mixed work incentives is in effect.

TWEAK: a dummy that equals one if a TANF program that is characterized as having weak work incentives is in effect.

As in the previous two specifications, the waiver dummy is turned off once a state implements TANF. The coefficients $\beta_{TSTRONG}$, β_{TMIXED}, and β_{TWEAK}, then, measure the impact of each particular TANF reform regime relative to the old AFDC system. Finding larger effects on employment under TANF in states with strong work incentives than in states with weak work incentives would strengthen the case that policy effects are in fact captured.[15]

IV. RESULTS

I begin by showing the results of regressions run on samples of all women, pooling over race and ethnicity. The summary statistics for these samples of women appear in Table 2. Results of regressions run using specification (1), in which the effects of TANF are constrained to be the same in states that had previously implemented waivers and in states that had never implemented waivers, appear in Panel A of Table 3. The first two rows of Panel A of Table 3 display the coefficient estimates for the waiver dummy and the TANF dummy for both the employment regression and the hours worked regression run on the sample of female high school dropouts. The analogous coefficient estimates for female high school graduates and females with at least some college education appear in the bottom four rows of Panel A.[16] The pre-reform (1989-1992) means of the dependent variables appear at the bottom of Table 3. Results of regressions run using specification (2), in which the effects of TANF are allowed to vary between waiver states and non-waiver states, appear in Panel B of Table 3. Lastly, results of regressions run using specification (3), in which the effects of TANF are allowed to vary between states that implemented TANF programs that were characterized as having either "strong," "mixed," or "weak" work incentives, appear in Table 4.

Table 3 indicates that while, under waivers, the employment of high school dropouts did not increase, weekly hours worked by that group did rise by 0.32 hours (an increase of 2.3 percent relative to the pre-reform mean of 14.1 hours). Under TANF, on the other hand, dropouts experienced both an employment increase (4.6 percent), and a labor supply increase (5.4 percent, p-value= 0.104). Further, the employment gain for dropouts under TANF was larger in waiver states (5.0 percent) than

Table 2. Summary statistics for employment samples (1989-2004)

	(ed<HS)	(ed=HS)	(ed>HS)
Waiver	0.107 (0.309)	0.091 (0.288)	0.102 (0.302)
TANF	0.463 (0.499)	0.474 (0.499)	0.538 (0.499)
Unemployment Rate	5.78 (1.40)	5.56 (1.39)	5.52 (1.40)
Unemployment Rate (lagged one year)	5.77 (1.44)	5.55 (1.42)	5.52 (1.44)
Employment Growth Rate	1.46 (1.83)	1.40 (1.81)	1.38 (1.81)
Employment Growth Rate (lagged one year)	1.60 (1.88)	1.54 (1.85)	1.49 (1.87)
Log Real Maximum Welfare Benefits	6.08 (0.46)	6.09 (0.42)	6.12 (0.42)
Age	34.23 (11.37)	36.49 (10.03)	37.00 (9.14)

The employment samples include women aged 16-54 who are not enrolled in school. Women whose CPS-ORG weight or employment status is missing are excluded from the samples. The means are weighted by the CPS-ORG weight and the standard deviations appear in parentheses. See the text for a more detailed description of the explanatory variables.

in non-waiver states (4.1 percent), and the increase in hours worked by dropouts under TANF was larger in waiver states (7.7 percent) than in non-waiver states (2.6 percent, p-value= 0.444). These findings suggest that states that had previously implemented waivers might have adopted more comprehensive reform measures under TANF than states that had had no prior experience with reform.[17]

Because the welfare participation rate for women with at least some college experience is so low, that weekly hours worked by this group are found to have risen under TANF is somewhat surprising.[18] The estimated labor supply increase for the high-education group, however, is quite small. For instance, in line with expectations, the increase in hours worked by women with some college experience under TANF in waiver states (1.7 percent) was, in percentage terms, less than one-fourth the size of the increase in hours worked by dropouts under TANF in those same states (7.7 percent).

Table 4 indicates that, under TANF, dropouts experienced the largest employment and labor supply gains in states that implemented TANF programs with strong work incentives. In such states, the employment of dropouts rose by 7.2 percent and hours worked by dropouts increased by 7.3 percent. In states with TANF programs characterized as having only mixed work incentives, the employment of dropouts rose by 4.2 percent and hours worked by dropouts increased by 5.4 percent

Table 3. Results for all women, TANF effects constrained and unconstrained (1989-2004)

Panel A	Employed	Hours Worked
Waiver (ed<HS)	0.0065 (0.0054)	0.323* (0.193)
TANF (ed<HS)	0.0199* (0.0102)	0.759 (0.466)
Waiver (ed=HS)	0.0013 (0.0032)	0.091 (0.137)
TANF (ed=HS)	0.0004 (0.0054)	0.039 (0.219)
Waiver (ed>HS)	0.0014 (0.0025)	0.174 (0.136)
TANF (ed>HS)	0.0026 (0.0035)	0.408** (0.182)

Panel B	Employed	Hours Worked
Waiver (ed<HS)	0.0074 (0.0060)	0.510** (0.218)
TANF (no waiver) (ed<HS)	0.0179* (0.0108)	0.365 (0.476)
TANF (had waiver) (ed<HS)	0.0216* (0.0121)	1.091* (0.564)
Waiver (ed=HS)	-0.0016 (0.0036)	0.081 (0.155)
TANF (no waiver) (ed=HS)	0.0060 (0.0063)	0.059 (0.245)
TANF (had waiver) (ed=HS)	-0.0051 (0.0065)	0.019 (0.262)
Waiver (ed>HS)	0.0011 (0.0028)	0.210 (0.153)
TANF (no waiver) (ed>HS)	0.0033 (0.0041)	0.336 (0.208)
TANF (had waiver) (ed>HS)	0.0021 (0.0041)	0.473** (0.214)

Means (1989-1992)	Employed	Hours Worked
ed<HS	0.433	14.1
ed=HS	0.689	23.9
ed>HS	0.786	27.6

***,**,* indicate statistical significance at the 1%, 5%, and 10% levels respectively. Robust standard errors that are corrected for clustering in state-year cells appear in parentheses. Regressions and means are weighted by the CPS-ORG weight. Results from regressions run using specification (1) appear in Panel A, and results from regressions run using specification (2) appear in Panel B.

Table 4. Results for all women, coding the intensity of TANF (1989-2004)

	Employed	Hours Worked
Waiver (ed<HS)	0.0066 (0.0054)	0.330* (0.192)
TANF (weak work incentives) (ed<HS)	0.0017 (0.0187)	0.031 (0.759)
TANF (mixed work incentives) (ed<HS)	0.0182* (0.0109)	0.755 (0.508)
TANF (strong work incentives) (ed<HS)	0.0310*** (0.0114)	1.036** (0.486)
Waiver (ed=HS)	0.0016 (0.0032)	0.099 (0.136)
TANF (weak work incentives) (ed=HS)	-0.0157 (0.0111)	-0.362 (0.463)
TANF (mixed work incentives) (ed=HS)	0.0019 (0.0059)	0.072 (0.229)
TANF (strong work incentives) (ed=HS)	0.0032 (0.0066)	0.124 (0.287)
Waiver (ed>HS)	0.0016 (0.0025)	0.187 (0.137)
TANF (weak work incentives) (ed>HS)	-0.0080 (0.0078)	-0.120 (0.344)
TANF (mixed work incentives) (ed>HS)	0.0055 (0.0036)	0.511** (0.200)
TANF (strong work incentives) (ed>HS)	0.0007 (0.0042)	0.408* (0.236)

Means (1989-1992)	Employed	Hours Worked
ed<HS	0.433	14.1
ed=HS	0.689	23.9
ed>HS	0.786	27.6

***,**,* indicate statistical significance at the 1%, 5%, and 10% levels respectively. Robust standard errors that are corrected for clustering in state-year cells appear in parentheses. Regressions and means are weighted by the CPS-ORG weight. The results are from regressions run using specification (3).

(p-value= 0.138). Lastly, dropouts in states with TANF programs characterized as having weak work incentives experienced neither an employment nor a labor supply gain.

The bottom of Table 4 indicates that, after the enactment of federal welfare reform, hours worked by women with at least some college education rose slightly in states that implemented TANF programs with both mixed work incentives (1.9 percent) and strong work incentives (1.5 percent). The labor supply increase for the high-education group in states with TANF programs characterized as having strong work incentives, in percentage terms, was only about one-fifth the size of the labor supply increase for dropouts in those same states. Arguably, the generosity of welfare was reduced more in states that implemented TANF programs with either strong or mixed work incentives than in states that implemented TANF programs with weak work incentives.[19] A possible reason why the labor supply of high-education women increased despite their low welfare participation rate is that they perceived the implementation of TANF as a signal that their state government would adopt legislation in the future that would make other government assistance programs, such as family leave programs, less generous, and responded by working more.

Next, I present the results of regressions run separately by race or ethnicity. The results of regressions run on samples of black women appear in Table 5. While black dropouts did not work more under waivers, they experienced large employment and labor supply gains under TANF (11.9 percent and 17.4 percent, respectively). As is found for all dropouts, pooling over race and ethnicity, the employment gain for black dropouts under TANF was larger in waiver states (13.2 percent) than in non-waiver states (10.9 percent, p-value= 0.103), and the increase in hours worked by black dropouts under TANF was larger in waiver states (23.3 percent) than in non-waiver states (13.0 percent). That large work gains are estimated for black dropouts under TANF is unsurprising since, as was noted previously, this group has such a high welfare participation rate.

<u>Table 5. Results for black women, TANF effects constrained and unconstrained (1989-2004)</u>

Panel A	Employed	Hours Worked
Waiver (ed<HS)	-0.0141 (0.0161)	0.134 (0.618)
TANF (ed<HS)	0.0434** (0.0207)	2.075*** (0.783)
Waiver (ed=HS)	0.0042 (0.0083)	0.145 (0.356)
TANF (ed=HS)	0.0036 (0.0131)	-0.041 (0.507)
Waiver (ed>HS)	-0.0052 (0.0071)	-0.025 (0.334)
TANF (ed>HS)	0.0054 (0.0097)	0.540 (0.456)
Panel B	Employed	Hours Worked
Waiver (ed<HS)	-0.0118 (0.0173)	0.470 (0.666)
TANF (no waiver) (ed<HS)	0.0397 (0.0243)	1.549* (0.893)
TANF (had waiver) (ed<HS)	0.0483* (0.0259)	2.772*** (1.021)

Waiver (ed=HS)	0.0060 (0.0093)	0.374 (0.409)
TANF (no waiver) (ed=HS)	0.0002 (0.0163)	-0.469 (0.611)
TANF (had waiver) (ed=HS)	0.0072 (0.0154)	0.402 (0.649)
Waiver (ed>HS)	-0.0094 (0.0079)	-0.073 (0.372)
TANF (no waiver) (ed>HS)	0.0143 (0.0111)	0.643 (0.569)
TANF (had waiver) (ed>HS)	-0.0023 (0.0125)	0.451 (0.534)

Means (1989-1992)	Employed	Hours Worked
ed<HS	0.365	11.9
ed=HS	0.638	22.4
ed>HS	0.785	28.3

***,**,* indicate statistical significance at the 1%, 5%, and 10% levels respectively. Robust standard errors that are corrected for clustering in state-year cells appear in parentheses. Regressions and means are weighted by the CPS-ORG weight. Results from regressions run using specification (1) appear in Panel A, and results from regressions run using specification (2) appear in Panel B.

The results of regressions run on samples of Hispanic women appear in Table 6. Hispanic high school dropouts and high school graduates realized similar work gains under waivers.[20] While, during that reform regime, Hispanic dropouts experienced an employment gain of 5.1 percent and a labor supply gain of 5.2 percent, Hispanic high school graduates experienced an employment gain of 5.3 percent and a labor supply gain of 4.3 percent. No work gains are estimated for any group of Hispanic women under TANF.

