New York State Economics Association

Proceedings of the New York State Economics Association

Volume 2

62nd Annual Meeting
Ithaca, New York
October 9-10, 2009

Editor: Richard Vogel

New York State Economics Association
Founded 1948

2008-2009 Officers

President:
William Kolberg, Ithaca College

Vice-President:
Jeffrey Wagner, Rochester Institute of Technology

Secretary:
Michael McAvoy, SUNY-Oneonta

Treasurer:
David Ring, SUNY-Oneonta

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

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

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

Web Coordinator:
Wade L. Thomas, SUNY-Oneonta

Board of Directors:
Cynthia Bansak, Saint Lawrence University
Robert P. Culp, Dalton State College
Richard Dietz, Federal Reserve Bank of New York – Buffalo Branch
Natsuko Iwasaki, SUNY-Farmingdale
William Kolberg, Ithaca College
Patrick Meister, Ithaca College
Peter Pasqualino, Fulton/Montgomery Community College
Manimoy Paul, Siena College
F. Pan Shu, SUNY Potsdam
Della Lee Sue, Marist College
Richard Vogel, SUNY-Farmingdale
Jeffrey Wagner, Rochester Institute of Technology

Published, May 2010, revised December 2010.

Table of Contents

Forecasting the New York State Economy with "Terraced" VARs and Coincident Indices

Eric Doviak[*] and Sean MacDonald[**]

ABSTRACT

This paper introduces "Terraced" Vector Autoregressive (VAR) models, an innovative twist on traditional VAR modeling, which allows the econometrician to simultaneously forecast both exogenous and endogenous variables and the confidence intervals around those forecasts. In an application of our Terraced VAR framework, we have estimated coincident indices of economic activity for the United States, New York State and the six largest metropolitan areas of New York State and incorporated them into Terraced VARs, which forecast the unemployment rate, total non-farm employment, real wages and average hours worked in manufacturing in those regions.

I. INTRODUCTION

Forecasting regional economic variables poses a difficult task because the data series often exhibit negative first-order serial correlation (i.e. the series looks "jagged" or "saw-toothed"). Perhaps more importantly, the frequent fluctuations in the series make it difficult to discern whether a one-month increase in the unemployment rate or a one-month decrease in non-farm employment represent a deterioration of local economic conditions or simply represent the natural fluctuations of the series. Figures 1 and 2, which plot Rochester's unemployment rate and non-farm employment level, illustrate the point.

To overcome this difficulty, we used Kalman's (1960) filtering algorithm to estimate a coincident index of economic activity from the unemployment rate, non-farm employment, average hours worked in manufacturing and real wages. Such an index of the Rochester economy is plotted in Figure 3.

Because the index does not exhibit negative first-order serial correlation, it follows a simple autoregressive process, which is relatively easy to forecast. And because the coincident index was estimated from the economic variables of interest to us, it is highly correlated with those variables and allows us to forecast them with more accuracy than we could achieve if we had not incorporated the coincident index into our forecasting models.

The reader of our reports will also appreciate the visual "smoothness" of the coincident index. Because the index is smooth, the reader can quickly discern the trend in regional economic activity.

This paper introduces and defines a methodology for constructing a coincident index of economic indicators that can estimate the current state of the economies of the United States, New York State and the six largest metropolitan statistical areas (MSAs) of New York State[1]. In explaining what a coincident

[*] New York State Banking Department, eric@doviak.net
[**] New York City college of Technology, City University of New York, smacdonald@citytech.cuny.edu

index is, the paper also describes the coincident indices of national, state and regional economic activity that other economists have developed.

We did not develop coincident indices for their own sake. We developed them for the purpose of forecasting variables that depend on the state of the economy. To that end, we incorporated the coincident indices into "terraced" vector autoregressive (VAR) models, which can be used to forecast variables of interest to policymakers, such as the unemployment rate and rates of foreclosure on residential mortgages.

The "terraced" VAR methodology is our primary contribution to the economic literature. Unlike traditional VAR forecasting, "terraced" VARs do not require exogenous and endogenous variables to be forecast in two separate steps. Instead, the exogenous and endogenous variables are forecast simultaneously, which allows the econometrician to obtain confidence intervals that depend only on the respective "predictor" variables.

For example, suppose that the New York City coincident index depends on past values of the national coincident index, but the reverse is not true. In such a simple, two-variable "terraced" VAR, the confidence interval around the forecast of the New York City index would depend on the residual variance of its regression equation, the residual variance of the national index's regression equation and the covariance between them. By contrast, the confidence interval around the forecast of the national index would only depend on the residual variance of its own regression equation. Consequently, the "terraced" VAR methodology allows us to simultaneously compute the appropriate confidence intervals around both the New York City and national coincident indices.

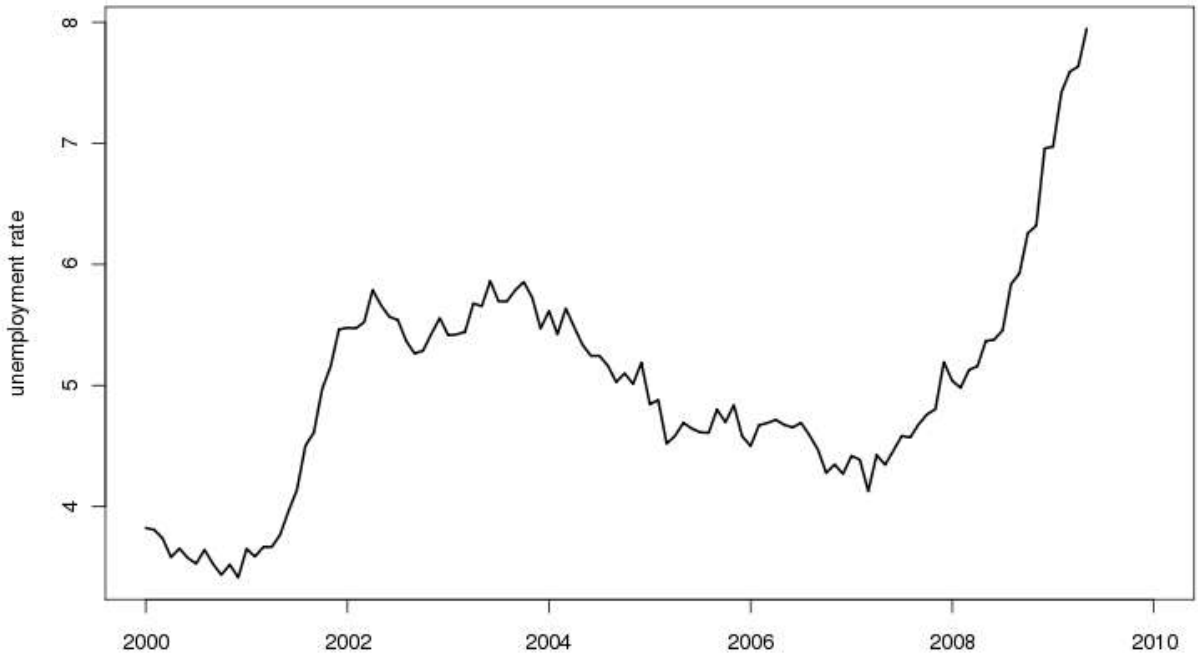
Thus, the paper seeks to extend and build upon the existing literature on state-level coincident indices, while introducing new applications, such as incorporating the coincident indices into "terraced" VARs to forecast "key economic variables." The indices can therefore help policymakers anticipate the future course of key economic trends. For example, the coincident indices can be used to forecast unemployment rates, total employment as well as indicators of banks' financial health (e.g. foreclosure filing rates and non-current loans).

II. COINCIDENT INDICES

1. Survey of Literature

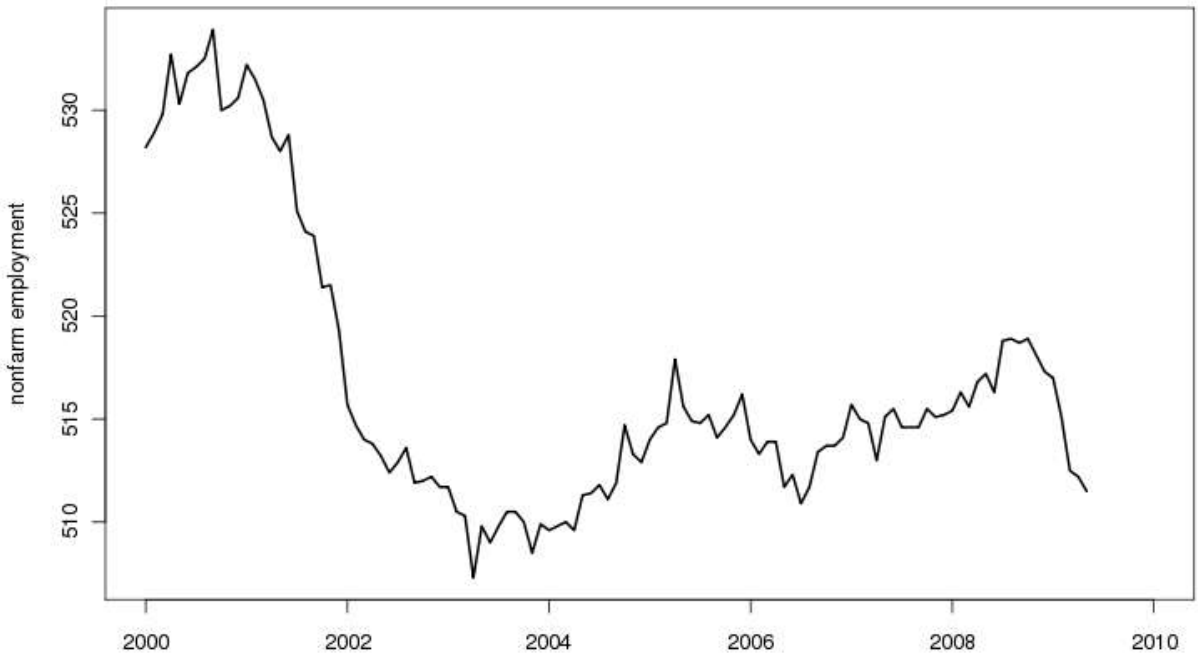
In common speech, people often refer to the "state of the economy" with statements such as: "The economy is strong," or "The economy is in bad shape right now." As a question of measurement however, one must wonder what they are referring to. Are they referring to growth of real Gross Domestic Product (GDP)? Are they referring to growth of real household income? Are they referring to job growth? Are they referring to the unemployment rate?

Figure 1: Rochester Unemployment Rate (seasonally adjusted)



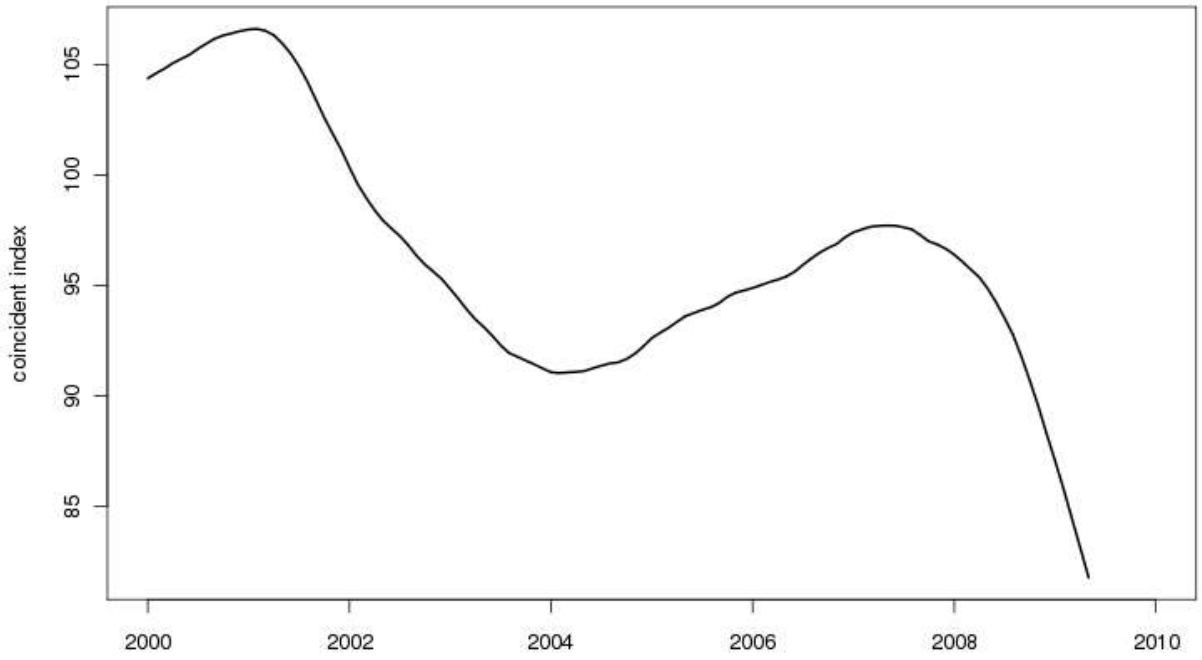
Note: The series depicted above ends in May 2009.

Figure 2: Rochester Nonfarm Employment (seasonally adjusted)



Note: The series depicted above ends in May 2009.

Figure 3: Coincident Index of the Rochester Economy



Note: The data used to generate the coincident index depicted above ends in May 2009.

Assuming that they are referring to some combination of those measures, how then do we evaluate a period in which real GDP grows, but the unemployment rate rises? Is the state of the economy "strong" or "weak?"

Obviously, one can easily define a weighted average of such measures, but what are the appropriate weights to apply to each measure? Should those weights be constant over time? One can easily imagine many scenarios in which they should not be constant over time. For example, if the economy is experiencing a structural shift, in which the manufacturing sector shrinks as the service sector grows, then it hardly makes sense to apply the same weight to manufacturing sector employment over time.

In addressing such questions, Stock and Watson (1989) made an original and valuable contribution to the econometric literature by observing that each of the measures described above depends on the underlying "state of the economy" and the particular characteristics that define each measure. They then used Kalman's (1960) filtering algorithm to identify the unobservable "state of the economy" in a single index, called the "coincident index."

Visually significant decreases in Stock and Watson's original coincident index matched the beginnings and endings of recessions, as defined by the National Bureau of Economic Research (NBER). Visually significant increases, of course, corresponded to economic expansions. Stock and Watson also used a Vector Auto-Regression (VAR) model to forecast changes in the coincident index. The forecast of the coincident index is called the "leading index."

More recently, two economists at the New York State Division of the Budget, Megna and Xu (2003), used Stock and Watson's methodology to develop coincident and leading indices for the New York State economy, which they then used to forecast changes in state tax revenue.

There is major difference between Stock and Watson's index and Megna and Xu's index however. Specifically, Stock and Watson were interested in developing an index based on co-movements in several macroeconomic time series, whereas Megna and Xu were interested in forecasting state tax revenue. Consequently, they used different data series to develop their indices.[1]

Economists at the Federal Reserve Bank of New York have also developed a coincident index of the New York State economy. Orr et al. (1999) developed such an index to predict changes in the regional economy that do not necessarily coincide with national trends.[2]

A recent update of their index conducted by Bram et al. (2009) suggests that the New York regional economy remained far more resilient than the national economy (in the sense that the New York region entered the current recession several months later than the nation overall). They also found that New York State's economic activity peaked in February 2008, while New York City's continued to expand through June 2008. They also note that New York City's economy has already experienced a "steeper downturn than a number of metropolitan areas in upstate New York."

Such trends also appear in our own indices. Our indices suggest that the United States entered the current economic recession almost one year before the economies of New York State and New York City. Specifically, our indices date the national recession as beginning in March 2007, while New York State and New York City entered recession in February 2008 and March 2008 respectively.

More broadly, Crone and Clayton-Matthews (2005) applied Stock and Watson's methodology to estimate a consistent set of coincident indicators for all 50 states. They developed their indices because the considerable lag with which the Bureau of Economic Analysis releases Gross State Product (GSP) series inhibits timely monitoring of state economic trends. In developing their index, they also found that the indices are a useful for comparing "the length, depth, and timing of recessions at the state level."

The Federal Reserve Bank of Philadelphia argues that Crone and Clayton-Matthews' state coincident indices are comparable across states because they are generated from a consistent set of models and variables, but we remain skeptical. Specifically, we suspect that some of the indices may appear more volatile than others because the parameter values in their filters differ across states.

Even if they are correct however, the fact that there is less data available to us at the sub-state level than there is at the national and state level would prevent us from constructing an index of regional economic activity that is comparable to a state or national index unless we were willing to discard valuable data.

Consequently, the index that we develop for one region is not directly comparable to another's because the coincident indices do not have a common unit of measurement. New York City's index is measured in "New York City coincident units," while Rochester's index is measured in "Rochester coincident units."

2. Data Used to Estimate the Coincident Index

With two exceptions, the data that we used to estimate the coincident index were the same four data series that Crone and Clayton-Matthews (2005) used to develop coincident indices for all 50 states. Three of the four series that Crone and Clayton-Matthews used – nonagricultural payroll employment, the unemployment rate and average hours worked in manufacturing – are monthly data series from the US Bureau of Labor Statistics.

Although the unemployment rate is generally identified as a lagging indicator of economic activity, its use here as a variable designed to contribute to obtaining an estimate of the *current* state of the economy, is consistent with the models established in the earlier works of Stock and Watson ((1989), Orr, et.al. (1999), Megna and Xu (2003), and Crone and Clayton-Matthews (2005). Only after estimating the coincident index is the index itself forecast.

Further, we reasoned that housing starts would not prove to be a strong predictor of the state of the economy, as the period of estimation here includes the period during which the housing bubble in the U.S. and New York State burst, and housing starts were actually very weak from January 2007 through mid-2009. An initial evaluation of forecast accuracy over the 1984-2008 period, with housing starts included, confirmed this view. Unfilled orders for durable goods proved to be a significantly stronger predictor overall.

Because the fourth series that Crone and Clayton-Matthews used, U.S. Bureau of Economic Analysis' quarterly real wages and salary disbursements, is not available at the sub-state level, we calculated real wages by adjusting the U.S. Bureau of Labor Statistics' total wage series for inflation with the U.S. consumer price index. For consistency, we also used the total wage series in our estimates of the U.S. and New York State coincident indices.

The fact that the total wage series is quarterly while the other series are monthly, does not pose a problem. In their paper, Crone and Clayton-Matthews show that interpolation is not necessary to handle mixed monthly and quarterly series. Instead, they show that the Kalman Filter can be adapted to handle any missing data simply by altering the dimensions of the relevant matrices. In fact, Petris' (2009) "dlm" package for R (R Development Core Team, 2008), which we used to implement the Kalman Filter, is designed to handle missing data.

In a second departure from Crone and Clayton-Matthews' framework, we did not use average hours worked in manufacturing in our sub-state coincident indices because the series is not available at the sub-state level. We also did not use the real wage series in our estimates of the Rochester and Buffalo-Niagara Falls coincident indices because the extreme volatility of those series inhibited our ability estimate the parameters of the Kalman Filter when we tried to include them.

Additionally, we used the US Census Bureau's X-12-ARIMA seasonal adjustment software (2007) to seasonally adjust the New York State "average hours" series, all of the quarterly wage series and the sub-state unemployment rates.

Finally, to account for serial correlation in the data series, we took differences of the unemployment rate series and log differences of the other series. Differencing also enabled us to overcome the problems associated with the change in industrial classification from SIC to NAICS, which was implemented by the U.S. Bureau of Labor Statistics in 1990. Since both classifications attempted to measure the same economic variable, we assumed that the log-differences of the two classifications are highly correlated with each other and switched from SIC to NAICS at the earliest possible point (i.e. from 1990 onward).

3. Estimation of the Coincident Index

To implement the Kalman Filter, we used a structure based loosely on the one used by Megna and Xu (2003)[3]. Specifically, we used the following structure for the measurement equation:

$$\begin{bmatrix} \Delta unemp_t \\ \Delta \ln(emp)_t \\ \Delta \ln(wages)_t \\ \Delta \ln(hours)_t \end{bmatrix} = \begin{bmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_0 & \beta_1 & 0 & 0 \\ \gamma_0 & \gamma_1 & \gamma_2 & 0 \\ \delta_0 & \delta_1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta c_t \\ \Delta c_{t-1} \\ \Delta c_{t-2} \\ \Delta c_{t-3} \end{bmatrix} + \begin{bmatrix} u_{unemp,t} \\ u_{emp,t} \\ u_{wages,t} \\ u_{hours,t} \end{bmatrix} \quad (1)$$

and the following structure for the transition equation:

$$\begin{bmatrix} \Delta c_t \\ \Delta c_{t-1} \\ \Delta c_{t-2} \\ \Delta c_{t-3} \end{bmatrix} = \begin{bmatrix} \phi_0 & \phi_1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta c_{t-1} \\ \Delta c_{t-2} \\ \Delta c_{t-3} \\ \Delta c_{t-4} \end{bmatrix} + \begin{bmatrix} v_t \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad (2)$$

where Δc_t is the change in the coincident index at time t and the residuals $u_{i,t}$ and v_t are assumed to have zero mean and constant variance.

As mentioned in Section II.2, average hours worked in manufacturing is not available at the sub-state level, so the last row of equation 1 does not appear in the local economies' measurement equations. We also used a slightly different structure for the measurement and transition equations of the Rochester and Buffalo-Niagara Falls Metropolitan Statistical Areas (MSAs). As mentioned in Section II.2, we did not use the real wage series in our estimates of their coincident indices. We also found that a different structure worked better for their economies. Specifically, with the data for the Rochester and Buffalo-Niagara Falls MSAs we used the following structure for the measurement equation:

$$\begin{bmatrix} \Delta unemp_t \\ \Delta \ln(emp)_t \end{bmatrix} = \begin{bmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_0 & \beta_1 & \beta_2 & 0 \end{bmatrix} \begin{bmatrix} \Delta c_t \\ \Delta c_{t-1} \\ \Delta c_{t-2} \\ \Delta c_{t-3} \end{bmatrix} + \begin{bmatrix} u_{unemp,t} \\ u_{emp,t} \end{bmatrix} \quad (3)$$

and the following structure for the transition equation:

$$\begin{bmatrix} \Delta c_t \\ \Delta c_{t-1} \\ \Delta c_{t-2} \\ \Delta c_{t-3} \end{bmatrix} = \begin{bmatrix} \phi_0 & \phi_1 & \phi_2 & \phi_3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta c_{t-1} \\ \Delta c_{t-2} \\ \Delta c_{t-3} \\ \Delta c_{t-4} \end{bmatrix} + \begin{bmatrix} v_t \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (4)$$

Although the transition equations are designed to yield a single index, Petris' (2009) "dlm" package, which we used to implement the Kalman Filter, does not output a single index. Instead, it outputs four (one for each row of the transition equation). In practice however, each successive index contains the lagged values of the previous one and (after accounting for the lags) the differences between the four indices can safely be ignored.

4. What the Coincident Index Tells Us

To evaluate the quality of our coincident indices, it is helpful to compare our national and New York State index with the national and New York State indices published by the Federal Reserve Bank of Philadelphia and with Megna and Xu's (2003) New York State index. Figures 4 and 5 provide a visual comparison of our indices with those published by the Philadelphia Fed and the text of this section provides a qualitative comparison. [4]

As mentioned in Section II.1, visually significant decreases in the coincident index indicate that the economy is in recession. Consequently, the peaks and troughs in the index can be used to mark the beginnings and endings of economic recessions. Therefore, the periods during which our coincident index indicates that the economy was in recession should roughly correspond to the recessionary periods in their indices.

Our indices exhibited longer periods of recession than their indices did. Generally speaking, our indices declined at about the same time or earlier than theirs and they recovered at about the same time or later than theirs.

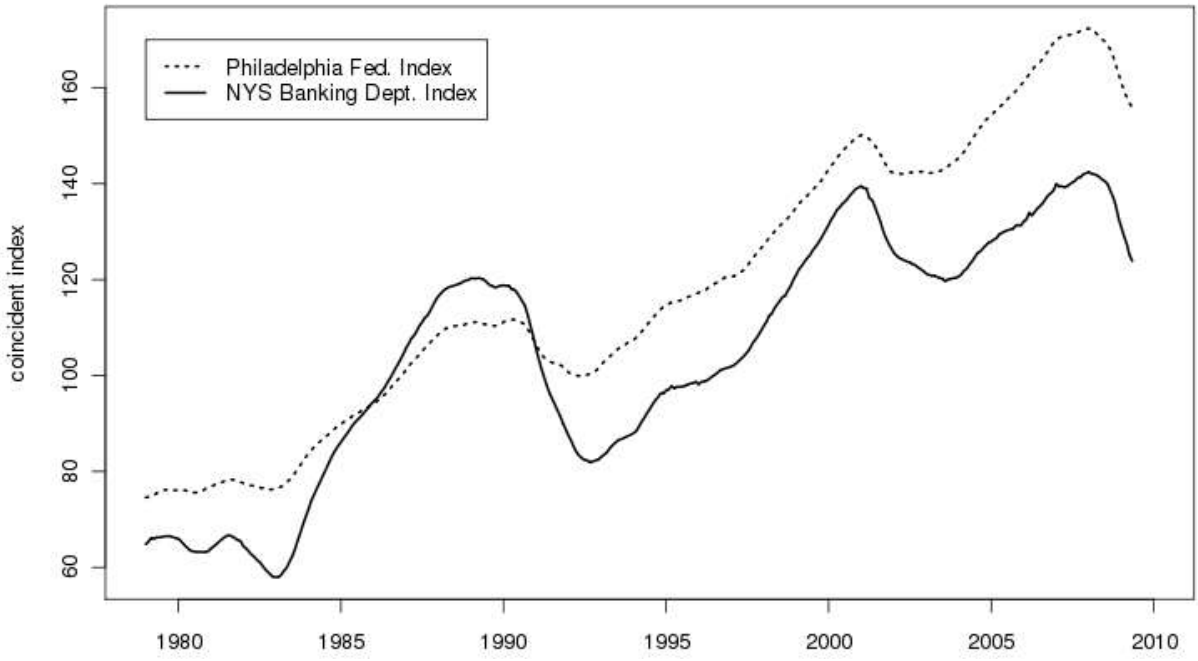
For example, our New York State index indicates that the New York State economy entered a recession in October 1979 and did not recover until November 1980. By contrast, the Philadelphia Fed index indicates that the New York State recession began in March 1980 and ended in July 1980, while Megna and Xu's index indicates that recession began in March 1980 and ended in September 1980.

All three indices indicate that the next New York State recession began in September 1981, but our index and Megna and Xu's index indicate that it ended in February 1983, while the Philadelphia Fed index indicates that it ended three months earlier, in November 1982.

At the end of the 1980s, both our index and the Philadelphia Fed index suggest that the New York State's economy plateaued for over a year before entering recession in the second quarter of 1990, but our index indicates that the recession lasted longer than the Philadelphia Fed index does. Specifically, our index indicates that it ended in September 1992, whereas the Philadelphia Fed index indicates that it ended in June 1992. By contrast, Megna and Xu's index shows a peak in June 1989 and a long recession that didn't end until November 1992.

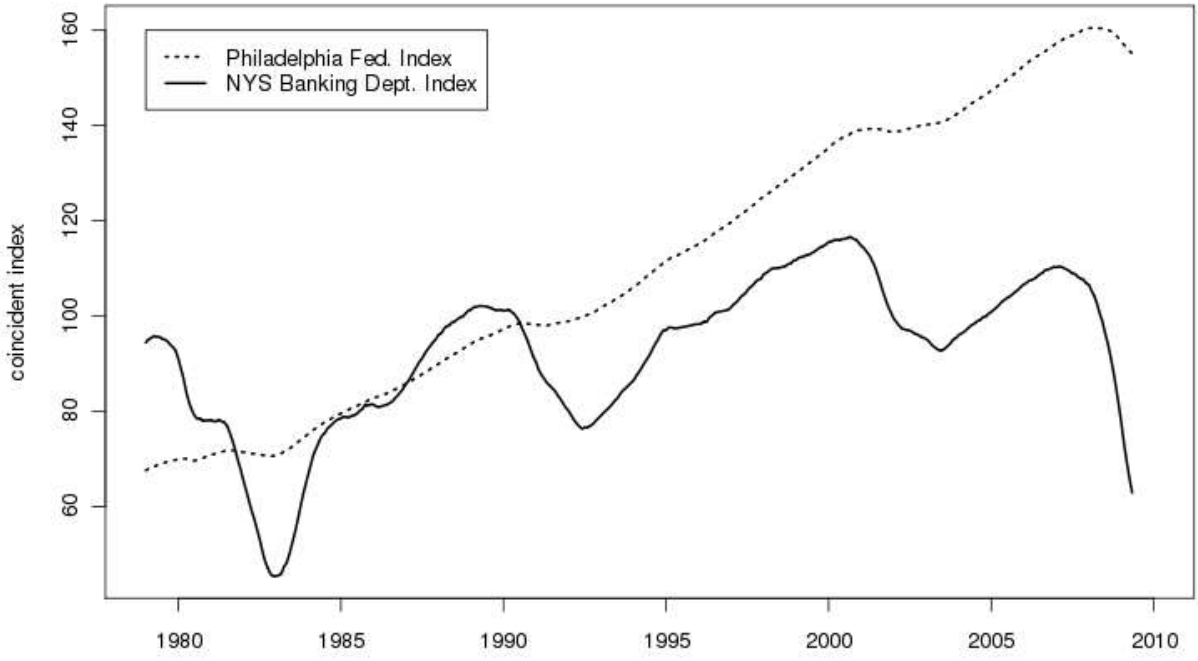
Our index and the Philadelphia Fed index also show a common beginning to the New York State recession that began in February 2001, but once again our index indicates that it lasted longer than the Philadelphia Fed index does. Specifically, their index indicates that it ended in March 2003, while our index indicates that it ended in August 2003.

Figure 4: Coincident Indices of the New York State Economy



Note: The series depicted above ends in May 2009.

Figure 5: Coincident Indices of the United States Economy



Note: The series depicted above ends in May 2009.

The last point of commonality between our index and the Philadelphia Fed index is the date when New York State entered the current economic recession: February 2008.

Sharp differences between our indices arise regarding when the U.S. economy enters recession and how long it lasts. Generally speaking, our index suggests earlier beginnings and longer durations of economic recessions than the Philadelphia Fed's index. For example, our index indicates that the United States entered the current economic recession in March 2007, whereas the Philadelphia Fed index indicates that the national economic recession began in April 2008.

Similar discrepancies are also seen in the dating of other recessions as well. At the extreme, our index indicates that the previous national recession began in October 2000 and lasted 33 months, while the Philadelphia Fed index places its beginning at June 2001 and indicates that it lasted only eight months.