Table 6. Results for Hispanic women, TANF effects constrained and unconstrained (1989-2004)

Panel A	Employed	Hours Worked
Waiver (ed<HS)	0.0199** (0.0086)	0.691** (0.331)
TANF (ed<HS)	0.0268 (0.0175)	0.605 (0.858)
Waiver (ed=HS)	0.0330*** (0.0110)	0.938** (0.472)
TANF (ed=HS)	0.0005 (0.0182)	0.312 (0.728)

Waiver (ed>HS)	-0.0141 (0.0098)	-0.307 (0.469)
TANF (ed>HS)	0.0045 (0.0124)	0.485 (0.644)

Panel B	Employed	Hours Worked
Waiver (ed<HS)	0.0199** (0.0100)	0.883** (0.401)
TANF (no waiver) (ed<HS)	0.0266 (0.0166)	0.118 (0.749)
TANF (had waiver) (ed<HS)	0.0269 (0.0207)	0.876 (1.024)
Waiver (ed=HS)	0.0261** (0.0131)	0.659 (0.585)
TANF (no waiver) (ed=HS)	0.0150 (0.0216)	0.893 (0.957)
TANF (had waiver) (ed=HS)	-0.0102 (0.0207)	-0.116 (0.771)
Waiver (ed>HS)	-0.0198 (0.0120)	-0.416 (0.564)
TANF (no waiver) (ed>HS)	0.0146 (0.0140)	0.677 (0.670)
TANF (had waiver) (ed>HS)	-0.0045 (0.0169)	0.314 (0.828)

Means (1989-1992)	Employed	Hours Worked
ed<HS	0.393	13.3
ed=HS	0.623	21.7
ed>HS	0.738	26.4

***,**,* indicate statistical significance at the 1%, 5%, and 10% levels respectively. Robust standard errors that are corrected for clustering in state-year cells appear in parentheses. Regressions and means are weighted by the CPS-ORG weight. Results from regressions run using specification (1) appear in Panel A, and results from regressions run using specification (2) appear in Panel B.

Table 7. Results for white women, TANF effects constrained and unconstrained (1989-2004)

Panel A	Employed	Hours Worked
Waiver (ed<HS)	0.0074 (0.0084)	0.265 (0.320)
TANF (ed<HS)	0.0196 (0.0140)	0.866 (0.529)
Waiver (ed=HS)	-0.0027 (0.0040)	-0.025 (0.167)
TANF (ed=HS)	-0.0010 (0.0069)	-0.069 (0.282)

Waiver (ed>HS)	0.0025 (0.0028)	0.219 (0.147)
TANF (ed>HS)	0.0027 (0.0037)	0.400** (0.200)

Panel B	Employed	Hours Worked
Waiver (ed<HS)	0.0086 (0.0093)	0.447 (0.356)
TANF (no waiver) (ed<HS)	0.0173 (0.0163)	0.506 (0.588)
TANF (had waiver) (ed<HS)	0.0220 (0.0155)	1.222** (0.614)
Waiver (ed=HS)	-0.0068 (0.0045)	-0.076 (0.186)
TANF (no waiver) (ed=HS)	0.0069 (0.0084)	0.029 (0.328)
TANF (had waiver) (ed=HS)	-0.0090 (0.0080)	-0.168 (0.323)
Waiver (ed>HS)	0.0035 (0.0032)	0.295* (0.161)
TANF (no waiver) (ed>HS)	0.0007 (0.0044)	0.252 (0.225)
TANF (had waiver) (ed>HS)	0.0044 (0.0044)	0.538** (0.233)

Means (1989-1992)	Employed	Hours Worked
ed<HS	0.478	15.3
ed=HS	0.707	24.4
ed>HS	0.792	27.7

***,**,* indicate statistical significance at the 1%, 5%, and 10% levels respectively. Robust standard errors that are corrected for clustering in state-year cells appear in parentheses. Regressions and means are weighted by the CPS-ORG weight. Results from regressions run using specification (1) appear in Panel A, and results from regressions run using specification (2) appear in Panel B.

Lastly, the results of regressions run on samples of white women appear in Table 7. While, under TANF, white dropouts worked 5.7 percent more weekly hours (p-value= 0.102), white women with at least some college education also experienced a small labor supply increase (1.4 percent).[21] For both groups, labor supply gains following the implementation of the federal welfare bill were concentrated in waiver states. In percentage terms, however, the increase in hours worked by high-education women under TANF in waiver states (1.9 percent) was less than one-fourth the size of the increase in hours worked by white dropouts under TANF in those same states (8.0 percent).

V. CONCLUSION

I find that female high school dropouts worked more hours under waivers, and experienced both an employment and labor supply increase under TANF. Although I also find that the labor supply of women with at least some college education rose under TANF, the estimated increase is quite small. That work gains under TANF are concentrated among dropouts, the education group that has the highest welfare participation rate and is thus most likely to be affected by reform, lends credibility to the notion that federal welfare reform caused women to work more. Further, I estimate that under TANF there were large gains in employment and labor supply for black dropouts, a group in which, prior to reform, over thirty percent lived in a household that received some AFDC income. This result strengthens the case that policy effects are captured.

I generally estimate that women experienced larger work gains under TANF than under waivers. This result is not surprising since states implemented more comprehensive reform packages under TANF than under waivers.[22] I also find that employment and labor supply gains under TANF were larger in waiver states than in non-waiver states. This finding suggests that states that had previously adopted waivers, having had more experience with welfare reform, might have implemented more far-reaching reform initiatives under TANF than states that had never adopted waivers. Lastly, consistent with expectations, I generally find that, after the enactment of federal welfare reform, women experienced the largest employment and labor supply gains in states that implemented TANF programs with strong work incentives. This finding lends further support to the notion that policy effects are indeed captured.

Since reform could have affected marriage and caused the pool of unmarried women to become more skilled, relying on a sample of some group of unmarried women or single mothers could produce upward-biased estimates of the effects of reform on employment. Hence, it is a concern that the only previous paper using post-1996 data that does not find that TANF increased female employment is the only one that does not restrict its sample to some group of unmarried women or single mothers. In this paper, however, using a sample of female high school dropouts that is not restricted to unmarried women or single mothers, I do find evidence that federal welfare reform contributed to the post-1996 employment and labor supply gains of low-skill women.

ENDNOTES

1. These figures come from the Bureau of Labor Statistics Website: http://www.bls.gov.
2. The age of the youngest child which would relieve women from having to participate in work activities varied from state to state. Some states did not allow work exemptions based on the age of the youngest child.
3. Both Blank (2002), in her literature review of the effects of welfare reform on various outcomes of interest, and Moffitt (2002), in his literature review of the effects of different welfare programs on

labor supply, point out that the results in the Schoeni and Blank (2000) paper differ from other results in the literature.

4. The data refer to calendar years 1989-1992, before the onset of reform. By the end of 1992, only three states had implemented a waiver program.
5. See Bitler et al. (2003).
6. I start my analysis in 1989 because that starting point precedes the 1990-1991 recession and the implementation of the first state waiver programs in 1992. A number of researchers begin their analysis in 1989 (Bitler et al. (2004), Bitler et al. (2006)).
7. This dependent variable takes on a value of one if the woman was employed last week and a value of zero otherwise. The dependent variable Schoeni and Blank (2000) use is employment status last year. When examining the impact of welfare reform on employment, using employment status last week as the dependent variable is probably preferable. Suppose that in a given state-year in which reform was implemented a woman is recorded in the March CPS as having been employed for part of that year. The researcher cannot determine whether that woman was employed before the implementation of reform, after the implementation of reform, or during both periods. When employment status last week is used as the dependent variable, on the other hand, the researcher can always tell whether or not the woman was employed while reform was in effect.
8. Waiver and TANF implementation dates come from Table A1 in Schoeni and Blank (2000). Between October 1992 and February 1997, 29 states implemented waivers. Between September 1996 and January 1998, all 51 states (including Washington DC) implemented TANF.
9. Schoeni and Blank (2000) use the March CPS surveys from 1977 to 1999 to examine the impact of welfare waivers and TANF on a variety of outcomes. Pooling all women into their sample, they create three education groups (less than 12 years, 12 years, more than 12 years) and four age groups (16-25, 26-34, 35-44, 45-54). The authors regress each variable of interest against a waiver dummy interacted with each education group, a TANF dummy interacted with each education group, education dummies, age group dummies, interactions between the age group dummies and the education dummies, interactions between state economic variables (current and lagged unemployment rate and current and lagged employment growth rate) and the education dummies, interactions between the log maximum AFDC benefit for a family of three and the education dummies, year effects, year effects interacted with the education dummies, state effects, linear state-specific time trends, a Hispanic dummy, and a non-Hispanic black dummy.

   Notably, Schoeni and Blank do not interact all of their right-hand-side variables with the education groups. Education group-specific state fixed effects, for instance, are potentially important omitted variables since the education differential between high-education and low-education women could differ from state to state for any number of reasons.

Using individual-level data from the CPS-ORG surveys from1989-2004, I regress employment status against the same right-hand-side variables listed above, as well as month effects and the interactions between the state effects and the education dummies. Performing an F-test, I reject the null that the interactions between the state effects and the education dummies equal zero (F-stat= 12.29, p-value=0.000). Pooling women over education group and not allowing the state effects to vary by education group is not a specification that is supported empirically.

10. The black samples consist of non-Hispanic black women and the white samples consist of non-Hispanic white women. When regressions are run separately by race or ethnicity, the non-Hispanic black dummy and the Hispanic dummy are dropped.
11. Blank (2002) notes that the specific reform policies that are typically coded are time limits, family caps, benefit reduction rates, work exemptions, work requirements, and sanctions.
12. Reviewing studies that investigate the effects of specific reform policies on various outcomes, Blank (2002) notes that there is a "regular occurrence of perverse signs on some of [the] policy variables" (1137), Bell (2001) argues that "the analysis of individual reform measures tells us little on which we can depend" (38), and Moffitt (2002) concludes that "there are almost no credible studies of the impact of different individual components of reform taken individually" (2423). Moffitt notes that Grogger's papers on time limits (2003 and 2004b) are exceptions.
13. See Table 3-5 in Blank and Schmidt (2001) for a detailed explanation of the coding scheme. The authors characterize 10 states as having TANF programs with weak work incentives (including Washington DC), 27 states as having TANF programs with mixed work incentives, and 14 states as having TANF programs with strong work incentives.
14. Blank and Schoeni (2003) find that between the 1992-1995 and 1997-2000, the family income gains for children without both parents present in states with strong work incentives were greater than the family income gains for children without both parents present in states with weak work incentives.
15. Because small sample sizes might produce unreliable estimates, I do not run regressions separately by race or ethnicity using specification (3). TANF programs that were characterized as having weak work incentives were implemented in only 10 states. Only 1,259 of the observations from the black high school dropout sample come from state-year cells in which a TANF program with weak work incentives was in effect. In two of the ten states with weak work incentives, the fraction of the female population aged 16-54 that was black was less than 1 percent, and in four other of those states, the fraction of the female population aged 16-54 that was black was less than 4 percent.
16. Coefficient estimates for additional explanatory variables for the employment regressions are displayed in Appendix Table 1.

17. The most recent data Schoeni and Blank (2000) use, because fewer years of data were available to them when they initially wrote their paper, are data that refer to 1998. In Appendix Table 2, I report my results from regressions run on all female high school dropouts, all female high school graduates, and all females with at least some college education when the samples are restricted to the 1989-1998 time period. Whether examining the 1989-2004 time period or the 1989-1998 time period, the effects of TANF on the employment and labor supply of female high school dropouts are estimated to be quite similar. While, when the longer time frame is used, female high school dropouts are found to have experienced an employment increase of 1.99 percentage points (p-value=0.052) and a labor supply increase of 0.759 hours (p-value=0.104) under TANF, when the shorter time frame is used, this group is found to have experienced an employment increase of 1.53 percentage points (p-value=0.106) and a labor supply increase of 0.654 hours (p-value=0.121) during the same reform regime.
18. Table 1 indicates that, prior to reform, 1.9 percent of women with at least some college education lived in a household that received some AFDC income during the year.
19. Blank and Schmidt (2001) note that, according to their coding scheme, a "state with unambiguously strong work incentives would have low benefit generosity, high earnings disregards, strict sanctions, and strict time limits" (85). Low benefit levels, strict sanctions, and strict time limits reduce the generosity of welfare. High earnings disregards, however, increase the generosity of welfare.
20. Table 1 indicates that, prior to reform, 13.9 percent of Hispanic high school dropouts and 7.5 percent of Hispanic high school graduates lived in a household that received some AFDC income during the year.
21. Table 1 indicates that, prior to reform, 10.8 percent of white high school dropouts and 1.2 percent of white women with at least some college education lived in a household that received some AFDC income during the year.
22. For instance, the five-year lifetime time limit and the two-year continuous-use time limit established in the federal welfare bill became components of every state's TANF program.

DATA APPENDIX

Unemployment rates and employment growth rates: BLS website, http://www.bls.gov.

CPI: BLS website, http://www.bls.gov/cpi/home.htm. I use the CPI for All Urban Consumers. AFDC benefit levels are expressed in 2000 dollars.

Maximum AFDC benefits for a family of three: *Green Book, Background Material and Data on Programs within the Jurisdiction of the Committee on Ways and Means*, various years. The values I use for the benefit levels in 1999, 2001, and 2004 are the previous year's benefit levels adjusted for inflation.

Waiver and TANF implementation dates: Table A1 in Schoeni and Blank (2000).

Coding the intensity of work incentives of each state's TANF program: Table 3-5 in Blank and Schmidt (2001).

REFERENCES

Bell, Stephen H. 2001. "Why Are Welfare Caseloads Falling?" *Assessing the New Federalism* Discussion Paper 01-02. Washington, DC: Urban Institute.

Bitler, Marianne P., Jonah B. Gelbach, and Hilary W. Hoynes. 2002. "The Impact of Welfare Reform on Living Arrangements." NBER Working Paper 8784. Cambridge, Mass.: National Bureau of Economic Research.

Bitler, Marianne P., Jonah B. Gelbach, and Hilary W. Hoynes. 2003. "Some Evidence on Race, Welfare Reform, and Household Income." *American Economic Review* 93(2): 293-298.

Bitler, Marianne P., Jonah B. Gelbach, and Hilary W. Hoynes. 2006. "Welfare Reform and Children's Living Arrangements." *Journal of Human Resources* 41(1): 1-27.

Bitler, Marianne P., Jonah B. Gelbach, Hilary W. Hoynes, and Madeline Zavodny. 2004. "The Impact of Welfare Reform on Marriage and Divorce." *Demography* 41(2): 213-236.

Blank, Rebecca M. 2001. "Declining Caseloads/ Increased Work: What Can We Conclude About the Effects of Welfare Reform?" *Economic Policy Review* 7(2): 25-36.

Blank, Rebecca M. 2002. "Evaluating Welfare Reform in the United States." *Journal of Economic Literature* 40(4): 1105-1166.

Blank, Rebecca M. and Lucie Schmidt. 2001. "Work, Wages, and Welfare," in *The New World of Welfare*. Rebecca M. Blank and Ron Haskins, eds. Washington, DC: Brookings Institution. pp. 70-102.

Blank, Rebecca M. and Robert F. Schoeni. 2003. "Changes in the Distribution of Children's Family Income over the 1990's." *American Economic Review* 93(2): 304-308.

Council of Economic Advisors. 1999. "Economic Expansion, Welfare Reform, and the Decline in Welfare Caseloads: An Update." Washington, D.C.: Executive Office of the President.

Fang, Hanming and Michael P. Keane. 2004. "Assessing the Impact of Welfare Reform on Single Mothers." *Brookings Papers on Economic Activity* 2004(1): 1-95.

Grogger, Jeffrey. 2003. "The Effects of Time Limits, the EITC, and Other Policy Changes on Welfare Use, Work, and Income Among Female-Headed Families." 2003. *The Review of Economics and Statistics* 85(2): 394-408.

Grogger, Jeffrey. 2004a. "Comment on 'Assessing the Impact of Welfare Reform on Single Mothers.'" *Brookings Papers on Economic Activity* 2004(1): 102-112.

Grogger, Jeffrey. 2004b. "Time Limits and Welfare Use." *Journal of Human Resources* 39(2): 405-424.

Kaestner, Robert and Neeraj Kaushal. 2005. "Immigrant and Native Responses to Welfare Reform." *Journal of Population Economics* 18(1): 69-92.

Kaushal, Neeraj, and Robert Kaestner. 2001. "From Welfare to Work: Has Welfare Reform Worked?" *Journal of Policy Analysis and Management*. 20(4): 699-719.

Meyer, Bruce D. and Dan T. Rosenbaum. 2001. "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers." *The Quarterly Journal of Economics* 116(3): 1063-1114.

Moffitt, Robert A. 1999. "The Effects of Pre-PRWORA Waivers on AFDC Caseloads and Female Earnings, Income, and Labor Force Behavior." NBER Working Paper 7627. Cambridge, Mass.: National Bureau of Economic Research.

Moffitt, Robert A. 2002. "Welfare Programs and Labor Supply," in *Handbook of Public Economics: Volume 4*. A.J. Auerbach and M. Feldstein, eds. New York, NY: Elsevier Science B.V. pp. 2393-2430.

O'Neill, June E. and M. Anne Hill. 2001. "Gaining Ground? Measuring the Impact of Welfare Reform on Welfare and Work." Civic Report 17, Center for Civic Innovation. New York, NY: Manhattan Institute Schoeni, Robert F. and Rebecca M. Blank. 2000. "What Has Welfare Reform Accomplished? Impacts on Welfare Participation, Employment, Income Poverty, and Family Structure." NBER Working Paper 7627. Cambridge, Mass.: National Bureau of Economic Research.

Schoeni, Robert F. and Rebecca M. Blank. 2000. "What Has Welfare Reform Accomplished? Impacts on Welfare Participation, Employment, Income Poverty, and Family Structure." NBER Working Paper 7627. Cambridge, Mass.: National Bureau of Economic Research.

Appendix Table 1. Detailed results for employment regressions (1989-2004)

	Employed (ed<HS)	Employed (ed=HS)	Employed (ed>HS)
Waiver	0.0065 (0.0054)	0.0013 (0.0032)	0.0014 (0.0025)
TANF	0.0199* (0.0102)	0.0004 (0.0054)	0.0026 (0.0035)
Unemployment Rate	-0.0088** (0.0042)	-0.0091*** (0.0021)	-0.0066*** (0.0018)
Unemployment Rate (lagged one year)	-0.0035 (0.0034)	-0.0036** (0.0018)	-0.0008 (0.0014)
Employment Growth Rate	-0.0020 (0.0018)	-0.0011 (0.0010)	0.0006 (0.0007)
Employment Growth Rate (lagged one year)	-0.0001 (0.0017)	-0.0003 (0.0010)	-0.0001 (0.0007)
Log Real Maximum Welfare Benefits	-0.0233 (0.0266)	-0.0052 (0.0165)	-0.0040 (0.0139)

Age	0.0197*** (0.0008)	0.0140*** (0.0006)	-0.0051*** (0.0005)
Age-squared	-0.0002*** (0.0000)	-0.0002*** (0.0000)	0.0001*** (0.0000)
Non-Hispanic Black	-0.0864*** (0.0050)	-0.0497*** (0.0035)	0.0083*** (0.0024)
Hispanic	-0.0581*** (0.0061)	-0.0604*** (0.0039)	-0.0355*** (0.0035)
Mean (1989-1992)	0.393	0.623	0.738

***,**,* indicate statistical significance at the 1%, 5%, and 10% levels respectively. Robust standard errors that are corrected for clustering in state-year cells appear in parentheses. Regressions and means are weighted by the CPS-ORG weight. Results are for regressions run using specification (1).

Appendix Table 2. Results for all women, TANF effects constrained and unconstrained (1989-1998)

Panel A	Employed	Hours Worked
Waiver (ed<HS)	-0.0029 (0.0057)	0.019 (0.197)
TANF (ed<HS)	0.0153 (0.0095)	0.654 (0.421)
Waiver (ed=HS)	-0.0005 (0.0035)	0.034 (0.151)
TANF (ed=HS)	-0.0035 (0.0051)	-0.114 (0.224)
Waiver (ed>HS)	0.0025 (0.0024)	0.160 (0.140)
TANF (ed>HS)	0.0042 (0.0035)	0.392** (0.183)

Panel B	Employed	Hours Worked
Waiver (ed<HS)	-0.0041 (0.0066)	0.050 (0.229)
TANF (no waiver) (ed<HS)	0.0177* (0.0101)	0.592 (0.417)
TANF (had waiver) (ed<HS)	0.0133 (0.0113)	0.706 (0.507)
Waiver (ed=HS)	-0.0046 (0.0042)	-0.039 (0.181)
TANF (no waiver) (ed=HS)	0.0038 (0.0059)	0.020 (0.253)
TANF (had waiver) (ed=HS)	-0.0106* (0.0061)	-0.242 (0.267)

Waiver (ed>HS)	0.0032 (0.0029)	0.213 (0.169)
TANF (no waiver) (ed>HS)	0.0029 (0.0041)	0.294 (0.209)
TANF (had waiver) (ed>HS)	0.0054 (0.0041)	0.480** (0.229)

Means (1989-1992)	Employed	Hours Worked
ed<HS	0.433	14.1
ed=HS	0.689	23.9
ed>HS	0.786	27.6

***,**,* indicate statistical significance at the 1%, 5%, and 10% levels respectively. Robust standard errors that are corrected for clustering in state-year cells appear in parentheses. Regressions and means are weighted by the CPS-ORG weight. Results from regressions run using specification (1) appear in Panel A, and results from regressions run using specification (2) appear in Panel B.

The Impact of Head Start Participation on the Criminal Behavior of Teenagers

Mark Gius*

ABSTRACT

The purpose of the present study is to estimate the effect of Head Start participation on the criminal behavior of teenagers. Using National Longitudinal Survey of Youth data, the present study finds that participation in the Head Start program does not reduce the likelihood that a person engages in criminal activity. In fact, results of the present study show that, holding all other factors constant, teenagers who had participated in the Head Start program as children were more likely to be arrested but were no more likely to commit a crime than a teenager who did not participate in the program as a child. These results are rather robust since factors such as race, sex, and family and peer influences are all held constant.

INTRODUCTION

Head Start is a federally-funded but locally administered program devoted to early childhood development. Begun in 1965, Head Start enrolls over 900,00 children in over 1600 programs nationwide. The vast majority of children attending these programs, which are run by local non-profits or school systems, are 3 or 4 years old. For 2004, the Head Start program was appropriated almost $6.8 billion, and the average cost per child was $7,222. All data and background information on the Head Start program were obtained from the Head Start Bureau website.

When the program began in the 1960s, the objective was to help local communities meet the needs of disadvantaged children. There was a general consensus that children from lower income families were much more likely to do poorly in school and drop out, thus increasing the probability that these children would have poor employment prospects and hence be unable to break out of the cycle of poverty. Early proponents of Head Start believed that if intervention could happen during the pre-school years, then possibly these economically-disadvantaged children would stand a better chance of succeeding in school and therefore greatly increase their earnings potential. In addition to improved academic performance, other benefits of early intervention that have been espoused include improved health and reduced likelihood of future criminal activity.

There has been a limited amount of research conducted on the benefits of the Head Start program. Most of this prior research has concentrated on the impact of Head Start on future academic

*Department of Economics, Quinnipiac University, Hamden, CT 06518, 203-582-8576, Fax: 203-582-8664, gius@quinnipiac.edu

performance (Cole and Washington 1986; Currie and Thomas,1995; Currie and Thomas 2000). Only one paper has examined the effect of Head Start participation on future criminal activity (Garces, Thomas, and Currie 2002). In this paper, the authors use data on adults from the Panel Study of Income Dynamics for 1995 and attempt to determine if Head Start had any effect on a variety of individual outcomes, including completion of high school, attending college, and being arrested. Using a fixed effects model with a sample size of 3255, they found that African-Americans who participated in Head Start were less likely to be charged with a crime; the probability that a white person was charged with a crime was not affected by Head Start participation.

The present study differs from the Garces, Thomas and Currie (2002) paper in several ways. First, instead of using the Panel Study of Income Dynamics, the National Longitudinal Survey of Youth is used. This data set is uniquely suited to the estimation of an economic model of crime for young adults, since it contains a vast amount of personal information about the survey respondents, including information about any crimes that the respondents may have committed. Second, the year examined in the present study is 1998, a more recent year than the Garces, Thomas, and Currie paper. Third, only teenagers are examined in the present study; hence, the effects of Head Start should be more pronounced than in a study looking only at adults. Finally, variables capturing family and peer influences are included in the present study; no prior study looking at the effects of Head Start has used family or peer influences as explanatory variables.

THEORETICAL FOUNDATIONS AND EMPIRICAL TECHNIQUE

In order to construct an equation that estimates the determinants of crime at the individual level, prior research in the area of criminal behavior was examined (Carr-Hill and Stern 1973; Sandelin and Skogh 1986; Britt 1994; Young 1993; Howsen and Jarrell 1987; Benson, Kim, Rasmussen, and Zuehlke 1992; Rodney, Tachia, and Rodney 1999; Williams, Ayers, Abbott, Hawkins, and Catalano 1999; Videon 2002; Paschall, Ringwalt, and Flewelling 2003; Kierkus and Baer 2003; Eamon and Altshuler 2004; Simons, Simons, and Conger 2004).

Most of these studies use as their data aggregate measures of crime, such as county, state, or even national-level crime statistics, and aggregate measures of socioeconomic characteristics. While these studies have shed light on the effects of socioeconomic characteristics and institutional factors on criminal activities, there has been some criticism of the use of aggregate data to model criminal behavior, a behavior that is an individual choice about lifestyles and income-generating opportunities.

In addition, many of these prior studies use crime statistics data that may not capture all criminal activity, but only reported criminal activity. Hence, aggregate studies that use reported crime statistics not only ignore the individualistic nature of crime, but also seriously under-report the level of criminal activity within any given jurisdiction.