A similarly large difference exists in the dating of the national recession of the early 1990s. Our index indicates that it lasted from April 1990 to June 1992, whereas the Philadelphia Fed indicates that it lasted from October 1990 to May 1991.

The differences in the dating are smaller however when we compare the dating of the two recessions in the early 1980s. Our index indicates that the first one began in July 1979 and lasted 14 months, while the Philadelphia Fed index indicates that the first one began in April 1980 and lasted four months. There is even more agreement on the dating of the second recession in the early 1980s. Our index indicates that it lasted from May 1981 until January 1983, whereas the Philadelphia Fed index indicates that it lasted from September 1981 until November 1982.

On balance, our indices generally agree more than they disagree. The dating of New York State's business cycles and the dating of the two national recessions of the early 1980s are similar enough that we can conclude that our indices adequately capture movement in overall economic activity in both the United States and New York State.

III. FORECASTING

Because the coincident index measures the state of the underlying economy, forecasts of the coincident index (called a "leading index") help us to foresee economic recessions and expansions. Because it is constructed from the unemployment rate, non-farm employment and average hours worked in manufacturing, the coincident index is also an excellent predictor of those variables, so we include forecasts of these "key economic variables" in our monthly reports. These forecasts have the added benefit of providing the reader with a concrete understanding of what the coincident index forecast means.

1. "Terraced" VARs

Our forecasting model consists of a set of "terraces." The top terrace forecasts exogenous variables that are useful in predicting changes in the national coincident index. The forecasts of those exogenous variables are passed down to the next terrace, which forecasts the change in the national coincident index. The national index forecast is then passed down to the next terrace, which forecasts the change in

a sub-national index. Finally, at the lowest level, the sub-national index forecast is used to forecast changes in the "key economic variables."

Forecasts on a given terrace are based on autoregressive models, which assume that the change in a given variable depends on past changes in that variable and on past changes in variables at higher level terraces. To be more concrete, this section will explain our forecasting model in terms of a simple three variable "terraced" vector autoregressive (VAR) model. In practice, there are many more variables in our model, but – to keep the example simple here – these will be discussed in further detail in Section III.4.

The specific example of a terraced VAR that we will use is as follows:

$$\begin{bmatrix} \Delta US_t \\ \Delta NYC_t \\ \Delta unemp_t \end{bmatrix} = \begin{bmatrix} \alpha_{US} & 0 & 0 \\ \beta_{US} & \beta_{NYC} & 0 \\ 0 & \gamma_{NYC} & \gamma_{unemp} \end{bmatrix} \begin{bmatrix} \Delta US_{t-1} \\ \Delta NYC_{t-1} \\ \Delta unemp_{t-1} \end{bmatrix} + \begin{bmatrix} u_{US,t} \\ u_{NYC,t} \\ u_{unemp,t} \end{bmatrix} . \quad (5)$$

The top row of equation 5 (or "top terrace") models the change in the U.S. coincident index at time *t* (denoted: ΔUS_t). In this case, it is assumed that ΔUS_t depends only on the previous value of the change in the US index (i.e. ΔUS_{t-1}) and the residual: $u_{US,t}$. In such a model, forecasts of the change in the U.S. coincident index will therefore depend exclusively on past values of the change in the US index.

The second row models the change in the New York City coincident index at time *t* (denoted: ΔNYC_t). According to equation 5, ΔNYC_t depends on ΔUS_{t-1}, its own past value (i.e. ΔNYC_{t-1}) and the residual: $u_{NYC,t}$. Forecasts of the change in the New York City coincident index will therefore depend on past values of the change in the New York City index as well as past values of the change in the U.S. index.

Finally, the third row models the change in the New York City unemployment rate at time *t* (denoted: $\Delta unemp_t$). In this case, it is assumed that $\Delta unemp_t$ depends on ΔNYC_{t-1}, its own past value (i.e. $\Delta unemp_{t-1}$) and the residual: $u_{unemp,t}$. Forecasts of the change in the New York City unemployment rate will depend on past values of the change in the New York City unemployment rate and past values of the change in the New York City coincident index. It is important to note that forecasts of the change in the New York City unemployment rate will also depend on past values of the change in the U.S. index (because the change in the New York City index depends on past values of the change in the U.S. index).

To set up the terraced VAR, we first created an ordinary VAR using Pfaff's (2008) "vars" package for R (R Development Core Team, 2008). We then used ordinary least squares (OLS) regression to estimate the coefficients of the terraced VAR and replaced the ordinary VAR's coefficients with the coefficients of the terraced VAR.

We also replaced the residuals in the VAR object with the rescaled residuals [5] from the OLS equations, so that the VAR object's residual covariance matrix could be used to compute the confidence intervals around our forecasts.

2. Long-Run Assumptions

As mentioned previously, a terraced VAR models *changes* in the variables of interest. We had to express each model in terms of changes to remove the spurious correlation that arises between two unrelated variables that trend upward (or downward) over time.

In practice, this means that we had to make sure that our terraced VARs do not contain a "unit root," so that the effect of any given shock will fade over time. (In other words, the model assumes that the events of 20 or 30 years ago do not have any influence on the month-to-month fluctuations in the variables we are examining).

Because all shocks fade over time and because our terraced VARs do not contain an intercept term, our models implicitly assume that the changes in the coincident index and changes in the key economic variables converge to zero in the long run. Since the changes are zero in the long run, the model assumes that the coincident index and key economic variables will remain constant in the long run.

In theory, this implies a prediction that the economy will never recover from the current economic crisis. In practice however, forecasts of exogenous leading economic indicators (discussed in Section X) act as an implicit intercept, which allow our model to predict both recovery from recession and decline from expansion.

3. Confidence Intervals

When forecasting, we must assume that the errors in equation 5 will equal zero. In other words, when forecasting the New York City coincident index we implicitly assume that we know the future course of the the US index with certainty. Similarly, when we forecast the New York City unemployment rate we implicitly assume that we know the future course of the the New York City index with certainty.

Of course, we do not know those values with certainty. All of our forecasts are prone to error, which we must account for. To that end, we estimate a range of forecast values, within which we can be 90 percent certain that the true values will fall. As mentioned in Section 3.1, these estimates are based on the residuals that we obtained when we estimated the coefficients of the terraced VAR.

To compute the confidence intervals, Pfaff's "vars" package uses forecast error variance decomposition, which is based on the Wold (1954) moving average decomposition for stable VAR-processes, which is defined as:

$$\mathbf{y}_t = \mathbf{\Phi}_0\mathbf{u}_t + \mathbf{\Phi}_1\mathbf{u}_{t-1} + \mathbf{\Phi}_2\mathbf{u}_{t-2} + \cdots \quad , \qquad (6)$$

where $\boldsymbol{y}_t$ is the vector of dependent variables at time t, $\mathbf{u}_{t-l}$ is the residual vector at time $t - i$ and the $\mathbf{\Phi}_i$ are matrices of parameters that depend on the coefficients of the VAR (or terraced VAR in our case). If the residuals are interpreted as unforeseen economic shocks, then equation 6 shows that current changes in our model's variables depend entirely on past shocks.

To compute the values of $\mathbf{\Phi}_i$, Pfaff's "vars" package sets $\mathbf{\Phi}_0$ equal to an identity matrix (with the same number of dimensions as the VAR has variables) and recursively computes $\mathbf{\Phi}_s$ as:

$$\mathbf{\Phi}_s = \sum_{j=1}^{s} \mathbf{\Phi}_{s-j} A_j \qquad \text{for s = 1,2,…,} \tag{7}$$

where A_j is matrix j of the p matrices of VAR coefficients [6] and where $A_j = 0$ for $j > p$. The values of $\mathbf{\Phi}_i$ are then used to compute the forecast error covariance matrix:

$$Cov\left(\begin{bmatrix} \mathbf{y}_{T+1} - \mathbf{y}_{T+1|T} \\ \vdots \\ \mathbf{y}_{T+h} - \mathbf{y}_{T+h|T} \end{bmatrix}\right) = \begin{bmatrix} I & 0 & \cdots & 0 \\ \mathbf{\Phi}_1 & I & & 0 \\ \vdots & & \ddots & 0 \\ \mathbf{\Phi}_{h-1} & \mathbf{\Phi}_{h-2} & \cdots & I \end{bmatrix} (\mathbf{\Sigma_u} \otimes I_h) \begin{bmatrix} I & \mathbf{\Phi}_1^T & \cdots & \mathbf{\Phi}_{h-1}^T \\ 0 & I & & \mathbf{\Phi}_{h-2}^T \\ \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & I \end{bmatrix}, \tag{8}$$

where T is last period for which data is available, h is the number of forecast periods, I_h is an identity matrix with h dimensions, $\mathbf{\Sigma_u}$ is the residual covariance matrix, the operator $\otimes$ is the Kronecker product and $\boldsymbol{y}_{T+h} - \boldsymbol{y}_{T+h\,|T}$ is the difference between the true value of the dependent variable vector at time $T+h$ and our forecast of it, given the information available to us at time T.

The square roots of the diagonal elements of the forecast error covariance matrix are the standard errors of the forecasted variables, which we then use to compute the confidence intervals.

Had we estimated the terraced VAR that was used as an example in Section VII and then computed the confidence interval around its forecasts, we would find that the confidence interval around the U.S. coincident index is computed as if the New York City variables were not in the terraced VAR at all.

The confidence interval around the New York City index would depend on the variances of the U.S. and New York City index residuals and the covariance between them (i.e. just as if it had come from a standard two-variable VAR), but would be blind to the presence of the New York City unemployment rate in the terraced VAR.

In the third forecast period (and thereafter), the confidence interval around New York City's unemployment rate would depend on all six variances and covariances (just as one would expect from the model).

There are two sources of error, however, that the forecast error variance decomposition cannot account for. One is revisions of the data used to generate the index and forecasts. The other is sampling error. Forecast error variance decomposition does not account for sampling error because it implicitly assumes that the VAR's coefficients are exactly equal to their true values. Had we estimated the coefficient from a different time interval, we would have obtained different coefficient estimates and thus gotten a different forecast.

As Feldstein (1971) has shown, sampling error implies that the forecast error of the New York City variables (in Section VII) would not be distributed normally. To overcome the problems associated with non-normality, McCullough (1996) proposed bootstrapping.

We have chosen not to use bootstrapped confidence intervals however for two reasons. First, given the complexity of our models, the bootstrapping code is very difficult to write and could introduce an unknown degree of human error into the bootstrapped confidence intervals. Second, the purpose of this

exercise is to provide the reader of our reports with a general sense of the direction in which the national, state and local economies are headed. We do not need precise estimates of the distribution of the forecast error.

4. Forecast Accuracy

It is simple to create a terraced VAR and compute the confidence intervals around its forecasts, but there are many possible terraced VAR structures to choose from. Therefore, we need to select the structure that has the most predicative power. We also wish to show that using coincident indices in our forecasting models increases the model's predicative power.

To that end, we estimated the coefficients of many possible terraced VAR structures using the available data between from 1990 to 2004. We then used the estimated coefficients and the rest of the available data (which ended in December 2008) to create 37 sets of 12-month forecasts and 11 sets of shorter forecasts. Finally, we compared the forecasts to the actual values and computed the Mean Squared Error (MSE) of the model's forecasts:

$$MSE_k = \frac{1}{n}\sum_{i=0}^{n-1}\left(y_{T+i+k} - y_{T+i+k|T+i}\right)^2 \quad ,$$

where k is the number of periods ahead for which the variable y_t is forecasted, $y_{T+i+k|T+i}$ is the forecasted value of y_{T+i+k} given the information available at time $T+i$ and n is the number of k-period ahead forecasts.

As Greene (2000) notes, the MSE does not have a scale which allows us to compare the MSE of one forecasted variable to the MSE of another. To get around this problem, we compared the MSE to the variance of the forecasted variable by creating a "relative MSE" statistic, which is similar in spirit to the coefficient of determination:

$$Rel.MSE_k = 1 - \frac{MSE_k}{\mathrm{var}(y_{t+k})}$$

where $var(y_{t+k})$ is the variance of the variable y_{t+k} in the forecast period.

A relative MSE of one implies perfect prediction, while lower values imply weaker forecasting performance. It is important to note that this statistic may be negative, but – because we cannot know what the variable's mean will be in the forecast period – negativity does not necessarily mean that the forecast is poor. Sometimes, a slightly negative relative MSE is the highest value one can obtain when forecasting a very volatile variable.

Table 1 uses relative MSE to compare our forecast models to both a random walk and an "alternative model," in which the region's coincident index is not used to forecast "key economic variables" (i.e. the unemployment rate, non-farm employment and average hours worked in manufacturing). To keep the table concise, the reported relative MSE statistic is the average of the 12 $Rel.MSE_k$:

$$Rel.MSE = \frac{1}{12}\sum_{k=1}^{12} Rel.MSE_k$$

Table 1 shows that our forecast models always outperform the "random walk" forecasts (i.e. $y_{T+k|T} = y_T$ for all *k*). In other words, our forecast models always provide better prediction than simply observing present economic conditions and assuming that the same condition will continue to hold for the foreseeable future.

With three exceptions, Table 1 shows that our forecast models always outperform the forecasts of the "alternative model," which is the best model that we could develop that does not include the region's coincident index as an explanatory variable. This shows that the region's coincident index is a good predictor of the variables that were used to create it.

The three exceptions were the unemployment rates of New York State and the Nassau-Suffolk MSA and non-farm employment of the Putnam-Rockland-Westchester MSA, but in each of these cases, the difference between the forecast model and the alternative model was small. Given the proximity of Nassau-Suffolk and Putnam-Rockland-Westchester to New York City, it may be the case that the labor markets in these two MSAs are more dependent on state of the New York City economy than they are on their own. Moreover, the fact that New York City accounts for approximately 43 percent of the New York State population, the New York State unemployment rate may also be heavily dependent on state of the New York City economy.

Interestingly, the coincident indices turned out to be good predictors of themselves. This may be attributable to the auto-regressive processes that the coincident indices follow (i.e. to the smoothness" of the coincident index series).

The U.S. coincident index also appears to be a good predictor of state and regional coincident indices. This may be attributable to the fact that regional economies are strongly influenced by national economic conditions.

While developing our forecasting models, we also experimented with other variables, such as the Federal Reserve's measures of industrial production and capacity, the Census Bureau's series on housing starts, durable goods and retail sales. The model that provided the most predicative power however only included the Census Bureau's durable goods unfilled orders series as an exogenous variable.

Taken as a whole, one can conclude from the reasonably high values of our forecast models' relative MSE statistics that our forecasting models capture a good share of the out-of-sample variation in the variables that the models predict.