In order to develop a model more appropriate for the purposes of the present study, the economic model of crime developed by Becker (1968) is used a theoretical basis. In this model, individuals are characterized as being rational decision makers who respond to opportunities, both legitimate and illegitimate. The rational individual decides which activities to undertake by examining the expected returns from all opportunities. If a legitimate activity for an individual has a greater return than all other activities, then that individual takes advantage of that opportunity. If, however, an illegal activity has a greater return for an individual than all other activities, then that individual commits the criminal act.

In the present study, however, it is assumed that a teenager does not work, and hence there are no returns from legitimate activities. In addition, it is assumed that a teenager is provided all necessary goods and services by his or her parents; hence, stealing is not economically necessary.

Within the context of the present study, these assumptions are reasonable. For the sample used, the average annual income of the teenager was \$2,272, while the average income of the household was \$58,561. Hence, teenagers earned less than 4 per cent of total household income. Therefore, it is reasonable to assume that teenagers could not provide for themselves financially and that they were dependent on their parents for most, if not all, of their necessities. Hence, crimes are committed by teenagers for utility (happiness or satisfaction) and not for economic or monetary purposes.

It is important to note that these averages were calculated using limited samples. Most respondents did not answer the income questions. In addition, the sample used to estimate the teenage income was different from the sample used to estimate the household income. Finally, due to these data constraints, income was not included as an explanatory variable in any of the regressions estimated in the present study.

Given the above, one may model criminal behavior as follows:

$$(1)\quad Max\ U = f(C(I_p, I_f, HEAD, \boldsymbol{X}),\ AG,\ L(\boldsymbol{X}))$$

$$(2)\quad s.\ t.\ T = C + L$$

where U is total utility, C denotes criminal activities, AG is all goods, **X** is a vector of variables that include socioeconomic characteristics and other variables that may explain the propensity to commit crimes, and L denotes leisure (non-criminal) activities. Assuming that the teenager has no income, and that all goods are provided by his or her parents, the only constraint would then be total time in a given day. Assuming T is less than 24, a teenager attempts to maximize his or her utility by allocating his or her time between criminal and non-criminal activities. In the present study, it is assumed that all non-violent crimes are property crimes, such as theft.

L is assumed to be exogenous while C is assumed to depend upon the influences of the teenager's parents (I_P), the influences of his or her friends (I_F), the influences of Head Start (HEAD), and other variables. Parental and peer influences and the influence of Head Start may be either positive or negative regarding criminal behavior. If all influences are positive in nature, then all time is spent pursuing non-criminal activities, since criminal activities are viewed as bads and not goods. If, however, the influences are negative in nature, then the individual would engage in criminal activities, since both criminal activities and leisure would be viewed as goods. Hence, in the first case, with criminal activities being viewed as a bad, a corner solution would result, and the teenager would only engage in non-criminal leisure activities. If, however, criminal activity is viewed as good, then an interior solution would result, and the individual would divide his/her time between non-criminal and criminal activities.

In attempting to empirically model participation in criminal activities, it is important to note that young adults make a discrete choice about whether or not to commit a crime. Using equations (1) and (2) as a theoretical basis, the following equation will be estimated:

$$\begin{aligned} CRIME = {} & a_1\, MALE + a_2\, SOUTH + a_3\, WHITE + a_4\, UR + a_5\, URBAN \\ & + a_6\, SIZE + a_7\, AGE + a_8\, PSMOKE + a_9\, PDRINK + a_{10}\, PGANG \\ & + a_{11}\, PDRUG + a_{12}\, PARENTS + a_{13}\, DINNER + a_{14}\, NOHS \\ & + a_{15}\, HEAD + a_{16}\, EDMOM + a_{17}\, EDDAD + u \end{aligned} \tag{3}$$

where CRIME equals one if person committed a criminal act and zero otherwise; MALE equals one if person is male and zero otherwise; SOUTH equals one if person is from southern states and zero otherwise; WHITE takes value of one if person is white and zero otherwise; UR equals one if the unemployment rate for the respondent's labor market of current residence is less than six percent and zero otherwise; URBAN equals one if person lives in urban area and zero otherwise; SIZE is size of respondent's household; AGE is respondent's age; PSMOKE equals one if respondent's peers smoke cigarettes and zero otherwise; PDRINK equals one if respondent's peers drink alcoholic beverages and zero otherwise; PDRUG takes value of one if respondent's peers use illicit drugs and zero otherwise; PGANG equals one if respondent's peers belong to a gang and zero otherwise; PARENTS equals one if respondent has two parents and zero otherwise; DINNER is the number of times the respondent eats dinner with his or her family in one week; NOHS equals one if person dropped out or was expelled from school; HEAD equals one if respondent participated in the Head Start program; EDMOM equals one if the mother has at most a high school diploma and zero otherwise; EDDAD equals one if the father has at most a high school diploma and zero otherwise; and u is a normally-distributed, random error term. The family influences are represented by PARENTS and DINNER, and the peer influences are represented by PDRINK, PSMOKE, PDRUG, and PGANG.

Given the theoretical foundations discussed above, MALE, UR, URBAN, and NOHS are all expected to have positive effects on the probability that a person will engage in criminal activities; these hypotheses are supported by theory and prior research. WHITE, AGE, and SOUTH are all expected to have negative effects on criminal acts. The above are all variables in the **X** vector of socioeconomic characteristics. It is expected that EDMOM, EDDAD, SIZE, PARENTS, DINNER, and HEAD will have negative effects on the probability that a teenager will commit criminal acts. Finally, PSMOKE, PDRINK, PDRUG, and PGANG will all positively affect the probability that a teenager will commit a crime.

EDMOM and EDDAD are used as proxies for household income. The reason for using proxies is because there was little income data available, and if the income variable was included in the present study, after eliminating all observations with missing data, the final data set would have had fewer than 50 observations. Hence, the educational attainment of the parents was used in order to capture the effects of family income on the criminal activity of teenagers.

Given the binary nature of the dependent variable, equation (3) is estimated using a probit model. This equation is estimated for the year 1998. There are two reasons for selecting this year. First, for the data set used in the present study, there were numerous survey questions asked in 1998 regarding criminal activities, and family and peer influences. Second, in 1998 for the NLSY, all respondents were less than 18 years of age.

DATA AND RESULTS

Data used in the present study were obtained from the National Longitudinal Survey of Youth. The NLSY was constructed to be a nationally representative sample of the civilian non-institutionalized population at the time of the initial survey in 1979. A second survey with a different cohort was started in 1997. The 1997 NLSY consisted of 8,984 men and women between the ages of 12 and 16 Interviews with NLSY respondents are conducted annually, and retention rates have been relatively high, averaging over 90 percent. Each age-sex cohort is represented by a multi-stage probability sample drawn by the Bureau of the Census from a list of sampling areas that had been constructed for the *Monthly Labor Survey*. The NLSY employed extensive household interviews in the selected sampling areas in order to obtain as random and as representative a sample as possible. In the present study, the 1997 NLSY was used.

In the survey, the respondent was asked if he or she performed any one of a variety of criminal acts. The five crimes reported by the NLSY are as follows: stolen anything worth less than $50; stolen anything worth more than $50; other property crimes; vandalism; and assault and battery.

In order to determine if the factors that affect criminal activity differ by the type or severity of crime committed, equation (3) was estimated for two types of crime: property crimes, which include stolen anything worth less than $50, stolen anything worth more than $50, and other property crimes; and violent crimes, which include vandalism (malicious behavior) and assault and battery

According to the NLSY data, not all criminal acts resulted in the arrest of the perpetrator; in fact, it is unknown if any of the individuals were arrested for any of the reported criminal behavior. Hence, for that reason, equation (3) is also estimated with the dependent variable ARREST, which equals one if person was arrested in the past year for one of the five crimes listed and zero otherwise.

Even though the data on illegal activities in the NLSY are based on self-reports, the mode of data collection employed by the NLSY may be as good or better than other methods of data collection. Although most studies in this area of data collection have questioned the validity of self-reported criminal data, the structure of the NLSY prevents the researcher from determining any potential over- or under-reporting by respondents. For a more complete discussion of the potential problems that arise from the use of self-reported criminal data, the reader is referred to the vast literature on this topic (Green 1990; Menard and Elliot 1990; Wyner 1980; Sampson 1985).

As previously noted, the NLSY surveys over 8,000 individuals every year; however, due to a variety of problems, there are sometimes missing responses. After eliminating all of the observations with missing responses, the sample used in the present study has 3288 observations. Of those 3288, 29.6 percent were involved in a criminal act in 1998, but only 5.2 percent were arrested. This descriptive statistic validates the estimation of both an arrest equation and a criminal act equation.

Descriptive statistics for all variables used in the present study are presented on Table 1. These statistics indicate that, for the sample examined in the present study, 66 percent are white, the average family size was 4.5, the average age was 15, 70 percent had two parents, the average number of family dinners a teenager had in a given week was 4.67, 4.4 percent of teenagers had either dropped out of school or were expelled, and 16.9 percent of respondents had participated in the Head Start program.

Table 1
Descriptive Statistics

Variable	**Minimum**	**Maximum**	**Mean**
VIOLENT CRIME	0	1	0.216
PROPERTY CRIME	0	1	0.172
ARREST	0	1	0.052
MALE	0	1	0.522
SOUTH	0	1	0.360
WHITE	0	1	0.663
UR	0	1	0.185
URBAN	0	1	0.730
SIZE	2	12	4.5
AGE	1	17	15
PSMOKE	0	1	0.368
PDRINK	0	1	0.185
PGANG	0	1	0.107
PDRUG	0	1	0.255
PARENTS	0	1	0.699
DINNER	0	7	4.67
NOHS	0	1	0.044
HEAD	0	1	0.169
EDMON	0	1	0.809
EDDAD	0	1	0.777

Notes:
N=3288

Regression results are shown in Tables 2, 3, and 4. Regarding both the property and violent crime regression results, HEAD is insignificant. This result indicates that, holding all other factors constant, including race, sex, and family and peer influences, teenagers who had participated in the Head Start program as children (5 years of age or younger) were no more likely to have been involved in criminal activity than a teenager who had not participated in Head Start. This result contradicts the result of Garces, Thomas and Currie (2002) and casts doubt on assertions made by such groups as Fight Crime: Invest in Kids that allege that Head Start graduates have lower crime rates.

Table 2
VIOLENT CRIME Regression Results

Variable	Coefficient	Standard Deviation	Test Statistic
MALE	0.432	0.0518	8.342***
SOUTH	-0.053	0.0537	-0.988
WHITE	0.130	0.0597	2.180**
UR	0.054	0.065	0.840
URBAN	0.118	0.059	2.006**
SIZE	0.0004	0.0188	0.021
AGE	-0.065	0.0092	-7.083***
PSMOKE	0.165	0.063	2.601***
PDRINK	-0.037	0.077	-0.474
PGANG	0.271	0.0819	3.317***
PDRUG	0.216	0.071	3.041***
PARENTS	-0.092	0.059	-1.56
DINNER	-0.061	0.011	-5.610***
NOHS	0.276	0.114	2.414**
HEAD	0.0407	0.071	0.573
EDMOM	0.0157	0.069	0.226
EDDAD	-0.081	0.0657	-1.230

Notes:

Log-Likelihood = -1625.414

n=3288

*** = Significant at 1%
** = Significant at 5%
* = Significant at 10%

Table 3
PROPERTY CRIME Regression Results

Variable	Coefficient	Standard Deviation	Test Statistic
MALE	0.258	0.0540	4.774***
SOUTH	-0.165	0.057	-2.893***
WHITE	0.077	0.062	1.228
UR	0.047	0.069	0.684
URBAN	0.097	0.062	1.559
SIZE	-0.034	0.0201	-1.699*
AGE	-0.0486	0.0095	-5.092***
PSMOKE	0.146	0.067	2.182**
PDRINK	-0.0037	0.079	-0.046
PGANG	0.097	0.086	1.120
PDRUG	0.267	0.073	3.635***
PARENTS	-0.1008	0.062	-1.62
DINNER	-0.059	0.011	-5.228***
NOHS	0.075	0.124	0.604
HEAD	0.0175	0.075	0.234
EDMOM	-0.041	0.072	-0.570
EDDAD	-0.086	0.0683	-1.260

Notes:
Log-Likelihood = -1446.036
n=3288
*** = Significant at 1%
** = Significant at 5%
* = Significant at 10%

Table 4
ARREST Regression Results

Variable	Coefficient	Standard Deviation	Test Statistic
MALE	0.356	0.081	4.414***
SOUTH	-0.242	0.086	-2.809***
WHITE	0.126	0.090	1.401
UR	0.024	0.099	0.244
URBAN	0.334	0.103	3.231***
SIZE	-0.0305	0.0288	-1.058
AGE	-0.125	0.0145	-8.591***
PSMOKE	0.0034	0.098	0.035
PDRINK	0.125	0.110	1.144
PGANG	0.219	0.113	1.945*
PDRUG	0.311	0.104	3.006***
PARENTS	-0.297	0.087	-3.42***
DINNER	-0.0556	0.016	-3.451***
NOHS	0.363	0.151	2.406**
HEAD	0.219	0.099	2.221**
EDMOM	0.115	0.111	1.033
EDDAD	0.0068	0.103	0.067

Notes:
Log-Likelihood = -613.1325
n=3288
*** = Significant at 1%
** = Significant at 5%
* = Significant at 10%

For the violent crime equation, MALE, WHITE, URBAN, AGE, PSMOKE, PGANG, PDRUG, DINNER, and NOHS are all statistically significant. These results indicate that family and peer influences are also significant in determining whether or not a teenager will commit a violent crime. Specifically, a young adult who eats many dinners with his or her family is much less likely to engage in violent crimes than a young person who has friends who use drugs, smoke, and are in gangs. In addition, the present study finds that young, white, urban males who dropped out of high school or were expelled are more likely to commit violent crimes than others.

For the property crime equation, MALE, SOUTH, SIZE, AGE, PSMOKE, PDRUG, and DINNER are all statistically significant. Once again, these results indicate that family and peer influences are significant in the determination of whether or not a teenager will commit a property crime. Specifically, a young adult who eats many dinners with his or her family is much less likely to engage in property crimes than a young person who has friends who use drugs or smoke. In addition, young males are more likely to commit property crimes than others.

Regarding the ARREST equation, HEAD is significant and positive, suggesting that Head Start graduates are more likely to be arrested. In fact, according to the results of the present study, Head Start graduates are 21.9 percent more likely to be arrested.

Concerning other significant explanatory variables in the ARREST regression, MALE, SOUTH, URBAN, AGE, PDRUG, PARENTS, DINNER, and NOHS are all statistically significant. Hence, young urban men who did not finish high school are more likely to be arrested. In addition, a teenager who has both parents and has dinner frequently with his or her family is much less likely to be arrested while a teenager who has friends who use drugs is much more likely to be arrested.

Interestingly, the educational attainment of the parents had no effect on either the criminal actions or the arrest of the teenage respondents. This result may suggest that the parents' educational attainment may be a less than satisfactory proxy for family income. Given the data constraints in the present study, however, no other suitable measure of family income was available.

CONCLUSION

The purpose of the present study was to estimate the effect of Head Start participation on the criminal behavior of teenagers. Using National Longitudinal Survey of Youth data and only examining individuals between the ages of 13 and 17, the present study found that participation in the Head Start program does not reduce the likelihood that a person engages in criminal activity. Results of the present study show that, holding all other factors constant, teenagers who had participated in the Head Start program as children were more likely to be arrested but were no more likely to commit crimes, either violent or property in nature. These results are rather robust since factors such as race, sex, and family and peer influences are all held constant.

These results may suggest one of two possible public policy implications. First, the Head Start program does not have as much of an affect on future individual behavior as its proponents would like to believe; hence, current funding levels are too high given the limited positive outcomes of the program. Second, the program is woefully under-funded; therefore, local agencies are not provided sufficient resources with which to engage in fruitful and long lasting interventionist programs with the economically-disadvantaged children of their local communities. Hence, funding levels should be greatly increased. Which implication is most likely the correct course of action is not addressed in the present study.

Finally, results of the present study also suggest that teenagers who have two parents and who eat dinner often with their families are much less likely to be arrested than teenagers who have friends that use drugs. In addition, young males, who never completed high school are much more likely to commit violent crimes and to be arrested. Unfortunately, family dinner opportunities and increased prestige for academic excellence among teenage peer groups are not areas in which government policies have had a great deal of success. Hence, the government's ability to reduce the root causes of the criminal activity of young adults may be rather limited.

REFERENCES

Becker, Gary . 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76(1): 169-217.

Benson, Bruce, Iljoong Kim, David Rasmussen, and Thomas Zuehlke. 1992. "Is Property Crime Caused by Drug Use or by Drug Enforcement Policy." *Applied Economics* 24 (7): 679-692.

Britt, Chester. 1994. "Crime and Unemployment Among Youths in the United States, 1958-1990." *American Journal of Economics and Sociology* 53(1): 99-109.

Carr-Hill, R. And N. Stern. 1973. "An Econometric Model of the Supply and Control of Recorded Offense in England and Wales." *Journal of Public Economics* 2(1): 289-318.

Cole, Jackson and Valora Washington. 1986. "A Critical Analysis of the Assessment of the Effects of Head Start on Minority Children." *The Journal of Negro Education* 55(1): 91-106.

Currie, Janet and Duncan Thomas. 1995. "Does Head Start Make a Difference?" *American Economic Review* 85(3): 341-364.

Currie, Janet and Duncan Thomas. 2000. "School Quality and the Longer-Term Effects of Head Start." *The Journal of Human Resources* 35(4): 755-774.

Eamon, Mary and Sandra Altshuler. 2004. "Can We Predict Disruptive School Behavior? *Children & Schools* 26(1): 23-37.

Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer-Tern Effects of Head Start." *American Economic Review* 92(4): 999-1012.

Green, Gary. 1990. "Resurrecting Polygraph Validation of Self-Reported Crime Data: A Note on Research Methods and Ethics Using the Deer Poacher." *Deviant Behavior* 11(2): 131-137.

Howsen, Roy and Stephen Jarrell. 1987. “Some Determinants of Property Crime.” *American Journal of Economics and Sociology* 46(4): 445-457.

Kierkus, Christopher and Douglas Baer. 2003. “Does the Relationship Between Family Structure and Delinquency Vary According to Circumstances? An Investigation of Interaction Effects.” *Canadian Journal of Criminology and Criminal Justice* 45(4): 405-429.

Menard, Scott and Delbert Elliott. 1990. “Self-Reported Offending, Maturational Reform, and the Easterlin Hypothesis.” *Journal of Quantitative Criminology* 6(3): 237-267.

Paschall, Mallie, Christopher Ringwalt, and Robert Flewelling. 2003. “Effects of Parenting, Father Absence, and Affiliation with Delinquent Peers on Delinquent Behavior Among African-American Male Adolescents.” *Adolescence* 38(1): 15-34.

Rodney, H. Elaine, H. Richard Tachia, and Laxley Rodney. 1990. “The Home Environment and Delinquency: A Study of African American Adolescents.” *Families in Society* 80(6): 551-559.

Sampson, Robert. 1985. “Sex Differences in Self-Reported Delinquency and Official Records: A Multiple-Group Structural Modeling Approach.” *Journal of Quantitative Criminology* 1(4): 345-367.

Sandelin, Bo and Goran Skogh. 1986. “Property Crimes and the Police: An Empirical Analysis of Swedish Data.” *Scandinavian Journal of Economics* 88(3): 547-561.

Simons, Leslie, Ronald Simons, and Rand Conger. 2004. “Identifying the Mechanisms Whereby Family Religiosity Influences the Probability of Adolescent Antisocial Behavior.” *Journal of Comparative Family Studies* 35(4): 547-563.

Videon, Tami. 2002. “The Effects of Parent -Adolescent Relationships and Parental Separation on Adolescent Well-Being.” *Journal of Marriage and Family* 64(2): 489-503.

Williams, James, Charles Ayers, Robert Abbott, J. David Hawkins, and Richard Catalano. 1999. “Racial Differences in Risk Factors for Delinquency and Substance Use Among Adolescents.” *Social Work Research* 23(4): 241-256.

Wyner, Gordon. 1980. “Response Errors in Self-Reported Number of Arrests.” *Sociological Methods and Research* 9(2): 161-177.

Young, Thomas. 1993. “Unemployment and Property Crime” *American Journal of Economics and Sociology* 52(4): 413-415.

Attendance in the NY-Penn Baseball League: Effects of Performance, Demographics, and Promotions

Rodney J. Paul, Kristin K. Paul, Michael Toma, and Andrew Brennan*

ABSTRACT

A regression model is specified for the NY-Penn Baseball League which uses independent variables that consist of demographic, team performance, timing of the game, and promotional variables. Fans appear to treat the NY-Penn League like they would most sports and entertainment activities. The game itself is a normal good and higher population areas attract more fans. Consumers of these games respond favorably to teams that post high win percentages and attend games in greater numbers when popular promotions are offered. The greatest attendance gains due to promotions are shown to be from fireworks shows, concerts, events, and merchandise giveaways.

I. INTRODUCTION

The NY-Penn League is a minor league baseball league which plays a short season (June-Sept). The teams all have affiliations with a Major League Baseball team, which stocks their rosters with drafted players and free agents. These players are typically in the early stages of their career, often right out of high-school or college. In the year studied, 2006, the NY-Penn League consisted of three divisions (McNamara, Pinckney, and Stedler) and a total of 14 teams (Staten Island, Brooklyn, Aberdeen, Hudson Valley, Auburn, Mahoning Valley, State College, Batavia, Jamestown, Williamsport, Tri-City, Oneonta, Lowell, and Vermont).

The main question to be studied in this research is what attracts fans to NY-Penn League games? This is a rather low-level minor league that has teams located across a wide-range of demographics. There are teams located in small towns in Pennsylvania, New York, Maryland, and Vermont and also teams (Brooklyn and Staten Island) located within New York City. Many studies have attempted to determine what drives attendance in major league sports, but questions arise when these factors are put to the test for minor league attendance, especially for such a low-level minor league. What factors really put fans in the seats for these games? Do demographics of the towns/cities matter? Could minor league baseball be an inferior good? Does winning make a difference in the eyes of the

*Rodney J. Paul – St. Bonaventure University, St. Bonaventure, NY 14778, rpaul@sbu.edu, (716) 375-2145, (585) 223-4008; Kristin K. Paul – St. Bonaventure University; Michael Toma – Armstrong Atlantic State University; Andrew Brennan – St. Bonaventure University

fans? Is attendance all about promotions, marketing, and gimmicks? This research attempts to determine what factors influence attendance at a low-level minor league.

A regression model is specified using common factors found in attendance modeling. The data set is relatively rich, as it contains information on all promotions and events for each game for each of the fourteen teams. These promotions and events are translated into dummy variables to allow for an estimation of their impact on attendance. Fan preferences for the performance of the local team, the excitement of the games, and the promotions and events involved in the minor league baseball experience are studied.

Economists have studied fan (consumer) behavior in relation to baseball through a variety of models. Independent variables included in existing models of baseball attendance include population, income per capita, star players, and recent success (Noll, 1974), televised games, quality of the team, and availability of substitutes (Demmert, 1973), expected probabilities of winning a championship (Whitney, 1988), salary structure (Richards and Guell, 1998), turnover in team rosters (Kahane and Shmanske, 1997), and the impact of interleague play (Butler, 2002; Paul, Weinbach, Melvin, 2004).

Nearly all studies of baseball attendance have been conducted for teams at the Major League level. The few studies of minor league attendance include Siegfried and Eisenberg (1980); Branvold, Pan, and Gabert (1997); Bernthal and Graham (2003); and Lee, Ryder and Shin (2003). As compared to the above studies, this inquiry is most like that of Seigfried and Eisenberg (1980) with a focus on regression analysis of actual attendance data, rather than factor analysis of survey data (Lee, Ryder and Shin (2003); Bernthal and Graham (2003)). Both Siegfried and Eisenberg (1980) and Branvold, Pan, and Gabert (1997) study home *season* attendance data for a sample of minor league teams collectively ranging in quality from AAA to rookie leagues. Our analysis differs from these in that we focus on attendance data for *each home game* of the 2006 season for all teams in the NY-Penn League.

We further extend the analysis by focusing on the effects of promotions on game-by-game attendance in the NY-Penn League. Promotions have been considered in the baseball attendance literature almost since the seminal work on the topic in the 1970s. Siegfried and Eisenberg (1980), Hill, Madura and Zuber (1982), and Marcum and Greenstein (1985) represent early examples of attendance research that includes a binary characterization of promotions to indicate the presence of game-day specific price or non-price promotions. MacDonald and Rascher (2000) extended the analysis of promotions from what generally had been a binary characterization along the lines of major/minor promotions to include a continuous variable measuring the value of the promotional item. A more recent contribution by Boyd and Krehbiel (2003) extends the characterization of the promotion variables to include special events that contribute to the entertainment value of the game without being price or merchandise concessions.

The paper proceeds as follows. Section II presents the regression model and the results of the model. Section III discusses the findings and concludes the paper.

II. A REGRESSION MODEL OF ATTENDANCE FOR THE NY-PENN BASEBALL LEAGUE

The data for this study was primarily obtained through the NY-Penn League website on www.minorleaguebaseball.com. The website has a link for each team with the season results that includes the date of the game, the game outcome, and the final score. Data on promotions were obtained from the website of each team, under their promotional schedule. Missing promotional dates were kindly provided by the teams themselves upon e-mail request of the authors. Data on other sources are noted in the variable descriptions below.