Table 1: Forecast Accuracy

	Rel. MSE	**Forecast Model**	Rel. MSE	**Alternative Model**	**R. Walk** Rel. MSE
United States					
Dur. Goods Unfill. Orders	-0.220	DGU(1,2,3,4,5,6)			-0.770
Coincident Index	0.609	DGU(4), US(1,2)			0.494
Unemployment Rate	0.348	US(1,2), UR(1,2)	-0.020	DGU(4), UR(1,2)	-0.320
Nonfarm Employment	0.615	US(1,2), EMP(1,2)	0.390	DGU(4), EMP(1,2)	0.273
Avg. Hours Worked Manuf.	0.080	US(1,2), HR(1,2)	0.001	DGU(4), HR(1,2)	-0.752
New York State					
Coincident Index	0.475	US(1,2), NYS(1,2), NYC(1,2)			0.069
Unemployment Rate	0.368	US(1,2), NYS(1,2), UR(1,2)	0.371	US(1,2), NYC(1,2), UR(1,2)	-0.449
Nonfarm Employment	0.185	NYS(1,2), EMP(1,2)	-0.045	US(1,2), NYC(1,2), EMP(1,2)	-0.640
Avg. Hours Worked Manuf.	0.047	US(1,2), NYS(1,2), NYC(1,2), HR(1,2)	0.028	US(1,2), NYC(1,2), HR(1,2)	-0.895
New York City					
Coincident Index	0.783	US(1,2), NYC(1,2)			0.408
Unemployment Rate	0.259	US(1,2), NYC(1,2), UR(1,2)	0.241	US(1,2), UR(1,2)	-0.703
Nonfarm Employment	0.147	NYC(1,2), EMP(1,2)	-0.234	US(1,2), EMP(1,2)	-0.468
Nassau-Suffolk					
Coincident Index	0.773	US(1,2), NYC(1,2), NS(1,2)			0.419
Unemployment Rate	0.263	US(1,2), NYC(1,2), NS(1,2), UR(1,2)	0.269	US(1,2), NYC(1,2), UR(1,2)	-0.509
Nonfarm Employment	0.166	NYC(1,2), NS(1,2), EMP(1,2)	0.134	US(1,2), NYC(1,2), EMP(1,2)	-0.740
Putnam-R'land-W'chester					
Coincident Index	0.661	US(1,2), NYC(1,2), PRW(1,2)			0.479
Unemployment Rate	0.229	PRW(1,2), UR(1,2)	0.216	US(1,2), NYC(1,2), UR(1,2)	-0.537
Nonfarm Employment	0.058	US(1,2), NYC(1,2), PRW(1,2), EMP(1,2)	0.095	US(1,2), NYC(1,2), EMP(1,2)	-1.030

Table continued on the next page

Table 1 (continued): Forecast Accuracy

	Rel. MSE	**Forecast Model**	Rel. MSE	**Alternative Model**	**R. Walk** Rel. MSE
Buffalo-Niagara Falls					
Coincident Index	0.641	US(1,2), BN(1,2)			0.456
Unemployment Rate	0.254	US(1,2), BN(1,2), UR(1,2)	0.207	US(1,2), UR(1,2)	-0.721
Nonfarm Employment	0.031	BN(1,2), EMP(1,2)	-0.012	US(1,2), NYS(1,2), EMP(1,2)	-1.055
Rochester					
Coincident Index	0.660	US(1,2), Rch(1,2)			0.505
Unemployment Rate	0.292	US(1,2), NYS(1,2), Rch(1,2), UR(1,2)	0.240	US(1,2), NYS(1,2), UR(1,2)	-0.603
Nonfarm Employment	-0.039	NYS(1,2), Rch(1,2), EMP(1,2)	-0.210	US(1,2), EMP(1,2)	-1.168
Albany-Sch.-Troy					
Coincident Index	0.721	US(1,2), NYS(1,2), AST(1,2)			0.496
Unemployment Rate	0.193	US(1,2), NYS(1,2), AST(1,2), UR(1,2)	0.169	US(1,2), NYS(1,2), UR(1,2)	-0.774
Nonfarm Employment	0.124	AST(1,2), EMP(1,2)	0.122	US(1,2), NYS(1,2), EMP(1,2)	-1.030

The numbers in parenthesis are the lag numbers of the variables that appear in the model's equation.

DGU – Durable Goods Unfilled Orders
US – United States Coincident Index
NYS – New York State Coincident Index
NYC – New York City Coincident Index
NS – Nassau-Suffolk Coincident Index
PRW – Putnam-Rockland-Westchester Coincident Index
BN – Buffalo-Niagara Falls Coincident Index
Rch – Rochester Coincident Index
AST – Albany-Schenectady-Troy Coincident Index
UR – Unemployment Rate
HR – Average Hours Worked in Manufacturing
EMP – Nonfarm Employment

IV. CONCLUSION AND DIRECTION FOR FUTURE RESEARCH

This paper has explained what a coincident index is and has detailed the methodology employed to estimate the values of the coincident index for the United States, New York State and the six largest MSAs in New York State. It has also explained how we use the coincident index to forecast the future course of “key economic variables” in those regions.

No forecast will ever be perfect; however, we account for the inevitable error in our forecasts by using Terraced VARs to simultaneously forecast all of the variables in the model and compute confidence intervals that depend solely on the variables that affect the value of the forecasted variables. The

confidence intervals are imperfect because they do not account for sampling error, but these imperfections occur in all VAR model forecast confidence intervals (not just terraced VARs).

As Section III.4 has shown, our forecasts are more accurate than those we could obtain from assuming a random walk and those that we would obtain had we not included the coincident index in our forecasting models.

In future work, we will explore the correlation between our index and other variables of interest. For example, to the extent that borrowers' ability to meet their monthly mortgage and consumer credit payments depends on the state of their local economy, then we can examine whether there is a correlation between loans past due and the coincident index. If there is a correlation, we can use the forecast of the coincident index to forecast loans past due and predict future delinquency trends.

ENDNOTES

1. The coincident index estimates exclude Utica-Rome, Syracuse, Elmira, and Binghamton due to insufficient data.
2. Stock and Watson used data on industrial production, personal income net of transfer payments, hours worked by employees of nonagricultural establishments and total manufacturing and trade sales to develop their index, while Megna and Xu used data on total employment, sales tax revenue (as a proxy for retail sales), hours worked in manufacturing and the unemployment rate.
3. Orr et al. developed their index from data on total employment, real wages and salaries, the unemployment rate and average hours worked in manufacturing.
4. It should be noted, however, that Megna and Xu's framework used monthly sales tax collection data as a proxy for retail sales instead of a real wage series.
5. We requested, but were unable to obtain, Megna and Xu's New York State index. Our comparisons with their results are based on their published paper.
6. Rescaling the residuals was necessary to account for the difference in degrees of freedom between the ordinary VAR and the OLS regression.
7. The number of VAR coefficient matrices, p, is equal to the lag order of the VAR.

ACKNOWLEDGMENTS

We would like to thank the New York State Banking Department for supporting our research and for providing many useful comments and suggestions. The views expressed in this article are our own opinions and do not necessarily reflect the opinions of the New York State Banking Department

REFERENCES

Bram, Jason, James Orr, Rae Rosen, Robert Rich and J. Song. 2009. "Is the Worst Over? Economic Indexes and the Course of the Recession in New York and New Jersey." *Current Issues in Economics and Finance*, 15(5).

Crone, Theodore M. and Alan Clayton-Matthews. Nov. 2005. "Consistent Economic Indexes for the 50 States." *Review of Economics and Statistics*, 87(4):593–603.

Federal Reserve Bank of Philadelphia. State Coincident Indexes (2008, 2009). http://www.philadelphiafed.org/research-and-data/regional-economy/indexes/coincident/.

Feldstein, Martin. 1971. "The Error of Forecast in Econometric Models when the Forecast-Period Exogenous Variables are Stochastic". *Econometrica,* 39(1):55–60.

Greene, William H. 2000. *Econometric Analysis,* Upper Saddle River, NJ, Prentice Hall, 4 edition. 20

Kalman, Rudolf E. 1960. "A New Approach to Linear Filtering and Prediction Problems." *Transactions of the ASME – Journal of Basic Engineering*, 82(1):35–45.

McCullough, B.D. 1996. "Consistent Forecast Intervals when the Forecast-period Exogenous Variables are stochastic." *Journal of Forecasting*, 15(4):293–304.

Megna, Robert and Qiang. Xu. 2003. "Forecasting the New York State Economy: The Coincident and Leading Indicators Approach." *International Journal of Forecasting*, 19(4):701–713.

Orr, Jason, Robert Rich and Rae Rosen. 1999. "Two New Indexes Offer a Broad View of Economic Activity in the New York-New Jersey Region*." Current Issues in Economics and Finance*, 5(14).

Petris, Giovanni. 2009. *dlm: Bayesian and Likelihood Analysis of Dynamic Linear Models*, R package version 0.99-0.

Pfaff, Bernhard. 2008. "VAR, SVAR and SVEC Models: Implementation Within R Package 'vars.'" *Journal of Statistical Software*, 27(4):1–32. http://www.jstatsoft.org/v27/i04/.

R Development Core Team. 2008. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2008. http://www.R-project.org.

Stock, James H. and Mark W. Watson. 1989. "New Indexes of Coincident and Leading Economic Indicators." *NBER Macroeconomics Annual*, 4:351–394.

US Census Bureau, Time Series Staff. *X-12-ARIMA*. Washington, DC, 15 June 2007. Version 0.3, 20.

Wold, Hermann. 1954. *A Study in the Analysis of Stationary Time Series*. Almqvist and Wiksell Book Co., Uppsala.

Consumer Demand for Durable Goods, Nondurable Goods and Services

John J. Heim*

ABSTRACT

Separate macroeconomic consumption demand functions are developed and tested for (1) durable goods, (2) nondurable goods and (3) services. These are compared for consistency with econometric studies of total consumer demand. U.S. data for 1960 - 2000 is used. The econometric method used was 2SLS with heteroskedasticity controls. The models explain 94% of the variance in demand for consumer durables, 86% of demand for nondurable consumer goods and 81% of services demand.

Keywords: *Consumption, Consumer Durables, Consumer Nondurables and Services, Demand*

1. INTRODUCTION

Every economics student learns that the level of what's produced, the GDP, is determined by the demand for business and consumer goods, demand by government for goods and services and (the net) demand for exports and imports in accordance with the formula

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

Recent studies have subjected many theorized determinants of consumer, investment and imports demand to testing to see if they are empirically verifiable (Heim 2008a&b). Knowing what drives demand for goods and services is critical for businesses trying to project future demand, and for economists trying to advise them or determine how demand can be stimulated by changing government policy (e.g., toward credit availability or tax levels). One of these studies showed that you can't treat different components of investment demand the same: its three components (plant and equipment demand, inventory demand and residential housing demand) are driven by different factors, or by the same factors, but with different time lags. It found that studies of total investment only often provide misleading inferences about how these determinants affect demand for any one type of investment.

This study examines whether or not the same might be true for consumer spending. Does knowledge of what factors drive overall demand for consumption give us all the information we need to know about what drives demand for the separate parts? Do all variables which influence overall consumer demand influence in the same way demand for consumption's three component parts: durable goods(like cars and furniture), nondurable goods (like food and clothing), and services (like entertainment, restaurant, lawyer and accountant services).

* Professor of Economics, Rensselaer Polytechnic Institute, Troy, NY

Heim (2008c) indicated that demand for U.S. consumer goods is driven by the following factors (listed in order of importance), which explained 92% of the variance in total consumer demand, 1960 - 2000:

- Disposable income
- credit constraints due to the "crowd out" effects of government deficits,
- consumer wealth
- interest rates
- the exchange rate

The study did not develop separate demand functions for the three subcomponents of total consumption. This paper addresses that need by econometrically developing and testing demand functions for the three separate types of consumption, also using U. S. data for the period 1960 – 2000.

2. THEORIES OF DEMAND FOR CONSUMPTION GOODS:

The middle 50 years of the 20th century was dominated by Keynesian theory of how the economy operates, which is driven by Keynes' theory of what determines consumption spending. Keynes argues in chapter 8 of the General Theory of Employment, Interest and Money (1936) that income, wealth, taxes (fiscal policy) and possibly the rate of interest might influence consumption. However, he felt

> … income…is, as a rule, the principal variable upon which the consumption-constituent of the aggregate demand function will depend…(though)…windfall changes in capital-values will be capable of changing the propensity to consume, and substantial changes in the rate of interest and in fiscal policy may make some difference… (pp.95-96)

where "fiscal policy" is a reference to tax levels and capital values a reference to wealth. In chapter 9 he also notes other factors that might affect the level of consumption spending: precautionary, saving for known future needs (like retirement), and saving to finance improvements in future standards of living.

In the late 1950's and early 1960's, Keynes theory was challenged by Franco Modigiani's Life Cycle hypothesis, (Modigiani 1963) and Milton Friedman's Permanent Income Hypothesis (Friedman 1957). Central to both theories was that average income perhaps including future expected as well as past income, not just current income alone, as Keynes had argued, determined current period levels of consumption spending. Heim 2008b exhaustively tested different variants of these rational and adaptive expectations based income average models and found that none of them could explain variation in consumption as well as Keynes' current income variable, even controlling for other significant variables influencing consumption, such as wealth, credit availability and interest rates. Since the difference in the Keynes and Modigliani/Friedman functions has huge policy implications for how well the economy reacts to government spending and tax stimulus, this was an important finding.

Hence, the key theoretically postulated determinants of consumption we will test below are those from Keynesian theory: income, taxes, interest rates, wealth, and credit availability (measured by 'crowd out" caused by the government deficit. A number of other variables will also be tested.

3. METHODOLOGY

The following variables constitute this study's initial hypothesis of the determinants of demand for durable, and non durable goods spending. The hypothesis for testing is:

(1) $C_{DorND} = \alpha + \beta_1 (Y-T_G) + \beta_2 (T_G-G) + \beta_3 (DJ_{-2}) + + \beta_4 (PR) + \beta_5 (XR_{AV0123})$

where

- C_{DorND} = Consumer spending on durable or non durable goods
- $(Y- T_G)$ = Disposable income
- $(T_G -G)$ = The government budget deficit/surplus or (T_G and G may be modeled separately)
- DJ = = The Dow Jones Composite Stock Index, a measure of changes in consumer wealth
- PR = The Prime Interest Rate(r), multiplied by the size of the GDP (Y) two years earlier
- XR_{AV0123} = The average Exchange Rate for the current and past three years. Preliminary studies showed the full effects of exchange rate changes are not felt until three years after the change.

To these five are added population growth (POP) and demand for new housing (HOUSE), and variables that test for the effects of differing age compositions of the population over time (162465b34) and the relative expensiveness of housing compared to income (HP/INC).

The hypothesized determinants of services demand tested included these variables and some others found to influence the demand for the residential construction component of investment demand. (The rental value of owner-occupied housing is added to the GDP each year as an estimate of the (consumer) services (protection from the elements) provided by housing. Since newly constructed housing in the GDP expands the number of houses providing consumers with housing services, variables found related to housing demand in the investment study were also tested as possible determinants of this type of consumer serves demand. Other variables were added, including the cost of housing relative to income (HP/INC), population size, the mortgage interest rate and the percentage of population in prime house-buying age groups in the population (162465b34)). Hence, our model of services demand becomes:

(2) $C_S = \alpha + \beta_1 (Y-T_G) + \beta_2 (T_G-G) + \beta_3 (DJ_{-2}) + \beta_4 (PR) + \beta_5 (XR_{AV0123})) + \beta_6(MORT) + \beta_7(HP/INC) + \beta_9(POP)$

where the first five determinants are defined as above. The remainder are defined as follows:

- C_S = Expenditure on consumer services

- ACC = The Accelerator (ΔY). It is a measure of the rate of growth of the economy, therefore a measure of business climate: i.e., the current boom/bust condition of the economy
- MORT = The mortgage interest rate, the interest rate found most systematically related to housing construction demand in the earlier investment study
- HP/INC = The cost of housing relative to income
- POP = The percentage of population in prime house-buying age groups in the population

Through econometric testing of the three different types of consumption we can determine which of these hypothesized determinates really do move in ways consistent with the hypothesis they are determinants of consumer demand. Testing will also suggest something about the marginal effects on consumption resulting from changes in these determinants, and how reliable our estimates of these marginal effects are likely to be.