The regression model used in this research is a simple Ordinary Least Squares estimation with per game attendance as the dependent variable. Although the model is run in levels, the model was also specified in logs without a change in the significance levels of the independent variables. Therefore, for ease of discussion, the results are discussed in level form to explain the results in terms of the number of fans affected by each of the independent variables. Every game of the 2006 season for each team is included in the model. Very few of the games in the sample were sellouts, so therefore a restricted dependent variable model was not necessary as attendance did not suffer from capacity constraints. Ticket prices are set at the beginning of the season, presumably to maximize profits for the teams. Therefore, ticket prices are assumed to be endogenous within the model and are not included as an independent variable.

The independent variables fall into different categories. To begin, the model is specified using a lag of the dependent variable as an independent variable. Autocorrelation was a concern within the original model, but Durbin H-tests revealed that autocorrelation was not a problem after the inclusion of the lagged dependent variable. The lag is generated per team for the sample, with the first observation lag, per team, assumed to be the same as the first paid attendance of the season. This specification allows for the inclusion of the first observation for each team, but dropping this observation, per team, did not lead to a significant change in the independent variables. To account for the inclusion of this first observation per team, a dummy variable for opening day is also included in the model to allow for any increase in attendance associated with the first home game of the season.

The second category of independent variables is the demographic data of the cities/towns. These variables include population and income per capita. These data were found on www.city-data.com, which presents demographic data for cities and small towns across the United States. Population and income per capita are both expressed as their actual values reported in the demographic data and are not expressed as fractions of actual values. Both population and income per capita are expected to have positive effects on attendance as bigger cities have more potential customers and higher incomes should lead to fans attending more games, if NY-Penn League baseball is a normal good. Noll (1974) and Bruggink and Eaton (1996) found that income has a negative and significant impact on major league attendance. Negative, but insignificant results were found in Coffin (1996). Demmert (1973) obtains positive, but insignificant results for income and major league attendance, while Seigfried and Eisenberg (1980) obtain a similar result for minor league attendance.

The next category of independent variables is related to the performance of the team itself. Win Percentage is calculated by taking the wins of the team and dividing it by the number of games played. This variable is calculated on a running basis per game and the win percentage going into the current game is used as the relevant observation for each data point of attendance. If fans prefer to see winning teams, this variable should have a positive and significant effect on attendance. Higher win percentages led to an increase in attendance for Major League Baseball (Hill, Madura and Zuber, 1982; Rascher, 1996; Bruggink and Eaton, 1996; MacDonald and Rascher, 2000; Coates and Harrison, 2005). Previous studies of minor league attendance find no relationship between team performance and attendance (Seigfried and Eisenberg, 1980; Branvold, Pan and Gabert, 1997; Bernthal and Graham, 2003).

Another aspect of the game that fans may prefer to see is high scoring. Teams that score many runs, or are typically involved in high-scoring games, may be more entertaining for fans to watch. Therefore, runs scored per game are included in the model as a proxy for excitement of the games themselves. Runs scored per game are calculated as a running average dividing the total runs scored by the number of games played. It should be noted that to win games, teams obviously need to score runs. Therefore, the interpretation of results is somewhat complicated by the multicollinearity that is likely to exist between runs scored per game and winning percentage. To account for this, an alternative regression specification is used where total runs scored per game, adding the runs scored per game and the runs allowed per game by the home team, is included as a proxy for expected excitement of the baseball game.

A third category of independent variables is the timing of the game, specifically, the day of the week and the month of the year. The opportunity cost of fans' time is very important and may vary significantly throughout the week depending upon work and other factors. Weeknights are typically more difficult draws for sports teams as most fans need to wake up early for work the next morning. The months of the year may matter in this sample, even though league games are played in the summer months. For example, August usually involves back-to-school activities that could keep fans away from the ballpark. These results have been explored before in Hill, Madura and Zuber (1982), Rascher (1996), and Bruggink and Eaton (1996).

The last category of independent variables addressed in this study is promotions. Promotions vary across teams and are arranged in many different ways. In looking at the lists of promotions for all of the teams, some fell into very neat categories, while others were a bit more difficult to label. In the end, we settled on the following categories, for which dummy variables were created to be used in the regression model: food, beer, merchandise giveaways, free or reduced-price tickets, fireworks, group and family nights, concerts, and other events. Food promotions include either free food or reduced prices on typical baseball stadium fare like hot dogs. Beer promotions include either free beer or reduced prices on beer at the game. Merchandise giveaways include any item that was given to fans for attending the game such as caps, t-shirts, towels, bobble heads, trading cards, etc. There are likely differences between the quality of promotions, but attempts to distinguish between high-value and low-value promotional merchandise did not yield significantly different results. Therefore, a single variable for promotions was settled upon. Free or reduced-price tickets are self-explanatory. Fireworks signify post-game shows, typically on holidays or on weekend nights. Group and family nights are themed-nights or reduced-price nights for specific local groups (Boy Scouts, workplace, and churches) or for families. Concerts are pre- or post-game concerts that are included with the price of admission. The dummy variable “events” is included for all other activities that do not fit into one of the above categories, such as run-the-bases nights, special appearances, etc.

The regression results are shown below for both possible specifications of runs scored (the proxy for excitement of games). Due to a season worth of sellouts, the Lowell Spinners were dropped from the regression model due to lack of variation in their attendance figures. Initial regression results indicated problems with heteroskedasticity within the data, therefore, White’s heteroskedasticity-constant standard errors and co-variance adjustment is used in reporting the results. Coefficient estimates and their t-statistics are noted. *-notation is used to identify statistical significance at the 10 percent (*), 5 percent (**), and 1 percent (***) levels.

Table 1: NY-Penn Baseball League Attendance Model: 2006

Variable	(1) Dependent Var: Attendance	(2) Dependent Var: Attendance	(3) Dependent Var: Attendance
Constant	-1,250.72*** (-3.7795)	-1,157.69*** (-2.8122)	-1,437.97*** (-2.6239)
Attendance_{t-1}	0.6974*** (18.8642)	0.6966*** (18.6240)	0.6982*** (18.9362)
Opening Day	222.4955 (0.9396)	232.9409 (0.9159)	201.8380 (0.8298)
Population	0.0007*** (5.9472)	0.0007*** (5.8762)	0.0007*** (5.9755)
Income Per Capita	0.0395*** (4.3288)	0.0388*** (4.1151)	0.0404*** (4.2491)
Win Percentage	912.8515** (2.1551)	1,010.87* (1.7862)	914.0583** (2.1898)
Runs For		-24.9202 (-0.3150)	
Total Runs			17.4868 (0.4214)
July	54.9445 (0.2990)	45.3793 (0.2437)	66.1656 (0.3660)
August	87.1857 (0.4694)	79.4190 (0.4272)	96.5716 (0.5311)

September	-49.9861 (-0.2384)	-61.2615 (-0.2922)	-35.2155 (-0.1714)
Sunday	-4.0707 (-0.0248)	-2.0327 (-0.0123)	-7.6228 (-0.0460)
Monday	117.2747 (0.7319)	116.8939 (0.7295)	116.4671 (0.7255)
Tuesday	332.2263** (2.0896)	333.3990** (2.0843)	328.3192** (2.0629)
Thursday	500.5214*** (3.1649)	500.6314*** (3.1571)	497.0683*** (3.1314)
Friday	450.3165** (2.5193)	451.2175** (2.5210)	449.2823** (2.5067)
Saturday	58.3382 (0.2896)	57.6275 (0.2860)	56.5477 (0.2816)
Food	-8.5691 (-0.0324)	-14.8022 (-0.0557)	-3.3638 (-0.0127)
Beer	64.2971 (0.2655)	64.8723 (0.2669)	63.0988 (0.2610)
Merchandise	229.5607** (1.9797)	221.2610* (1.9352)	237.2754** (2.0476)
Fireworks	856.9910*** (5.3753)	845.0417*** (5.2635)	870.7227*** (5.3824)
Free Tickets	145.8076 (0.8525)	139.6261 (0.8161)	159.2970 (0.9180)
Group Nights	198.9388 (1.5268)	193.4272 (1.4827)	203.3272 (1.5414)
Concerts	652.8377* (1.6476)	654.3599* (1.6347)	656.5517* (1.6730)
Events	279.6452** (2.0052)	269.2264* (1.8974)	289.6061** (2.0487)
R-squared	0.7711	0.7711	0.7712
Adj. R-squared	0.7602	0.7598	0.7598

The intercept and the lag of attendance were found to be significant at the 1 percent level of significance. The opening day dummy was found to be positive, but not significant. This could be due to the way that the lag of attendance was specified for the first observation or opening day could simply not have as strong of an effect at a lower-level minor league.

The demographic variables took the expected signs and were found to be significant at a 1 percent level. Both population and income per capita were found to have positive effects on attendance. Larger cities/towns tend to attract more fans to the games. In addition, the NY-Penn League appears to be a normal good, even though it is a relatively low minor league, as higher income per capita cities/towns tend to have higher attendances. These results are in contrast with Demmert (1973) and Seigfried and Eisenberg (1980), which did not find significant results. It is also the opposite effect that has been found in some Major League Baseball studies (Noll, 1974; Bruggink and

Eaton, 1996) as income per capita for the NY-Penn league was found to have a positive effect. This could be due to the wide range of income per capita across the teams in this 14-team league.

The team variables also revealed interesting results. First, the success of the team is an important determinant of attendance. Teams with higher win percentages tend to attract more fans. Depending upon the specification, an additional increase in win percentage by 0.100 leads to an increase of around 90-100 fans, which accounts for an approximate 3 percent increase in attendance. This supports previous findings in Major League Baseball (Hill, Madura, and Zuber, 1982; Rascher, 1996; Bruggink and Eaton, 1996; MacDonald and Rascher, 2000; Coates and Harrison, 2005), but contradicts the results found in previous minor league baseball studies (Seigfried and Eisenberg, 1980; Branvold, Pan, and Gabert, 1997; Bernthal and Graham, 2003). This could be because the preferences of minor league fans have changed in recent years, or it could be due to the structure of this study which analyzes game-by-game attendance, instead of aggregating. Runs scored, in either specification, were not found to be significant. Therefore, fans in the NY-Penn league respond to teams that win, but not necessarily to teams that score more often.

The months of the year were found to not be significantly different from each other. The games are played in the summer months and there does not appear to be significant differences between the months. The days of the week were found to have significant differences between them. The omitted dummy variable day was Wednesday. Compared to Wednesday, the dummy variables for Friday, Thursday, and Tuesday were found to be positive and significant. Somewhat surprisingly, Saturday and Sunday were not found to be significant. This could indicate that there are alternative entertainment possibilities in these cities/towns that are available on the weekend that may not be available to local residents during the week (perhaps outdoor activities that tend to take larger allotments of time such as hiking, swimming, biking, etc.).

The results for the promotional dummy variables revealed some expected and some unexpected results. Fireworks had the largest effect on attendance and were found to be significant at the 1 percent level. Games that were followed by post-game fireworks shows attracted around 850 more fans to the game, about 25 percent more than the average attendance per game. Events, such as special appearances by former players or celebrities, shows, or activities such as kids run the bases, led to more than 250 additional fans per game with the results significant at the 5 percent level. Merchandise giveaways were also found to have a positive and significant effect on attendance. Promotional giveaways led to around 220 more fans and were found to be significant at the 5 percent level. Post-game concerts were also found to have a positive and significant effect on attendance.

Food and beer-related promotions were not found to have a significant effect on attendance. Although beer is often considered to be an important component in a fan’s decision to attend a game, fans did not turn out in significant numbers for reduced-price beer. It is likely that some fans do enjoy free or inexpensive beer at games, but this is countered by other fans. For example, reduced price or free beer might cause families to not attend a baseball game. Reduced prices or free food also did not appeal enough to fans to have a significant impact.

Group-related promotions and free or reduced-price tickets were both found to have positive effects on attendance, but were not found to be significant. These promotional activities did not attract enough fans to have a noticeable effect on attendance. Fans may not be overly sensitive to price changes, as tickets are relatively inexpensive to begin with.

In summary of the regression results, it appears that NY-Penn fans respond favorably to winning teams and to certain promotions (fireworks, merchandise, events, and concerts). Demographic variables had the expected results as population and income per capita each had positive effects on attendance. The days of the week were also found to have some interesting effects as weekend days (Saturday and Sunday) were not as popular as expected, but days such as Tuesday, Thursday, and Friday had positive and significant increases in attendance.