3.1 DATA USED

1960 – 2000 data taken from the *Economic Report of the President, 2002,* appendix tables B2, B7, B26, B28, B54, B73, B82, B95 was used to test the models. Data are in real, rather than nominal values, deflated where necessary using the most appropriate chained price index (base year =1996) from Table B7. The nominal prime interest rate is deflated using the average of the past two years consumer price index from Table B60.

3.2 THE ECONOMETRIC TESTING PROCEDURE

A “stepwise” regression procedure was employed to determine which of the variables explained the most variance in consumption during the 40 year period studied. Consumption theory rarely tells us how long it takes for a change in one of consumption’s determinants to bring about a change in consumption, yet the lag hypothesized can critically influence test results (Heim 2008b).

To determine the appropriate lags to use with a variable, each of the individual variables cited above was tested in a preliminary model containing only one other explanatory variable: current year disposable income (testing had indicated the current year or ‘zero” lag value of this variable explained the most variance). Each added variable was tested using seven different lags (+3 in the future to -3 in the past). The lag level that added the most to explained variance was picked as the lag level to use subsequently in larger tests of all the hypothesized determinants of consumption, unless the sign on the variable was theoretically wrong, or if the result suggested the direction of causation was backward. Subsequently, these variables, with their chosen lags, were used in a stepwise regression procedure to determine how much of the variance in consumption each explained.

A problem with the stepwise technique first made obvious by Goldberger (1961) and others since then is that when using stepwise regression, order of entry itself can influence estimates of how much variance a variable explains. Generally speaking, the earlier in the process a variable is entered, the larger the percent of variance it will explain. This is because, typically, explanatory variables are intercorrelated. Hence entering one allows it to "pick up" the variance it uniquely explains, plus some of the variance best explained by another variable, but one not yet entered in the regression. Hence, our findings should be considered as providing some ordinal, not cardinal, information, about relative importance, and not necessarily conclusive.

An additional use of the stepwise procedure results is to provide a way of assessing the stability (robustness) of the marginal effects (regression coefficient) to the model as new variables are added or subtracted. Regression coefficients, point estimates of marginal effects, often vary significantly as variables are added or subtracted from the regression, if variables added are correlated with variables already in the regression. This study will show the more variance a regression currently explains, the less likely marginal effect estimates will change when additional variables are added. Hence, in incompletely specified regressions, especially when important explanatory variables are left out, the possibility of overstating the marginal effects of a variable, or its statistical significance, are substantial, and results are subject to major change when the omitted variable is added. (Goldberger, 1961)

Since the base problem affecting the stability (or robustness) of our estimates is intercorrelation among the explanatory variable set, we can enhance the likelihood of stable point estimates by reducing the intercorrelation before running the regressions. One way of doing this that can be successful is to use "first difference" rather than "levels" of the data when estimating regression coefficients. In general, the first differences of two time series variables are substantially less intercorrelated than are their underlying data in levels. This technique has the added advantage in time series data sets, such as the one used here, of reducing autocorrelation problems (Griffiths, Hill, Judge, 1993), and can reduce nonstationarity.

Consumption, the dependent variable, and the (Y) component of the explanatory variable ($Y\text{-}T_G$) are simultaneously determined, since C_t is part of Y_t. Two Stage Least Squares regression is the appropriate form of regression to use to avoid this simultaneous equations bias (Griffiths, Hill, Judge, 1993)).

Evidence of heteroskedasticity was found in preliminary testing. Newey West heteroskedasticity corrections were made in all tests. (Newey West, 1987)

4.0. FINDINGS: DETERMINANTS OF THE TOTAL DEMAND FOR CONSUMER GOODS

Stepwise testing of the theorized determinants of total consumption yielded the Table 1 marginal impacts for variables determining total consumer demand in the U.S. 1960-2000:

The results in Table 1 are presented for each step in the stepwise regression procedure. Showing each stepwise equation's results individually provides information about regression coefficient stability. The stability of regression coefficients increases significantly as the total variance (R^2) explained by the

Table 1: Stepwise Addition to the Regression of Hypothesized Determinants of Total Consumption

$\Delta C_t = f\ [\ \beta_1 \Delta(Y\text{-}T_G)_t, + \beta_{2T\&2G} \Delta(\text{Crowd Out: } T_G, G)_t, + \beta_3 \Delta DJ_{-2}, + \beta_4\ \Delta PR, + \beta_5 \Delta\ XR_{AV0123} + \beta_6\ \Delta PR\ ]$

	$\Delta(Y\text{-}T_G)$	ΔT_G	ΔG	ΔDJ_{t-2}	ΔPR	ΔXR_{AV0i23}	$\Delta P0P$
R^2/Adj. R^2 (DW).	$\beta_1(t)$.	β_{2T} (t) .	$\beta_{2G}(t)$.	$\beta_3(t)\|$.	$\beta_4(t)$.	$\beta_5(t)$.	$\beta_6(t)$.
68% (1.3)	.82(12.6)						
84% (1.7)	70 (21.7)	.43 (3.9)	.25 (1.2)				
89% (1.9)	.65 (21.0)	.32 (3.5)	.21 (1.0)	.81 (4.5)			
91% (1.8)	.65 (31.8)	.48 (4.8)	.07 (0.4)	.71 (5.1)	-6.72(-3.9)		
92% (2.0)	.66 (29.2)	.49 (5.7)	.04 (0.3)	.62 (4.9)	-6.93(-3.2)	2.83 (3.2)	
93/91%(1.9)	.58 (12.5)	.54 (6.2)	-.08(0.6)	.59 (4.9)	-7.12(-3.3)	3.48 (3.8)	.006(1.8)

(*) t- statistics of 2.0 = 5% significance; t - statistics of 2.7 = 1% level of significance.

(**) The Crowd Out Variable (T_G- G) Reported As two separate Deficit Variables, ΔT_G and ΔG because of noticeable different effects.

model increases. Until the model explains approximately 85% or more of the variation in consumption, regression coefficients change so much when variables are added or subtracted they are simply unreliable. Other variables tested, thought a priori to influence total consumer demand were nominal, current period mortgage interest rates and real housing prices as a percent of real per capita income. Neither was statistically significant, so they were dropped from the regression. The mortgage interest rate variable was also tested to see if it better explained the effect of interest rate changes on overall consumer demand than the prime interest rate. It did not. When substituted for the prime rate, it was statistically insignificant. Population growth (POP) is left in, though its statistical significance is only at the .09 level, below out typical .05 cutoff. It is left in because evidence introduced later will show that it negatively seems to impact durables demand, but positively affect non durables demand (ceteris paribus). This may be because population growth (family size growth) reduces discretionary income available for durables and necessitates increased nondiscretionary consumption (e.g, food, clothing).

Table 2 shows the results of stepwise addition of variables to the model. Overwhelmingly, as Keynes predicted, current year disposable income is the most influential factor. Next in importance, somewhat unexpectedly, was our measure of crowd out – limitations on consumer credit availability. One might expect the interest rate underlying most consumer credit (the Prime Rate used here) would reflect such limitations completely. This would probably be true if the Prime Rate were a market determined rate, but it not. It is an administered rate, set by banks to maintain a 3 point spread between itself and the (Federal Reserve administered) federal funds rate (Heim 2009, p. 58). Hence, the amount of loanable funds available absorbed by the government as it attempts to finance deficits by borrowing also becomes a key measure of consumer spending.

TABLE 2: Contributions to Explained Variance

Variable	(Adds To R^2)
Disposable Income	68 %
Crowd Out	16
Wealth	5
Interest Rate	2
Exchange Rate	1.2
Population Growth.	0.5 .
Explained Variance	92.7%

4.1 DETERMINANTS OF CONSUMER DEMAND FOR DURABLE GOODS

Table 3: Stepwise Addition - Regression of Hypothesized Determinants of Durable Goods Demand

$\Delta C_D = f$ [$\beta_1 \Delta(Y\text{-}T_G)_t$, + $\beta_{2T\&2G} \Delta$(Crowd Out: T_G,G)$_t$, + $\beta_3 \Delta XR_{AV0123}$ + $\beta_4 \Delta DJ_{-2}$, + β_5 ΔPR + β_6 ΔPOP (plus either ΔMort or ΔHouse)]

	Δ(Y-T_G)	Δ T_G	ΔG	ΔXR_{AV0i23}	Δ DJ_{t-2}	ΔMORT	ΔPR	ΔHOUSE	ΔPOP
R^2/Adj.(DW) .	β_{1t}(t) .	β_{2T} (t) .	β_{2G}(t) .	β_3(t) .	β_4(t) .	$\beta 6_7$(t) .	β_5(t) .	β_8(t) .	β_6(t) .
49/49% (0.6)	.16(9.3)								
67/66% (0.9)	.14(7.0)	.18 (4.6)	-.10(-1.0)						
81/79% (1.4)	.14 5.6)	.17 (5.6)	-.15(-1.8)	3.23 (5.0)					
86/84% (1.9)	.12(8.6)	.14 (4.8)	-.17(-2.3)	2.83 (4.9)	.28 (3.5)				
89/87% (2.1)	.12(9.2)	.14 (5.4)	-.16(-2.6)	2.47 (4.7)	.30 (4.2)	-5.36(-3.0)			
90/89% (2.2)	.12(9.8)	.19 (5.6)	-.21(-3.2)	2.59 (5.1)	.26 (3.8)	-4.02(-2.2)	-1.90(-2.0)		
91/90% (2.0)	.08(5.9)	.08 (3.0)	-.03 (-0.4)	1.89 (3.8)	.41 (5.9)			.31 (4.5)	
92/91% (2.2)	.09(6.3)	.12 (3.3)	-.08 (-1.1)	2.06 (4.1)	.37 (5.1)		-1.37(-1.6)	.26 (3.6)	
92/90% (2.2)	.09(5.7)	.12 (3.1)	-.08 (-1.0)	2.06 (4.1)	.37 (4.9)	-.06(-0.0)	-1.37(-1.6)	.26 (2.6)	
93/91% (2.2)	.13(5.2)	.07 (2.7)	.01 (0.1)	1.72 (3.6)	.40 (6.2)			.26 (3.7)	
94/92% (2.2)	.14(5.7)	.12 (3.4)	-.05 (-0.7)	1.89 (4.1)	.35 (5.3)		-1.59(-2.0)	.20 (2.7)	-.003(2.2)
									-.004(2.5)

(*) t- statistics of 2.0 = 5% significance; t - statistics of 2.7 = 1% level of significance.

(**)Crowd Out Variable (T_G- G) Reported As Two Separate Variables).

Table 3 shows stepwise addition of variables to the model. The results indicate the factors whose variation explained most variance in consumption during the 1960 – 2000 period:

- Disposable income (49%)
- Crowd out (as indicated by the government deficit) (18%)
- The Exchange Rate (14%)
- Wealth (as indicated by the Dow Jones Composite average) (5%)
- Demand for Residential Housing (5%)
- Population Growth (2%)

- (Real) Prime Interest Rate (1%)

Alternatively, if we exclude the residential housing demand variable (HOUSE), and instead let the mortgage interest rate to pick up the effect of housing market conditions on demand for consumer durables, the last two variables (after wealth), in order of importance, become

- Population Growth
- (Real) Prime Interest Rate

With population growth in, the mortgage interest becomes statistically insignificant and is left off this list.

However, our preference is for the first interpretation. Other studies (Heim 2008c) have found that other variables not in the formulations above (e.g., housing price levels relative to income, or the accelerator) powerfully influence housing demand. This it turn powerfully influences durables demand (appliances and furniture for new housing). Neither component of housing demand shows statistically significant when used separately in the above model of durables demand, yet clearly play a role in determining housing demand, which does. Hence, our preference for the formulation that includes housing demand as one of the determinants. Housing demand is also systematically related to the mortgage interest rate in the Heim 2008c study. In table 4 above, the suspected reason it is significant when housing demand is left out, is that it is proxying for its effect on housing demand. When the housing demand variable is left in the model, the mortgage interest rate becomes insignificant. Its influence on demand for durables is already picked up by the housing demand variable.

Overwhelmingly, as Keynes predicted for consumption in general, current year disposable income is the most influential factor. Next in importance, again somewhat unexpectedly, was our measure of crowd out – limitations on consumer credit availability, which explained 18% of the variance. As was explained earlier, one might expect the interest rate underlying most consumer credit (the Prime Rate used here) would pick up variations in demand resulting from credit shortages. This would probably be true if the Prime Rate were a market determined rate, but it not. It is an administered rate, rigidly set by banks to maintain exactly a 3 point spread between itself and the federal funds rate, which is directly set by the Federal Reserve. Hence, the amount of loanable funds available which are taken up by the government as it attempts to finance its budget deficits becomes a key measure of how much credit is available to consumers, even more so than the prime rate itself.

The next most important variable affecting demand for consumer durables was the exchange rate (unlike our results for consumption overall), explaining 14% of the variance in durable goods demand, perhaps not unexpectedly, since durables such as foreign autos, constitute a significant part of durables demand. Wealth, as represented by the Dow Jones Composite Index, explained 5% of the variance, housing demand (5%) and the prime interest rate 1%

The population growth variable is negatively related to demand for durables. We suspect this may be because, holding income constant, as in our tests, increased population (family size) requires reductions in discretionary spending (durables) because of increased non discretionary spending on nondurables (food, clothing). Other variables found non-significant were the age distribution of the population (ratio of

16-24 year olds to those 65 and over), the accelerator, and the affordability of homes (ratio of house prices to per capita income).

4.2 DETERMINANTS OF CONSUMER DEMAND FOR NONDURABLE GOODS

In this section, we test the statistical significance of each of the theorized determinants of consumer demand for nondurable goods. Results are presented in table 4 below.

Table 4: Stepwise Addition - Regression of Hypothesized Determinants of Nondurable Goods

$$\Delta C_{ND} = f\,[\,\beta_1\,\Delta(Y\text{-}T_G)_t, + \beta_{2T\&2G}\,\Delta(\text{Crowd Out})_t, + \beta_3\,\Delta DJ_{-3}, + \beta_4\,\Delta PR, + \beta_5\,\Delta POP]$$

	$\Delta(Y\text{-}T_G)$	$\Delta\ T_G$	ΔG	$\Delta\ DJ_{-3}$	ΔPR	ΔPOP
R^2/AdjR^2(DW)	$\beta_1(t)$	$\beta_{2T}(t)$	$\beta_{2G}(t)$	$\beta_3(t)$	$\beta_4(t)$	$\beta_5(t)$
60/60% (1.3)	.22(16.8)					
77/76% (1.7)	.18(13.4)	.13 (5.0)	.04 (0.5)			
83/81% (2.2)	.16(12.3)	.12 (5.0)	.01(0.2)	.29 (3.6)		
85/83% (2.1)	.16(13.1)	.17 (5.5)	-.03(-0.4)	.26 (3.3)	-1.96(-2.3)	
86/84% (2.1)	.13(5.5)	.18 (5.9)	-.07(-1.1)	.28 (3.7)	-1.96(-2.4)	.003 (1.7)

(*) t- statistics of 2.0 = 5% significance; t - statistics of 2.7 = 1% level of significance.

(**) Using Separate T_G, G Variables to Represent Crowd Out.