III. CONCLUSIONS

The regression results for the NY-Penn League appear to correspond to what we typically believe the demand function for sporting events looks like. Fans will come to the ballpark to see teams that consistently win and they like the promotions that usually accompany a minor league game. The promotions that fans in the NY-Penn league were attracted to the most were fireworks, merchandise giveaways, events, and concerts. Other promotions, such as reduced prices on food, beer, or even the tickets themselves, do not appear to be as popular. Demographic variables appear to matter as areas with greater populations attract more fans and the NY-Penn League appears to be a normal good as higher income per capita areas also attract a greater number of paying customers. The most surprising result is that Saturday and Sunday games were not as popular as would be anticipated, as mid-week games seemed to fare better at the gate than in other leagues.

The most striking result of this research, to the authors, is that the NY-Penn league looks so much like what we envision a typical sports league to be. There does not appear to be the typical concerns that minor league baseball could be an inferior good, nor that winning games does not matter. Short season baseball fans in the NY-Penn league do respond to teams that are winning and the game itself appears to be a normal good. Promotions add extra value, either as a giveaway or some added event, and fans like these items and activities as well. It appears that the NY-Penn League could be modeled like any typical business that offers entertainment activities in the United States.

These results may or may not hold across different leagues of minor league baseball. There are likely to be regional differences, differences in leagues where teams are more or less geographically tied together, and at different levels of minor league play (AAA, AA, A, etc.). There appears to be advantages to studying leagues individually, instead of solely relying upon aggregated attendance, to decipher the similarities and differences that may exist across leagues. This will help researchers better understand the interests of consumers, and, perhaps more importantly, help teams in their quest to offer a better product and thereby maximize profits.

REFERENCES

Bernthal, Matthew J. and Peter J. Graham. 2003. "The Effect of Sport Setting on Fan Attendance Motivation: The Case of Minor League vs. Collegiate Baseball." *Journal of Sport Behavior* 26: (3) pp: 223-239

Boyd, Thomas C. and Timothy C. Krehbiel. 2003. "Promotion Timing in Major League Baseball and the Stacking Effects of Factors that Increase Game Attractiveness." *Sport Marketing Quarterly* 12: (3) pp: 173-183

Branvold, Scott E., David W. Pan, and Trent E. Gabert. 1997. "Effects of Winning Percentage and Market Size on Attendance in Minor League Baseball." *Sport Marketing Quarterly* 6: (4) pp: 35-42

Bruggink, Thomas H. and James W. Eaton. 1996. "Rebuilding Attendance in Major League Baseball: The Demand for Individual Games." *Baseball Economics*: Current Research pp: 9-31

Butler, M. (2002) "Interleague Play and Baseball Attendance," *Journal of Sports Economics*, 3, 4, pp. 320-334.

Coates, Dennis and Thane Harrison. 2005. "Baseball Strikes and the Demand for Attendance." *Journal of Sports Economics* 6: (3) pp. 282-302

Coffin, Donald A. 1996. "If You Build It, Will They Come? Attendance and New Stadium Construction." *Baseball Economics*: Current Research pp: 33-46

Demmert, H.G. (1973) The Economics of Professional Team Sports, Lexington Books, Lexington, MA.

Hill, James Richard, Jeff Madura and Richard A. Zuber. 1982. "The Short Run Demand for Major League Baseball." *Atlantic Economic Journal* 10:2 pp. 31-35

Kahane, L. and Shmanske, S. (1997) "Team Roster Turnover and Attendance in Major League Baseball," *Applied Economics*, 29.

Lee, Soonhwan, Cynthia Ryder, and Hee-Joon Shin. 2003 "An Investigation of Environmental Motivation Factors Among Minor League Baseball (MiLB) Fans." *The Sport Journal* 6: (3)

Marcum, J., and Greenstein, T. (1985). "Factors Affecting Attendance of major League Baseball: A within Season Analysis." *Sociology of Sport Journal* 4: 314-322.

McDonald, Mark and Daniel Rascher. 2000. "Does Bat Day Make Cents? The Effect of Promotions on the Demand for Major League Baseball." *Journal of Sport Management* 14: pp. 8-27

Noll, R.G. (1974) Attendance and price setting, in Government and the Sport Business (ed.) R.G. Noll, Brookings Studies in the Regulation of Economic Activity, Washington, DC.

Paul, R.J., Weinbach, A.P., and Melvin, P. (2004) "The Yankee Effect: The Impact of Interleague Play and the Unbalanced Schedule on Major League Baseball Attendance." *New York Economic Review,* 35.

Rascher, Daniel. “A Test of the Optimal Positive Production Network Externality in Major League Baseball.” *Sports Economics*: Current Research: pp. 33-45

Richards, D.G. and Guell, R.C. (1998) "Baseball Success and the Structure of Salaries," *Applied Economics Letters*, 5, 5.

Siegfried, John J. and Jeff D. Eisenberg. 1980. “The Demand for Minor League Baseball.” *Atlantic Economic Journal* 8: (2) pp: 59-69

Whitney, J.D. (1988) "Winning Games Versus Winning Championships: The Economics of Fan Interest and Team Performance," *Economic Inquiry*, 26.

Behavioral Finance and Football Betting: A Note

Ladd Kochman and Randy Goodwin*

ABSTRACT

Behaviorists argue that investors' fear of regret causes them to favor stocks that are popular and familiar. If bettors share that fear, they are more likely to place wagers on favorites vis-à-vis underdogs. Such a preference would inflate point spreads and possibly explain why underdogs in the National Football League produced a significantly nonrandom wins-to-bets ratio of nearly 52 percent over the 1991-2004 period.

BACKGROUND

Forty years ago, Pankoff (1968) drew an analogy between investing and betting and reasoned that regular profits from the latter activity would represent no less an exception to the theory of efficient markets than those from the former. When his imaginary wagers on National Football League (NFL) games produced only breakeven results, Pankoff concluded that participants in the football betting market—like those in the securities market—are rational profit-maximizers. Since that seminal article, researchers have become increasingly aware of the impact of psychology on the decision-making of individuals. Not only do people sometimes behave in a less-than-rational manner, but those actions can often take predictable forms.

One such pattern was termed the "disposition effect" by behaviorists Shefrin and Statman (1984). It claims that people tend to feel sorrow and grief after having made an error in judgment. This fear of regret will cause investors to avoid selling stocks that have suffered losses. The disposition effect can also explain why investors will prefer popular stocks—rationalizing that any resulting losses would not be peculiar to them. In the context of sports betting, the avoidance of regret could arguably translate into a preference for favorites; losing a bet on a favored team would be easier to accept than doing likewise on an underdog. This "buying pressure" on favorites should cause their "prices" to rise—making underdogs the beneficiaries of inflated point spreads and giving them a betting advantage. The emotional decision to prefer favorites is precisely the kind of irrational behavior that Kochman and Goodwin (2006) argued will create inefficiencies in the sports betting market.

METHODOLOGY

If bettors overbet favorites and bookmakers adjust point spreads up, profit opportunities should accrue to those wagering on underdogs. To test that hypothesis, we placed imaginary bets on all

*Kennesaw State University

underdogs in the National Football League (NFL) during the 1991-2004 seasons as identified by Las Vegas oddsmakers and reported by the Internet site for a popular football newsletter: www.goldsheet.com. Not unlike Gandar et al. (2001), our resulting wins-to-bets ratios were then tested for nonrandomness per Equation (1).

$$(1)\quad Z = \frac{(W/B - 0.50)}{\{[(0.50)(1 - 0.50)]/B\}^{1/2}}$$

Z = statistic for testing the breakeven null hypothesis

W = number of winning bets

B = total number of bets.

RESULTS

By wagering exclusively and consistently on underdogs in the NFL over the 14 consecutive seasons through 2004, we achieved a significantly nonrandom wins-to-bets ratio of 51.90 percent. It is also evident from Table 1 that winning bets on underdogs outnumbered losing bets in 11 (or roughly 80 percent) of those years. Cumulatively, the first half of our 14-year measurement period generated a significantly nonrandom wins-to-bets ratio of 52.28 percent while the cumulative W/B ratios for years 12, 13 and 14 were likewise significantly nonrandom.

Table 1

Wins-to-bets ratios for NFL underdogs (1991-2004)

Season	Bets	Wins	Wins-to-bets	Cumulative Wins-to-bets
1991	226	112	49.56%	
1992	223	116	52.02%	50.78%
1993	231	125	54.11%	51.91%
1994	224	118	52.68%	52.10%
1995	244	130	53.28%	52.35%
1996	240	121	50.42%	52.02%
1997	234	126	53.85%	52.28%*
1998	239	109	45.61%	51.42%
1999	245	130	53.06%	51.61%
2000	247	132	53.44%	51.81%
2001	245	125	51.02%	51.73%
2002	260	144	55.38%*	52.06%**
2003	256	126	49.22%	51.83%**
2004	258	136	52.71%	51.90%**
Totals	3372	1750	51.90%**	

*significant at $p < 0.10$
**significant at $p < 0.05$

CONCLUSIONS

Random returns from wagers are generally interpreted to mean that the market for bets is efficient. While our results suggest that the market for bets may be less than efficient, regular profits remain elusive. Our wins-to-bets ratio of 51.9 percent stops short of the 52.4-percent[1] mark needed for nonrandom profitability. Behaviorists could argue that our results imply that individuals behave irrationally—creating profit opportunities for those who are alert to the inclination of bettors to eschew underdogs in favor of favorites. While the temptation might be to view market efficiency and behavioral finance as competing explanations, they may, in truth, be complementary. Irrational (rational) behavior would account for the (absence of) mistakes that, in turn, produce abnormal (normal) returns. Thus, it seems fair to say that our nonrandom returns not only illuminate an attractive strategy for bettors but also energize the argument that individuals gravitate toward popular choices to avoid the fear of regret.

While no distinction between heavy and slight underdogs was made in this study, we would encourage future researchers to consider doing so. It could be that bettors' alleged aversion to (losing wagers on) underdogs is positively related to the size of the spread and that heavy underdogs are therefore the likeliest beneficiaries of inflated points. To the extent that a clearer definition of underdogs facilitates higher wins-to-bets ratios, our behaviorist approach to betting may finally pay off … literally!

ENDNOTES

1. Having to risk $11 to win $10, bettors must win 11 of 21 bets (or 52.4 percent) to break even.

REFERENCES

Gandar, J., R. Zuber, and R. Lamb. "The Home Field Advantage Revisited: A Search for the Bias in Other Sports Betting Markets." Journal of Economics and Business, July-August, 2001.

Kochman, Ladd, and Randy Goodwin. "An *Under* Bias in the Football Betting Market: Fact or Fiction?" New York Economic Review, Fall 2006.

Pankoff, Lyn. "Market Efficiency and Football Betting." *Journal of Business*, April 1968.

Shefrin, Hersh, and Meir Statman. "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence." *Journal of Finance*, 1984.

REFEREES

1. Darius Conger
2. Lester Hadsell
3. Elia Kacapyr (2)
4. Kent Klitgaard
5. Thomas Kopp
6. Frank Musgrave
7. David Pate
8. David Ring
9. Ted Schmidt
10. Martha Wojtowyz

The New York State Economics Association (NYSEA)
59th Annual Conference

2006

At Siena College,
Loudonville, NY

Friday and Saturday, September 29-30, 2006

Friday, September 29th
6:00 pm - 7:30 pm – Friday Reception, Crowne Plaza, Downtown Albany, NY

Saturday, September 30th
8:00 am - 4:30 pm – Conference: Siena College, Sarazen Student Union

8:00 am - 8:45 am – Registration/Continental Breakfast
8:45 am - 10:00 am – Welcome and Plenary Session
10:00 am - 10:15 am – Coffee Break
10:15 am - 11:45 am – Technical Session I
12:00 pm - 1:30 pm – Luncheon
1:45 pm - 3:15 pm – Technical Session II
3:15 pm - 3:30 pm – Coffee Break
3:30 pm - 4:45 pm – Technical Session III
5:00 pm - 6:00 pm – Business Meeting (All are Welcome)

Conference Sessions

8:00 am - 8:45 am - Registration/Continental Breakfast, Maloney Great Room (SSU 240)

8:45 am - 10:00 am - Welcome and Plenary Session, SSU 243
Session Chair: Kris Principe (Canisius College)

Frank W. Musgrave, Professor of Economics, Ithaca College, and Author of *The Economics of U.S. Health Care Policy: The Role of Market Forces* (2006)

A. Dale Tussing, Professor of Economics, Syracuse University, and Author of *How Ireland Cares: The Case for Health Care Reform* (2006)

10:00 am - 10:15 am - Coffee Break, Maloney Great Room (SSU 240)

10:15 am - 11:45 am - Technical Session I

Session 1A: Macroeconomics, SSU 241
Session Chair: Tom Kopp (Siena College)

Ordean Olson (Nova Southeastern University)	The Federal Funds Rate and Transmission of Monetary Policy: Measuring Monetary Policy Effectiveness *Discussant:* Tom Kopp (Siena College)
William T. Ganley (Buffalo State College)	An Early Heterodox Theory of the Business Cycle *Discussant:* Wade Thomas (SUNY - Oneonta)
David W. Ring (SUNY - Oneonta)	Is Fiscal Policy Countercyclical? *Discussant:* W. Scott Trees (Siena College)
Manimoy Paul (Siena College)	Balance of Payments for a State Considering Variations in Cost of Living Across States *Discussant:* Elia Kacapyr (Ithaca College)

Session 1B: Development, Security, and Regional Challenges, Dot Com Room (SSU 207)
Session Chair: Steve Onyeiwu (Alleghany College)

Steve Onyeiwu (Allegheny College)	Determinants of Success Amongst Small Enterprises: Evidence from the Tooling and Machining Cluster in Northwestern Pennsylvania *Discussant:* Robert Jones (Skidmore College)
Bala Veeramacheneni and Richard Vogel***** (SUNY - Farmingdale) and E.M. Ekanayake (Bethune - Cookman College) ***Presenter**	Impact of IT Outsourcing on Indian Agriculture *Discussant:* Mary Ellen Mallia (Siena College)
John M. Polimeni (Albany College of Pharmacy); Raluca Iorgulescu Polimeni, Richard L. Shirey***** and, W. Scott Trees (Siena College) ***Presenter**	Protecting the Global Food Supply from a Terrorist Attack *Discussant:* Alfred Lubell (SUNY - Oneonta)
Pellegrino Manfra (Queensborough CC/City University of New York)	The Rise of Foreign Direct Investment in the United States: Challenges and Opportunity

Discussant: John Piccione (JWP Consulting)

Session 1C: Healthcare and Disease, Molinari Room
Session Chair: Steve Neun (Utica College)

Kris Principe***** (Canisius College) and E. Kathleen Adams (The Rollins School of Public Health at Emory University) ***Presenter**	The Impact of Hospital Competition and Market Structure on a Private Hospital's Medicaid Share
	Discussant: Frank Musgrave (Ithaca College)
Bruce Carpenter (Mansfield University) and Stephen P. Neun***** (Utica College) ***Presenter**	The Factors that Contribute to the Prevalence of Heart Disease in New York State: A Production Function Approach
	Discussant: Joseph Cheng (Ithaca College)
Duane A. Matcha (Siena College)	Are Health Economists asking the Right Questions? A Medical Sociology Perspective *Discussant:* Bryan McCannon (Elmira College)
A. Dale Tussing (Syracuse University) and Maev-Ann Wren (Trinity College, Dublin)	Reunions Remuneration in Ireland: Private Patients in Public Hospitals
	Discussant: Duane A. Matcha (Siena College)

Session 1D: Teaching and Learning Local and State Economy, SSU 315
Session Chair: Ning Fu (SUNY - Potsdam)

Matthew Siver (Student, SUNY - Potsdam)	New York State Electricity
	Discussant: Michael McAvoy (SUNY- Oneonta)
Henry Sieg (Student, SUNY - Potsdam)	State and Local Taxes
	Discussant: Ed Portugal (SUNY - Potsdam)
Florence Shu (SUNY - Potsdam)	Teaching and Learning Local and State Economy
	Discussant: Richard Deitz (Federal Reserve Bank, Buffalo Branch)

12:00 pm - 1:30 pm – Luncheon, SSU 243
Luncheon Speaker: Dr. Stephen Sheppard
James Phinney Baxter III Professor of Public Affairs, Williams College
“The Causes and Consequences of Global Urban Expansion”

1:45 pm - 3:15 pm - Technical Sessions II

Session 2A: Winners, Losers, and Quality, SSU 241
Session Chair: Edward Howe (Siena College)

Mark Gius (Quinnipiac University)	Using Panel Data to Estimate the Economic Determinants of CEO Compensation *Discussant:* William O'Dea (SUNY - Oneonta)
Gee San and Wen-Jhan Jane* (Graduate Institute of Industrial Economics, National Central University, Taiwan) ***Presenter**	Wage Dispersion and Organizational Performance: Evidence from a Small Professional Baseball Team in Taiwan *Discussant:* Stephen P. Neun (Utica College)
Della Lee Sue (Marist College)	Application of an Unemployment Index to the Phillips Curve *Discussant:* Manimoy Paul (Siena College)
Joseph Cheng***** and Don Simmons (Ithaca College) ***Presenter**	A Model on Optimal Quality Improvement for Small Parts *Discussant:* Gary Clendenen (Siena College)

Session 2B: Welfare and Well-Being, Molinari Room
Session Chair: Bryan McCannon

Elia Kacapyr (Ithaca College)	An Index of American Progress *Discussant:* Richard L. Shirey (Siena College)
Mary Ellen Mallia (Siena College)	The Role of Income in Subjective Well-Being *Discussant:* David W. Ring (SUNY - Oneonta)
Joseph G. Eisenhauer (Canisius College)	Severity of Illness and the Welfare Effects of Moral Hazard *Discussant:* Duane A. Matcha (Siena College)
Bryan McCannon (Elmira College)	Punishing Repeat Offenders *Discussant:* Eric Kacapyr (Ithaca College)

Session 2C: Panel on Geographic Information Systems (GIS) in Economic Research and Teaching, SSU 315
Session Chair: James Booker (Siena College)

Robert Jones (Skidmore College)
James Booker (Siena College)
Craig Roberts (Canisius College)

3:15 pm - 3:30 pm - Coffee Break, Maloney Great Room, (SSU 240)

3:30 pm - 4:45 pm - Technical Sessions III

FALL 2007

Session 3A: Environment and Attributes Markets, SSU 241
Session Chair: William Kolberg (Ithaca College)

Robert P. Culp (Penn State University Lehigh Valley)	Predicting a Home's Time on Market: The Influence of Environmental and Home Attributes *Discussant:* Patrick Meister (Ithaca College)
Ben Hoen (Bard College)	Impacts of Windmill Visibility on Property Values in Madison County, New York *Discussant:* William Kolberg (Ithaca College)
William Kolberg (Ithaca College)	Advertising Elasticity, Advertising Cross- Elasticity, and Market Relationship *Discussant:* Joseph G. Eisenhauer (Canisius College)

Session 3B: Property and Performance, Dot Com Room, (SSU 207)
Session Chair: David Ring (SUNY - Oneonta)

Shuo Chen* (SUNY - Geneseo), John Conlon and William Shugart (University of Mississippi) ***Presenter**	Open Source Software: A Solution to the Intellectual Property Dilemma? *Discussant:* Bryan McCannon (Elmira College)
Fahrettin Dingil (SUNY - Alfred)	Testing Convergence of Property Crimes across the States *Discussant:* Wade L. Thomas (SUNY - Oneonta)
William O'Dea***** and David Ring (SUNY - Oneonta) ***Presenter**	The Impact of APLIA on Student Performance in Intermediate Microeconomic Theory *Discussant:* Chris Fedoryshyn (Siena College)

Session 3C: Regional Development, Molinari Room
Session Chair: Richard Vogel (SUNY – Farmingdale)

Craig Rogers (Canisius College)	Shifting Patterns of Spatial Employment Clusters in the Buffalo-Niagara Falls, New York Metropolitan Statistical Area (MSA): 1990-2000 *Discussant:* Florence Shu (SUNY - Potsdam)
Richard Vogel (SUNY - Farmingdale)	Incubators and the New York State Economy *Discussant:* W. Scott Trees (Siena College)
Mehmet Odekon and Robert Jones***** (Skidmore College) ***Presenter**	Factors Affecting New York State Out-Migration

Discussant: Alfred Lubell (SUNY - Oneonta)

5:00 pm - 6:00 pm - Business Meeting (All are Welcome), Dot Com Room-SSU 207

2006 List of Participants

James F. Booker
Professor of Economics
Siena College
jbooker@siena.edu

Shuo Chen
Assistant Professor of Economics
SUNY - Geneseo
chens@geneseo.edu

Joseph Cheng
Associate Professor of Finance and International Business
Ithaca College
cheng@ithaca.edu

Gary W. Clendenen
Associate Professor of Quantitative Business Analysis
Siena College
gclendenen@siena.edu

Robert P. Culp
Assistant Professor of Economics
Penn State Lehigh Valley
DrCulp@rcn.com; DrCulp@psu.edu

Richard Deitz
Sr. Economist
Federal Reserve Bank, Buffalo Branch
Richard.dietz@ay.frb.org

Fahrettin Dingil
Assistant Professor
SUNY - Alfred
dingilf@alfredstate.edu

Joseph G. Eisenhauer
Professor of Economics
Canisius College
eisenhauer@canisius.edu

Chris Fedoryshyn
Student
Siena College
christopher.fedoryshyn@siena.edu

Ning Fu
Student
SUNY - Potsdam
Fu73@potsdam.edu

William T. Ganley
Professor of Economics and Finance
Buffalo State College
ganleywt@buffalostate.edu

Mark Gius
Professor of Economics
Quinnipiac University
gius@quinnipiac.edu

Mr. Ben Hoen
Bard Center for Environmental Policy
Bard College
benhoen2@earthlink.net
Edward Howe
Professor of Economics
Siena College
howe@siena.edu

Mr. Wen-Jhan Jane
Graduate Institute of Industrial Economics
National Central University, Taiwan
s0444003@cc.ncu.edu.tw; krisenwerk@gmail.com

Robert Jones
Associate Professor & Chair of Economics
Skidmore College
rjones@skidmore.edu

Elia Kacapyr
Department of Economics
Ithaca College
kacapyr@ithaca.edu

William Kolberg
Department of Economics
Ithaca College
kolberg@ithaca.edu

Tom Kopp
Assoc. Professor of Finance
Siena College
kopp@siena.edu

Alfred Lubell
Interim Dean of Graduate Studies, Continuing Ed, and Summer Sessions
SUNY - Oneonta
lubellam@oneonta.edu
Mary Ellen Mallia
Lecturer in Economics
Siena College
mmallia@siena.edu

Pellegrino Manfra
Associate Professor, Social Sciences
Queensborough CC
City University of New York
rinomanfra@aol.com

Duane A. Matcha
Professor of Sociology
Siena College
matcha@siena.edu

Mike McAvoy
Associate Professor of Economics
SUNY – Oneonta
MCAVOYM@oneonta.edu

Bryan McCannon
Assistant Professor of Economics
Elmira College
bmccannon@elmira.edu

Patrick Meister
Assistant Professor of Economics
Ithaca College
pmeister@ithaca.edu

Frank Musgrave
Professor of Economics
Ithaca College
musgrave@ithaca.edu

Stephen P. Neun
Professor of Economics
Utica College
sneun@utica.edu

William O'Dea
Professor of Economics
SUNY - Oneonta
ODEAWP@oneonta.edu

Dr. Ordean Olson
Associate Professor
Nova Southeastern University
olson@nova.edu

Steve Onyeiwu
Assistant Professor of Economics
Allegheny College
sonyeiwu@allegheny.edu

Manimoy Paul
Assistant Professor of Quantitative Business Analysis
Siena College
mpaul@siena.edu

John Piccione
Investment Advisor, Adjunct Instructor
JWP Consulting
jwp.consulting@att.net

Ed Portugal
Associate Professor of Economics and Department Chair
SUNY – Potsdam
portugej@potsdam.edu

Kris Principe
Economics/Finance
Canisius College
principk@canisius.edu

David W. Ring
Associate Professor of Economics SUNY - Oneonta
ringdw@oneonta.edu

Craig Rogers
Assistant Professor, Department of Economics & Finance
Canisius College
rogersc@canisius.edu

Gautam Sethi
Assistant Professor of Economics, Center for Environmental Policy
Bard College
sethi@bard.edu

Stephen C. Sheppard
Professor of Economics
Williams College
Stephen.C.Sheppard@williams.edu

Richard L. Shirey
Assoc. Professor of Economics
Siena College
shirey@siena.edu

Florence Shu
Associate Professor of Economics
SUNY - Potsdam
shufp@potsdam.edu

Henry William Sieg
Student
SUNY - Potsdam
Sieg82@potsdam.edu

Matthew Siver
Student
SUNY - Potsdam
MSL21_25@yahoo.com

Della Lee Sue
Assistant Professor of Economics
Marist College
della.lee.sue@marist.edu

Wade L. Thomas
Professor and Associate Dean/Economics & Business
SUNY - Oneonta
THOMASWL@oneonta.edu

W. Scott Trees
Professor of Economics
Siena College
trees@siena.edu

A. Dale Tussing
Professor of Economics
Syracuse University
Tussing@syr.edu

Richard Vogel
Department of History, Economics, and Politics
SUNY - Farmingdale
richard.vogel@farmingdale.edu