The factors whose variation is most associated with change in consumption during the 1960 - 2000 period, in order of importance were

- Disposable income (60%)
- Crowd out (as indicated by the government deficit) (17%)
- Wealth (as indicated by the Dow Jones Composite average lagged 3 years) (6%)
- (Real) Prime Interest Rate (2%)
- Population Growth (1%)

Overwhelmingly, as Keynes predicted for consumption in general, current year disposable income is the most influential factor, explaining 60% of the variance. Next in importance, again somewhat unexpectedly, was our measure of crowd out – limitations on consumer credit availability, which explained 17% of the variance. One might expect the interest rate underlying most consumer credit (the Prime Rate used here) would pick up variations in demand resulting from credit shortages. This would probably be true if the Prime Rate were a market determined rate, but it not. It is an administered rate, rigidly set by banks to maintain exactly a 3 point spread between itself and the federal funds rate, which is directly set by the Federal Reserve. (Heim 2009, p. 58) Hence, the amount of loanable funds available which are taken up by the government as it attempts to finance its budget deficits becomes a key measure of how much credit is available to consumers. The next most important variable affecting demand for consumer

nondurables was wealth, as represented by the Dow Jones Composite Index, explained 6% of the variance, and finally, the prime interest rate, explaining 2%. A number of other variables were tested as possible determinants in the stepwise process, but none were found statistically significant. They included the exchange rate, population size, population 16-24 as a percent of population over 65 (and separately, this same variable divided by aggregate income), residential housing demand, and the ratio of real house prices to real income. However, population size, though insignificant at the 5% level, population size was significant at the 9 % level.

4.2 DETERMINANTS OF CONSUMER DEMAND FOR SERVICES

In this section, we test the statistical significance of each of the theorized determinants of consumer demand for services. Results are presented in table 5 below.

The results below are presented for each step in the stepwise regression procedure. Showing each stepwise equation's results individually provides information about regression coefficient stability. The stability of regression coefficients increases significantly as the total variance (R^2) explained by the model increases. Until the model explains approximately 70% of the variation in services consumption, regression coefficients change so much when variables are added they are simply unreliable.

Table 5 below shows stepwise addition of variables to the model. The results indicate the factors whose variation is most associated with change in consumption during the 1960 - 2000 period.

Table 5: Hypothesized Determinants of Consumer Demand for Services (T_g, G Variables For Crowd Out)

$\Delta C_s = f\ [\ \beta_1 \Delta(Y\text{-}T_G)_t, + \beta_{2T\&2G} \Delta(\text{Crowd Out})_t, + \beta_3 \Delta POP + \beta_4 \Delta DJ_{-2}, + \beta_5 \Delta(16\text{-}24)/65, + \beta_6 \Delta\ MORT\ ]$

	Δ(Y-T_G)	Δ T_G	ΔG	ΔPOP	DJ $_{-2}$	Δ16-24/65	ΔMORT
R^2/Adj R^2 (DW)	β_1(t)	β_{2T} (t)	β_{2G}(t)	β_3(t)	β_4(t)	β_5(t)	β_6(t)
40/40% (2.2)	.44(21.0)						
58/56% (2.2)	.37(15.3)	.08(1.6)	.39 (3.1)				
70/67% (1.7)	.26 (7.2)	.13 (2.9)	.21 (1.8)	.009 (3.8)			
77/74% (1.7)	.22(6.3)	.09 (2.1)	.18 (1.8)	.010 (4.2)	.37 (3.5)		
79/76% (1.8)	.21 (5.9)	.08 (2.1)	.17 (1.7)	.011 (4.7)	.39 (3.9)	-247.7(-2.1)	
79/76% (1.5)	.19 (5.2)	.10 (2.5)	.14(1.4)	.013 (4.8)	.37 (3.7)	na	-5.5(-2.0)
81/78% (1.6)	.18 (5.1)	.10 (2.4)	.13 (1.4)	.013 (5.1)	.39 (4.0)	-212.9(-1.8)	-4.7(-1.7)

(*) t- statistics of 2.0 = 5% significance; t - statistics of 2.7 = 1% level of significance

(**) There is a little overlap in the variance explained by the mortgage interest rate variable and the % of young people in the population variable, leaving both significant at the 5% level when entered alone, but reducing their significance to the 7.5 and 9.7% level when used together. We will view the results as indicating both make a small, reasonably systematic, contribution to explaining consumer demand for services.

As Keynes predicted for consumption in general, current year disposable income is the most influential factor related to consumer spending on services, explaining 40% of the variance. Next in importance, we again found the crowd out variables important, adding 18% to explained variance. Third most important, explaining 12% of the variation in demand for services was population growth. Fourth most important was consumer wealth, adding 7%. Fifth most important were the percentage of younger people in the population relative to those 65 and over, adding 2% to explained variance. Spending on services was inversely found related to the percentage of young people in the population, suggesting that families with younger members have less to spend on services, as might be expected. Sixth .most important was the mortgage interest rate, which added an additional 2% to explained variance. It also was found inversely related to demand for services. This may be because mortgage payments are often the largest single item in the consumer's budget and many are variable rate mortgages. The inverse relationship may indicate the need to accommodate the need to increase monthly mortgage payments due to interest rate increases by cutting back on other, more discretionary spending such as services.

Other variables were subsequently entered into the regression in stepwise order. None were found even marginally statistically significant. They included the prime interest rate, housing the ratio of real housing prices to real income, housing demand and the exchange rate.

REFERENCES

Heim 2008a. "What Determines Market Demand For Investment Goods?". *International Journal of Business Strategy*. Vol. 8, No. 1. 2008

Heim, J.J. 2008b. "The Consumption Function". *Review of Business Research*, Vol. 8, No.2 (June 2008)

Heim, J.J. 2008c. ""Determinants of Demand for Different Types of Investment Goods". *RPI Working Papers in Economics, No. 902.* Department of Economics, Rensselaer Polytechnic Institute.

Heim, J.J. 2009. "Which Interest Rate Seems Most Related To Business Investment?" *American Society of Business and Behavioral Sciences E - Journal.* Vol. 5., No. 1. 2009

Keynes, J.M. 1936. *The General Theory of Employment, Interest and Money*. New York. Harcourt, Brace and World, Inc. (1964 Printing)

Newey, W.K. and West, K.D. 1987. "A Simple Semi-definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix" *Econometrica*, pp. 703-708.

Friedman, M. 1957. *A Theory of the Consumption Function.* Princeton University Press. Princeton.

Modigliani, F. & Ando, A. 1963. "The Life Cycle Hypothesis of Saving", *American Economic Review,* Vol. 53 (1).

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

Griffiths, W. Hill, R and Judge, G. 1993. *Learning and Practicing Econometrics*. New York: John Wiley and Sons.

A Method for Separating Income and Substitution Effects of Exchange Rate Changes

John J. Heim*

ABSTRACT

Regression estimates of exchange rate total effects on aggregate demand are broken into separate income and substitution effects. Total effects (substitution and income) estimates can seem contrary to theory and common sense. Separating them into their two components shows this is not the case. The separation method also provides a simple test to determine if imports are normal or inferior goods. The paper finds consumer imports are normal goods, but investment imports are inferior goods. The paper shows that if import total effects exceed domestic total effects, imports are a normal good. If smaller, they are inferior goods. (JEL: E00, F40, F43)

***Keywords**: Macroeconomics, International Trade, Imports, Exports, Exchange Rate*

1.1. INTRODUCTION: CONSUMER DEMAND

A recent study indicated the U.S. exchange rate was systematically related to the level of consumer spending, particularly on imports (Heim 2008). In this study, demand for domestically produced or imported consumer goods was regressed on a range of variables commonly held to be determinants of consumer spending, including disposable income, interest rates, consumer wealth and the relative price of imports compared to domestic goods, as measured by the exchange rate. Higher exchange rate values indicate more foreign currency can be bought per dollar, which in turn can mean cheaper import prices. An additional determinant, measured by the government deficit, provides a measure of the extent to which consumers are crowded out of the credit market by government borrowing. The spending and tax variables are reported separately, rather than as a net figure, since preliminary testing indicated deficit increases due to increased government spending restrict consumer credit less than tax cuts. Key regression findings are summarized in Table 1 below:

Subscripts of zero on variables, or no subscripts at all, mean the current period value of the variable is used. Subscripts with negative signs indicate the number of years the variable is lagged. The equations are estimated using first differences of the data e.g., $\Delta(C)_0$ to help reduce multicollinearity and autocorrelation problems, and sometimes non-stationarity problems in otherwise highly correlated data.

Adding the exchange rate variable seemed to have a major influence on demand for imported consumer goods, adding 6%-points to explained variance. However, explained variance did not increase in the other two models. These exchange rate coefficients show the total effect of an exchange rate

* Professor of Economics, Rensselaer Polytechnic Institute, Troy, NY

change on consumer demand. Below this effect will be separated into its income and substitution effect components.

Table 1: Determinants of Demand for Total, Imported and Domestically Produced Consumer Goods

$\Delta(C)_0$	$=.63\Delta(Y-T_G)$	$+.47\Delta T_{G)}$	$+ .06\Delta G_0$	$- 6.22\Delta PR_0$	$+.60\ \Delta DJ_{-2}$	$+ 2.69\Delta XR_{AV}$	R^2=91%.
(t)	(18.2)	(4.3)	(0.5)	(-3.4)	(4.0)	(1.5)	D.W.=1.7
$\Delta(M_{m\text{-}ksm})_0$	$=.06\Delta(Y-T_G)$	$+.27\Delta T_G$	$-.18\ \Delta G_0$	$- 3.94\ \Delta PR_0$	$+.26\ \Delta DJ_{-2}$	$+4.33\ \Delta XR_{AV}$	R^2=83%
(t)	(2.8)	(4.6)	(-1.9)	(-2.6)	(1.8)	(2.9)	D.W.=1.5
$\Delta(C\text{-}M_{m\text{-}ksm})$	$=.57\Delta(Y-T_G)$	$+.20\Delta T_G$	$+.24\ \Delta G$	$- 2.28\ \Delta PR$	$+.34\Delta DJ_{-2}$	$- 1.64\ \Delta XR_{AV}$	R^2=74%
(t)	(16.0)	(1.6)	(1.3)	(-0.7)	(2.8)	(-1.0)	D.W.=1.8

* Nominal Exchange Rate Used)

Where

C = Total Consumption

$M_{m\text{-}ksm}$ = Consumer Imports

$C- M_{m\text{-}ksm}$ = Consumer Goods Domestically Produced

$Y-T_G$ = Disposable Income

T_G = Government Receipts

G_0 = Government Spending on Goods & Services

PR_0 = Real Prime Interest Rate

DJ_{-2} = A Wealth Measure: the Dow Jones Composite Average

XR_{AV} = The average nominal exchange rate (trade weighted) for the current and past three years

The Federal Reserve's trade weighted nominal Broad exchange rate was used above; a related study (Heim 2009) used the Federal Reserve's real exchange rate in the same models, and yielded similar results except the exchange rate variable, whose results varied somewhat, as expected. See Table 2.

Adding the exchange rate variable to the total, imports, and domestically produced consumer goods models in Table 2 increases explained variance by 2%, 8% and 0% respectively. Notice the estimated total effect of exchange rates on consumer demand for imports is larger than the estimated total effect on domestic goods, and that the estimated total effect on domestic goods is negative. Later we will show that this implies consumer imports are normal goods and that the substitution effect outweighs the income effect.

Table 2: Determinants of Demand for Total, Imported and Domestically Produced Consumer Goods

$\Delta(C)_0$	$=.66\Delta(Y-T_G)$	$+.49\Delta T_{G)}$	$+ .04\Delta G_0$	$- 6.92\Delta PR_0$	$+.62\ \Delta DJ_{-2}$	$+ 2.83\Delta XR_{AV}$	R^2=91%.
(t)	(29.2)	(5.7)	(0.3)	(-3.2)	(4.9)	(3.2)	D.W.=1.7
$\Delta(M_{m\text{-}ksm})_0$	$=.11\Delta(Y-T_G)$	$+.30\Delta T_G$	$-.20\ \Delta G_0$	$- 5.00\ \Delta PR_0$	$+.34\ \Delta DJ_{-2}$	$+3.03\ \Delta XR_{AV}$	R^2=85%
(t)	(6.3)	(5.0)	(-2.0)	(-3.5)	(4.5)	(5.6)	D.W.=1.8
$\Delta(C\text{-}M_{m\text{-}ksm})$	$=.55\Delta(Y-T_G)$	$+.19\Delta T_G$	$+.24\ \Delta G$	$- 1.92\ \Delta PR$	$+.28\Delta DJ_{-2}$	$- .20\ \Delta XR_{AV}$	R^2=74%
(t)	(16.2)	(1.5)	(1.3)	(-0.6)	(1.9)	(-0.2)	D.W.=1.8

*Real Exchange Rate Used

1.2. INTRODUCTION: INVESTMENT DEMAND

Similarly, the (2008a) study indicated the exchange rate played the following role in determining the level of spending on domestic and imported investment goods:

Table 3: Determinants of Demand for Total, Imported and Domestically Produced Investment Goods*

$\Delta I(t) =$									
.28ΔACC	+.95ΔDEP	+1.48ΔCA_{-1}	+.52ΔT	-.63ΔG	-6.40Δr_{-2}	-.20 ΔDJ_{-2}	+.16 ΔP_{-2}	+6.92ΔXR	R^2=.89
(7.9)	(3.2)	(1.0)	(5.6)	(-2.9)	(-3.8)	(-0.8)	(0.9)	(3.8)	DW 1.9
$\Delta(M_{ksm})=$									
.04ΔACC	+.38ΔDEP	+1.52ΔCA_{-1}	+.07 ΔT	-.22ΔG	+1.54Δr_{-2}	+.19 ΔDJ_{-2}	-.10ΔP_{-2}	+2.52ΔXR	R^2=.70
(1.8)	(4.1)	(2.2)	(2.5)	(-2.1)	(1.3)	(2.6)	(-1.1)	(-2.1)	DW2.4
$\Delta(I-M_{ksm})=$									
.24ΔACC	+.57ΔDEP	-.04ΔCA_{-1}	+.44 ΔT	-.41ΔG	-8.00Δr_{-2}	-.38ΔDJ_{-2}	+.26ΔP_{-2}	+4.39ΔXR	R^2=.85
(8.7)	(1.9)	(-0.0)	(5.4)	(-2.0)	(-4.9)	(-1.5)	(1.5)	(1.8)	DW1.6

*Nominal Exchange Rate Used

Where

I = Total Investment Demand

M_{ksm} = Demand for Imported Investment Goods

$I-M_{ksm}$ = Demand for Domestically Produced Investment Goods

ACC = The Accelerator, a measure of the growth rate of the GDP each year

DEP = Depreciation Levels of Capital Equipment

CA_{-1} = % of manufacturing capacity currently being utilized, lagged one year

ΔP_{-2} = Corporate profits, lagged two years

Δr_{-2} Real prime interest rate, lagged two years

Other variables used are defined in the consumption equations. Subscripts have the same meanings as before and first differences of the data are again used. Notice that, unlike the consumption equations, in the investment equations the larger estimated total effect of a change in exchange rates on demand is in the domestic demand equation, not import demand. We will show later that this is a sign that investment imports are inferior goods.

These coefficients estimates of the exchange rate total effects, will be separated below into income and substitution effects. The exchange rate does add significantly to the explanatory power of some equations. The exchange rate appears to have a major influence on demand for imported investment goods, adding 6%-points to explanatory power, but also adds 4% to explained variance when added to the total investment demand and 2% to domestically produced investment goods demand models.

We note that the regression results indicate that for every single - point (~ 0.8%) decline in nominal Broad exchange rate from 2000 levels, making imports more expensive, there appears to be a $4.39 billion decrease in demand for domestically produced investment goods as well as a 2.52 billion decrease in demand for imported investment goods. We will show later this finding for domestic goods is not nearly as irrational as it appears to be at first blush. It is totally consistent with our estimates of the sum of

income and substitution effects for investment goods, particularly our estimates showing investment goods, as a group, are inferior goods and that this trait is transmitted through the substitution effect.

Results for the same investment demand model using the real Broad exchange rate (Heim 2009), are presented in Table 4 below. They were very similar; with the total effect estimates above not changing much for any of the variables, except the exchange rate, which was expected:

Table 4: Determinants of Demand for Total, Imported and Domestically Produced Investment Goods*

ΔI (t) =									
.28ΔACC	+1.37ΔDEP	+.69ΔCA_{-1}	+.52ΔT	-.61ΔG	-8.46Δr_{-2}	-.10 ΔDJ_{-2}	+.35 ΔP_{-2}	+4.97ΔXR	R^2=.89
(6.9)	(4.7)	(0.4)	(5.3)	(-3.4)	(-3.5)	(-0.4)	(2.0)	(4.2)	DW 2.3
$\Delta(M_{ksm})$=									
.05ΔACC	+.46ΔDEP	+1.25ΔCA_{-1}	+.07 ΔT	-.14ΔG	+1.12Δr_{-2}	+.30 ΔDJ_{-2}	-.11ΔP_{-2}	-.40ΔXR	R^2=.64
(1.9)	(4.5)	(1.4)	(2.0)	(-1.7)	(0.7)	(3.4)	(-1.09)	(-0.7)	DW2.1
$\Delta(I-M_{ksm})$=									
.24ΔACC	+.91ΔDEP	-.15ΔCA_{-1}	+.45 ΔT	-.47ΔG	-9.59Δr_{-2}	-.40ΔDJ_{-2}	+.47ΔP_{-2}	+5.37ΔXR	R^2=.88
(7.8)	(3.0)	(-0.4)	(6.0)	(-2.9)	(-7.3)	(-1.9)	(4.1)	(4.1)	DW2.1

*Real Exchange Rate Used

Adding the exchange rate variable increases explained variance in the total, imports and domestically produced goods models by 4, 0 and 5% respectively.

Ultimately, the sign of the total effect of an exchange rate change on spending is the sum of its two parts: the "pure" income effect and the substitution effect. If the substitution effect is negative, and large, it may "swamp" the positive income effect, and leave a negative sign. If not, the sign may be positive. For example, by separating the total effects of a change in exchange rates in tables 1-4 above into their separate income and substitution effects, we can explain results which otherwise seem illogical, or at least puzzling. We (again) note that the regression results indicate that for every single - point decline in the nominal exchange rate (imports more expensive) there appears to be a $4.39 billion decrease in demand for domestically produced investment goods and a 2.52 billion decrease in demand for imported investment goods. The decrease in import demand is understandable since if import prices are increasing, the real income effect is negative. Further, if investment goods are normal goods, the increase in relative import prices should cause further movement out of imports. However, we do not find demand for domestically produced investment goods increases. This may be because the negative income effect swamps the positive substitution effect, or if imports are an inferior good, the overall effect may be out of domestic goods because both effects are negative. To know for certain, we must parse out the separate income and substitution effects.

There is a simple method for breaking down the total effect of an exchange rate change into its income effect and substitution effect components, so that we can resolve such questions. The method will use information we already have on the estimated total effect of exchange rate changes on and the information inherent in the following identities for domestic (D) and imported (M) goods:

Total \$ Effect (T_D) = \$ Income Effect (I_D) + \$ Substitution Effect (S_D)

Total \$ Effect (T_M) = \$Income Effect (I_M) + \$ Substitution Effect (S_M)

Since economic theory holds that *the real value* of (pure) substitution effects are symmetric except for sign, this means that in money terms

$S_D = -(S_M)$

i.e., $T_D - I_D = -(T_M - I_M)$

Since T_D and T_M are known from regression analysis, this leaves us with one equation in two unknowns: I_D and I_M, *the dollar value* of the pure income effect. However, we will show that these two pure income effects must be the same. Therefore, we have but one equation in one unknown to solve, which is a simple task. This will be done further below.

2. OTHER EMPIRICAL ESTIMATES OF INCOME AND SUBSTITUTION EFFECTS

Elmendorf (1996) has excellent estimates the total effect of changes in interest rates on consumption, but does not break them down into income and substitution effects. Baker, Gruber and Milligan (2003) examined the impact of Canada's government retirement programs on work incentives, but, again, did not attempt to separate the total effect of the onset of retirement income into income from substitution effects. Others have attempted to separate these two effects, but have used "hypotheticals", such as "what if" survey responses, instead of data to estimate one of the effects. The other effect is then inferred from indirect evidence. For example, Kimball and Shapiro (2008) estimated income effects of a income increase by asking survey respondents "what if" they won a sweepstakes. How would their work habits change if they won an independent income for life? Using these responses as "income effects", restrictions from labor theory, and known total effects of approximately zero, they inferred substitution effects of a magnitude similar to income effects.

3. INCOME AND SUBSTITUTION EFFECTS IN THEORY

In standard economic theory, utility is derived from consumption. Utility varies as the combination of goods consumed changes. The combinations considered here are domestically produced goods and imported goods. (D) represents the bundle of domestically produced goods consumed by businesses (investment goods) or consumers (consumer goods); (M) represents the bundle of investment or consumer imports. The utility relationship is given as

Utility (U) = f (D, M)

One example of this relationship might be

, $U = D*M$

→ $D = U/M$

i.e., utility grows in both D, M subject to diminishing returns . This function provides us with an example of a standard – shaped hyperbolic indifference curve in which U is increasing in D, M.. For example, we might find (were utility cardinally countable),

$U = 100 = D_{10} * M_{10}$

$U = 400 = D_{20} * M_{20}$

Where the subscripts on D, M represent the real quantities consumed. Consumers (and businesses) choose utility maximizing combinations of D and M, given their budget constraints

$P_D * D = \text{Budget (B)} - P_M * M$

If the budget is 20 and prices are $P_D = P_M = \$1$, the feasible combinations of goods the consumer can buy with a the budget is given by (D = 20 – M), where "20" might be interpreted as $20 billion and U =100 at utility maximization. The combination of goods that fully expends the budget and provides the highest utility level is D=10 and M=10. Other purchasable combinations provide lower utility, for examples

$D_{=18} = 20 - M_{=2} = \rightarrow \quad U = 15*5 = 36$

$D_{=15} = 20 - M_{=5} = \rightarrow \quad U = 15*5 = 75$

$\underline{D_{=10} = 20 - M_{=10} = \rightarrow \quad U = 10*10 = 100}$

$D_{=5} = 20 - M_{=15} = \rightarrow \quad U = 15*5 = 75$

$D_{=2} = 20 - M_{=18} = \rightarrow \quad U = 15*5 = 36$ Etc….

We can see the *dollar amounts of each good which maximize utility are the same*: ($1*D = $1*M =10). In this specific case, the quantities are also the same. In the more general case dollar equivalence will remain, but quantity equivalence will not. Theory suggests the ratio of the goods selected is inversely related to the ratio of their prices, which implies that for budget (income level) changes, prices remaining constant, utility maximization requires allocation of equal money amounts to both products when income changes. This is true for any income change, from zero income on up. We can measure this "pure income effect" by simply increasing the consumer's budget without changing the relative prices of (D,M). If income doubles, the vertical and horizontal intercepts on the budget constraint double, but the slope of the budget line (P_M/P_D) remains the same. It now touches the new (and higher) indifference curve where the new curve has the same marginal rate of substitution (MRS) as before. Therefore, both before and after the change

$MRS = \partial D/\partial M = P_M/P_D$ (Prager, 1993)

Which implies $\partial D * P_D = \partial M * P_M$

Or in discrete terms $\Delta D * P_D = \Delta M * P_M$

Clearly this indicates that (except for sign) the money value of substitution effects must be identical. The dollar amount substituted out of one good must equal the amount substituted into the other.

This formulation also clearly indicates that if incomes change *not due to price changes, i.e., prices remaining constant*, we have a "pure income effect". In order for the above condition above to be met if income is increasing, *spending on both goods must change by the same amount*, no matter what the initial income level. This must hold for all budget levels, e.g., $\Delta 2D * P_D = \Delta 2M * P_M$, etc.

Hence we conclude this standard theoretical formulation shows that the money value of the "pure" income effect must be the same for two goods when income changes, prices constant. The money value

of pure substitution effects are also the same (except for sign). These two findings are of key importance in inferring income and substitution effects from data on total effects.

As an example, suppose the consumer or businesses' budget is doubled, with prices remaining constant at \$1 = P_D = P_M. The budget constraint then becomes (\$1*D = \$40 –\$1*M) or (D = 40 –M). As shown below, no other combination of goods purchasable with the \$40 budget yields as much utility as twice the original quantities and money expenditures on each.

$D_{=30}$ = 40 – $M_{=10}$ = → U = 10*30 = 300

$D_{=25}$ = 40 – $M_{=15}$ = → U = 15*25 = 375

$D_{=20}$ = 40 – $M_{=20}$ = → U = 20*20 = 400

$D_{=25}$ = 40 – $M_{=15}$ = → U = 25*15 = 375

$D_{=30}$ = 40 – $M_{=10}$ = → U = 30*10 = 300 Etc.

Of course, utility is counted ordinally, not cardinally, so we don't know that 100 is really the initial value of U, or that 400 its later value. Nonetheless, standard utility theory shows what is shown in the utility curve above: utility increases in D and M, utility curves are convex, everywhere dense, and don't cross (Wold and Jureen, 1953). This gives them the same general shape, leading to the same results.

We can see the dollar equivalence of the pure income effect, since any optimal solution leaves the ratio of goods consumed inversely equal to the ratio of their prices, as before. Hence. $D*P_D$ stays equal to $M*P_M$. The total effect is a pure income effect, since relative prices remain the same. The money value of the income effect is the same for both imports and domestically produced goods.

The result is not dependent on the indifference curve shape, provided it is broadly convex to the origin, i.e., allows for diminishing marginal utility. Other forms will yield the same result.

Hence, the results are the same: the money value of income effects for D, M are the same and the substitution effects are the same (except for sign). Further support for this conclusion is given in the next section, where we show that the only way to explain our empirical results for total effects is by assuming this same equivalence of income and substitution effects.

Does this also hold if the relative prices of goods is different from the ratio used above (i.e., 1/1)? The example below assumes a relative price ratio of (2_D/1_M = 2), yielding a budget constraint of

\$2D= 20 - \$1M → D =10-1/2M Then we have

$D_{=9}$ = 20 – $M_{=2}$ = → U = 9*2 = 18

$D_{=8}$ = 20 – $M_{=4}$ = → U = 8*4 = 32

$D_{=5.5}$ = 20 – $M_{=9}$ = → U = 5.5*9 = 49.5

$D_{=5}$ = 20 – $M_{=10}$ = → U = 5*10 = 50

$D_{=2}$ = 20 – $M_{=16}$ = → U = 2*16 = 32

$D_{=1}$ = 20 – $M_{=19}$ = → U = 1*19 = 19 Etc.

If the consumer's income doubles,

$D_{=18}$ = 40 – $M_{=4}$ = → U = 18* 4 = 72

$D_{=16} = 40 - M_{=8} = \rightarrow$ U = 16* 8 = 128

$D_{=10.5} = 40 - M_{=19} = \rightarrow$ U = 10.5* 19 = 199.5

$D_{=10} = 40 - M_{=20} = \rightarrow$ U = 10*20 = 200

$D_{=4} = 40 - M_{=32} = \rightarrow$ U = 4*32 = 128

$D_{=1} = 40 - M_{=38} = \rightarrow$ U = 1*38 = 38

Etc.…

Again we notice that the utility – maximizing combination of goods doubles for both goods when income doubles. Other examples will show the same. Since price ratios are unchanged, there are no substitution effects to modify this proportionality result.

In the current case

$MRS = \partial D/\partial M = P_M/P_D$ (Prager, 1993)

Which implies $\partial D*P_D = \partial M*P_M$

Or in discrete terms $\Delta D*P_D = \Delta M*P_M$

i.e., $\Delta D*(\$2) = \Delta M*(\$1)$

which implies M = 2D. Pure income effects do not always result in equal quantities of both goods selected, but always do result in equal money expenditures on the two goods when incomes change from any level to another.

4. DERIVING INCOME AND SUBSTITUTION EFFECTS OF EXCHANGE RATE CHANGES WHEN TOTAL EFFECTS ARE KNOWN

Regression coefficients (β) in the demand functions above provide estimates of the total effect on domestic and imported goods demand, i.e.,

Total Effect (T_{DorM}) = (β_{DorM}) = Income Effect (I_{DorM}) + Substitution Effect (S_{DorM})

When only pure income effects are considered, the money (or “real”) value of income effects are the same for both groups of goods: imports and domestic. Substitution effects are also the same in money value, except for sign. Using this means that our domestic and imported investment goods equations above become two equations with two unknowns (I and S) to be solved, where, from above, we take (S_D) = (- S_M) and $I_D = I_M$.

4.1. INVESTMENT DEMAND: DERIVING INCOME AND SUBSTITUTION EFFECTS FROM ESTIMATED TOTAL EFFECTS:

Below, six cases are evaluated to determine income and substitution effects of exchange rate changes on investment goods. An additional six cases test consumer goods in the same way.

- Cases 1 and 3 test, whether imported investment goods, as a group, are normal goods. Nominal and real exchange rate changes are tested separately).

- Cases 2 and 4 test whether imports should be considered an inferior good, with the inferiority trait passing through the substitution effect (again, nominal and real exchange rates changes are tested separately).

Using the method described, these four tests, applied separately to investment and consumption, lead this study to conclude

- consumer imports are normal goods (as a macroeconomic grouping).
- investment imports are inferior goods (as a macroeconomic grouping), with the inferiority trait passing through a negative substitution effect.

Details of these tests are provided in sections 4.1.(1-4) and 4.2.(1-4) below.

4.1.1. CASE 1: IMPORTS ARE NORMAL INVESTMENT GOODS; NOMINAL EXCHANGE RATE USED TO ESTIMATE TOTAL EFFECTS

Assume income(I) and substitution (S) effects are normal for investment goods when import prices fall due to an increase in exchange rates: i.e.,

$I_D , I_M > 0$; $(I_D = I_M)$; $S_D < 0, S_M > 0$, $(-S_D = S_M)$

Let

Income Effect$_D$ (I_D) + Substitution Effect$_D$ (S_D) = \$4.39B = Total Effect$_D$ Estimate

Income Effect$_M$ (I_M) + Substitution Effect$_M$ (I_M) = \$2.52B = Total Effect$_M$ Estimate

Total Effect: $= I_D + S_D = \$4.39B$	$\| = I_M + S_M = \$2.52B$
$\rightarrow S_D = - I_D + 4.39$	$\| \rightarrow \quad S_M = - I_M + 2.52$
$\rightarrow \quad -S_D = I_D - 4.39$	| .
$-S_D$	$\| = S_M$

therefore

$I_D - 4.39 = -I_M + 2.52$

$(I_D + I_M) = 2 I_D = 6.91$ (since $I_M = I_D$)

$I_D = \$3.455B$

But, though we know income and total effects, we cannot deduce from them substitution effects consistent with our assumption of normality, which requires substitution out of domestic goods and into imports. No negative number for substitution effects, when added to positive income effects of 3.455B, will give use a total effect greater than the income effect alone, i.e., 4.49B for domestic goods. No positive substitution effect into imports, when added to a positive income effect of 3.455B will give us the smaller total effect of 2.52 billion we have estimated..

Income Effect:	\$ +3.455B I_D	\$ +3.455B I_M
+Substitution Effect	\$ - ? . S_D	\$ + ? . S_M
(Must =)Total Effect	\$ = 4.39 B T_D	\$ = 2.52 BT_M

Conclude: Hypothesis that investment imports as a group are a normal goods leads to irrational result

4.1.2. CASE 2: IMPORTS ARE INFERIOR INVESTMENT GOODS; NOMINAL EXCHANGE RATE USED TO ESTIMATE TOTAL EFFECTS

However, If we assume that imports are an inferior good, i.e., the substitution effect is negative i.e., out of imports as they become cheaper and positive i.e., into domestic goods, and that the inferiority trait is passed through the substitution effect, we get

Income Effect:	$ +3.455B I_D	$ +3.455B I_M
Substitution Effect	$ + .935B. S_D	$ - .935B. S_M
Total Effect	$ = 4.39 B T_D	$ = 2.52 B T_M

which is consistent with our earlier estimates for income effects and total effects, and consistent with our earlier finding that income and substitution effects (in absolute terms) had to be the same for both groups of goods. This example indicates that as a macroeconomic aggregate, imported investment are an inferior good based on our regression coefficient estimates of total effects of a change in nominal exchange rates. The conclusion, theory requires the money value of "pure" income effects to be the same for both types of goods. The conclusion, theory requires substitution effects for both types of goods to be the same in money terms, except for sign

4.1.3. CASE I.3: IMPORTS ARE NORMAL INVESTMENT GOODS; REAL EXCHANGE RATE USED TO ESTIMATE TOTAL EFFECTS

Using the same methods as above, it can be shown that the income effect is $2.485 billion and that no normal goods substitution effect, when added to it, can give us the total effects previously obtained.

Income Effect:	$ +2.485B I_D	$ +2.485B I_M
Substitution Effect	$ - ? . S_D	$ + ? . S_M
Total Effect	$ = 5.37 B T_D	$ = - .40 B T_M

Conclude: Irrational result: no substitution out of D and into M can lead to our estimated total effects of $5.37B and $-0.40B respectively. This implies incorrectness of original hypothesis that D, M are normal goods.

4.1.4. CASE I.4: IMPORTS ARE INFERIOR INVESTMENT GOODS; REAL EXCHANGE RATE USED TO ESTIMATE TOTAL EFFECTS

However, If we (again) assume that imports are an inferior good, i.e., the substitution effect is negative for imports as they become cheaper and positive for domestic goods, and that the inferiority trait is passed through the substitution effect, we get

Income Effect:	$ +2.485B I_D	$ +2.485B I_M

Substitution Effect $ +2.885B. S_D $ - 2.885B. S_M

Total Effect $ = 5.37 B T_D $ = -.40 B T_M

which is consistent with our earlier estimates for income effects and total effects. This example, using real exchange rates, matches our earlier results using nominal rates. It indicates that as a macroeconomic aggregate, imported investment goods are inferior goods, based on 1) our regression coefficient estimates of total effects of a change in real exchange rates; 2) the conclusion, theory requires the money value of "pure" income effects to be the same for both types of goods; 3) the conclusion, theory requires substitution effects for both types of goods to be the same in money terms, except for sign.

4.2. CONSUMPTION DEMAND: DERIVING INCOME AND SUBSTITUTION EFFECTS FROM TOTAL EFFECT ESTIMATES OF A CHANGE IN EXCHANGE RATES

Are consumer imports also inferior goods? Using the above methods we now examine the sensitivity of consumer demand to a change in either the nominal or real exchange rate, using estimates of the total effect from the regression coefficient results presented in Section 1 above.

Coefficient On Exchange Rate Variable

	Nominal .	Real .
Total Effect Estimate (T_D):	$ - 1.64B	$ - .20B
Total Effect Estimate (T_M):	$ + 4.33B	$ +3.03B

4.2.1. CASE 1: IMPORTS ARE NORMAL CONSUMER GOODS; NOMINAL EXCHANGE RATE USED TO ESTIMATE TOTAL EFFECTS

Assume: When import prices fall due to a change in exchange rates, income and substitution effects are normal: i.e.,

$$I_D , I_M > 0;\ I_D = I_M \qquad S_D < 0,\ S_M > 0,; -S_D = S_M$$

Let

Income Effect$_D$ (I_D) + Substitution Effect$_D$ (S_D) = - 1.64 Total Effect$_D$ Estimate

Income Effect$_M$ (I_M) + Substitution Effect$_M$ (I_M) = + 4.33 Total Effect$_M$ Estimate

Total Effect = $I_D + S_D$ = $ -1.64B | $I_M + S_M$ = $4.33B

→ $S_D = -I_D$ -1.64 B | $S_M = -I_M$ + 4.33B

→ $-S_D = I_D$ + 1.64B |

$-S_D$ | = S_M

$(I_D + 1.64) = (-I_M + 4.33)$

And since $I_M = I_D$ $I_D + I_M = 2\ I_D = 2.69$

I_D = \$1.345B

Because we know Income and total effects, we can deduce the substitution effects from this information. Our estimates are consistent with the assumption of normality for consumer imports:

Income Effect:	\$ +1.345B I_D	\$ +1.345BI_M	\$ +1.345B I_D \$ +1.345B I_M
Substitution Effect	\$ - ? . S_D	\$ + ? .S_M	\$ -2.985B . S_D \$ +2.985B . S_M
Total Effect	\$ = -1.64B T_D	\$ =+4.33BT_M	\$ = -1.64B T_D \$ =+4.33B T_M

Conclude: The results support the correctness of original hypothesis that unlike investment goods, imported consumer goods, as a group, are normal goods.

4.2.2. CASE C.2: M ARE INFERIOR CONSUMER GOODS; NOMINAL EXCHANGE RATE USED TO ESTIMATE TOTAL EFFECTS

However, If we had assumed that imported consumer goods are an inferior good, i.e., the substitution effect is negative for imports as they become cheaper, and positive for domestic goods, we can see that with our estimated income effect of \$1.345B for both types of goods, no negative number for the import substitution effect could be added to the income effect(+1.345B) to get our estimated total effect for imports (+4.33B). Similarly, no positive valued substitution effect into domestic consumption could be added to the 1.345B income effect to get our estimated total effect of (-1.64B). We therefore reject the hypothesis that consumer imports are inferior goods.

4.2.3. CASE C.3: IMPORTS ARE NORMAL CONSUMER GOODS; REAL EXCHANGE RATE USED TO ESTIMATE TOTAL EFFECTS

Using the methods described earlier, we know Income and total effects, the substitution effects we deduce from this information our results are consistent with the assumption of normality for consumer imports:

Income Effect:	\$ +1.415B I_D	\$ +1.415B I_M	\$ +1.415B I_D \$ +1.415B I_M
Substitution Effect	\$ - ? . S_D	\$ + ? . S_M	\$ -1.615B . S_D \$ +1.615B . S_M
Total Effect	\$ = - .20B T_D	\$ =+3.03B T_M	\$ = - .20B T_D \$ =+3.03B T_M

Conclude: The original hypothesis that D, M are normal goods appears correct, assuming our income and total effect estimates are correct.

4.2.4. CASE C.4: IMPORTS ARE INFERIOR CONSUMER GOODS; REAL EXCHANGE RATE USED TO ESTIMATE TOTAL EFFECTS

However, If we had assumed that imports are an inferior good, i.e., the substitution effect is negative for imports as they become cheaper and positive for domestic goods, we can see that with our estimated income effect of $+1.415B for both types of goods, no negative number for the import substitution effect could be added to the income effect to get our estimated total effect for imports ($3.03B). Similarly, no positive valued substitution effect for domestic consumption could be added to the $+1.415B income effect to get our estimated total effect of $ - .20B. We therefore reject the hypothesis that consumer goods imports are an inferior good.

REFERENCES

Baker, M., and J. Gruber, and K. Mulligan. 2003. “The retirement Incentive Effects Of Canada’s Income Security Programs”. *The Canadian Journal of Economics*. Vol. 36 (2).

Elmendorf, Douglas. 1996. “The Effect of Interest Rate Changes on Household Saving and Consumption: A Survey”. Finance and Economics Discussion Series 96-44. Board of Governors of the Federal Reserve System. June.

Heim, John J. 2008. “How Falling (Nominal) Exchange Rates Have Affected The U. S. Economy And Trade Deficit”. *Journal of International Business and Economics*. Vol. VIII, (1).

Heim, John J. 2009. “How Falling (Real) Exchange Rates Have Affected The U. S. Economy And Trade Deficit”. *Proceedings of the 61st Annual Conference of the New York State Economics Association, Vol. 1: 44-53.*

Kimball, Miles, and Matthew Shapiro. 2008. “Labor Supply: Are the Income and Substitution Effects Both Large Or Both Small?” Working Paper 14208, National Bureau of Economic Research. July.

Prager, J. 1993. *Applied Microeconomics.* Boston: R. J. Irwin, Inc.

Wold, H., and L. Jureen. 1953. *Demand Analysis.* Wiley Publications in Statistics. New York: John Wiley & Sons.

Biophysical Economics and the Failed Growth Economy

Kent A. Klitgaard*

ABSTRACT

The traditions of both mainstream and ecological economics presuppose the inherent nature of a market economy is to grow. However the traditions of classical and modern political economy posit an economy that is subject to limits to growth and one that tends towards secular stagnation. This paper explores the stagnation thesis in the contexts of the current financial crisis, political impasse and the limits posed by the potential conjuncture in time of peak oil and climate change.

INTRODUCTION

The standard model of neoclassical economics posits a self-contained and self-regulating economy. This approach is embodied in the familiar circular flow of materials, factors and money. In contrast, ecological economics embraces a pre-analytical vision of an open and growing economy embedded within a finite and non-growing biophysical system. The economic subsystem must obey the laws of the primary system as regards the flow of energy, and its growth is necessarily limited by the primary system. Moreover, the resilience of natural systems depends upon certain redundancies and biodiversity to function properly in their own rights. This further limits human ability to appropriate the whole of net primary productivity for economic uses.

A great deal of excellent work has been accomplished by economists and ecologists in the past fifteen years estimating how big the economy may grow relative to the limits of the supporting biophysical systems. In the 1980s and 1990s a number of studies focused upon the carrying capacity of the planet (Vitousek, *et al.*1986, Daily and Ehrlich, 1992, Postel, 1984, Rees and Wackernagel, 1992), while others sought to improve the accuracy of the income and product accounts as regards assessing the impact of humanity upon the finite and non-growing biophysical system (Daly and Cobb 1989, Cobb and Cobb 1994, Costanza, *et al.* 1997).

Unfortunately less work has been done assessing the nature of the economy, its inner dynamics, and how it operates as a system. The idea of the growing economy is so taken for granted that few ecological economic analyses have been devoted to understanding formally the patterns of growth, the or structural and institutional changes over time. The uncritical acceptance of growth may actually hinder a broader and more comprehensive understanding of mature market economy as a system. We are witnessing just the latest in a long series of such economic declines. I believe it is crucial to understand the dynamics and limits to the economic system if we are to more fully understand the interaction between the economic and biophysical systems.

* Department of Economics, Wells College, Aurora, NY 13026

The study of political economy is comprised of just this focus on the dynamics of accumulation, distribution and institutional change. The purpose of this paper is to begin the process of integrating the principles and insights of political economy and institutional economics into the corpus of biophysical and ecological economics.

THE NATURE OF LIMITS

I assert that the growth of a mature industrially-based economy located within a democratic political system is constrained by three sets of limits. To begin with there is a set of *biophysical limits* imposed by the laws of the finite and non-growing primary system in the age of hydrocarbons. These limits take the form of a declining resource base, for example peak oil, and the accumulation of heat trapping gases in the atmosphere. In addition the economy faces a set of *internal limits* posed by the inability to find sufficient spending outlets for the growing economic surplus in the real economy. This is most commonly recognized by Keynesian economists as the problem of inadequate effective demand. Cost reductions accomplished by means of technological change, increased productivity and efficiency must find comparable outlets in rising aggregate demand if they are to result in profitability rather than chronic unemployment and excess capacity. Finally the final impediment to fundamental change is a set of *political limits.* Bills are fashioned to pass the legislative bodies rather than to solve the problems at hand, and arcane rules perpetuate the *status quo.* When long-term strategies for growth succumb to internal or biophysical limits they do not immediately give way to superior strategies. Instead the nation often becomes mired in a period of political impasse. No one perspective is sufficiently strong to impose its vision, yet at the same time is strong enough to keep another vision from being implemented. The twentieth century was dominated by periods of impasse. The accepted vision of the self-regulating economy was dashed by the severity of the great depression, and after nearly half a decade a new vision emerged in the form of the New Deal in the United States. Keynes' theoretical insights would not come until the seventh year of the depression, and were never fully implemented.

This vision of a managed economy and a positive state survived until the limits of effectiveness were reached in the early 1970s. The liberal growth agenda, driven by increasing household incomes, could not survive the peak of US oil production, the decline of US political power internationally, and the onset of stagflation. A long period of impasses followed, only broken in the early 1980s with the election of Ronald Reagan to the presidency and the implementation of a conservative growth agenda grounded in the reduction of production costs and the expansion of unregulated financial markets. We are currently witnessing another such period of political impasse. Our current economic and political systems depend on growth; however there is simply no way one can grow their way into a smaller impact on the planet. Attempts to decarbonize the economy in the face of globally peaking oil will reorder fundamentally the power relations of American industry and finance, as well as create myriad macroeconomic dislocations as well. As much as our economic and political structures depend upon growth, the scientific evidence of limits to growth is overwhelming.

The result of the historical conjuncture of these limits results in *The Failed Growth Economy*: an economy that *must* produce growth in order to provide for profits and employment yet at the same time simply *cannot* produce this requisite growth. Since the 1970s in the United States the failed growth economy creates a fundamental dilemma. **Growth rates of the real economy have not been sufficiently high to provide for full employment, while at the same time aggregate growth has been too high to provide for sustainability within finite and non-growing biophysical limits.** We grow too much and too little at the same time.

Historically these periods of economic crisis and political impasse were transcended by transforming nature to meet human needs. In other words society could push back the biophysical limits in order to transcend the political and economic limits. We are most probably entering a historically unprecedented era, one for which no models exist. As we approach the era of globally peaking oil and the limits posed by atmospheric carbon dioxide we may no longer have the ability to ignore the consequences of reaching biophysical limits. Instead we must lean to live within these limits: to draw our energy from renewable solar flow rather than limited terrestrial stocks subject to depletion. We must maximize human welfare while simultaneously reducing the flow of matter and energy through a finite and non-growing system. In short we must cease growing. But how to accomplish that, without perpetuating the worst aspects of the failed growth economy: unemployment and lack of opportunity, is a difficult endeavor. It must begin with an understanding of the dynamics of the growth economy itself.

GROWTH AND ACCUMULATION IN CLASSICAL POLITICAL ECONOMY

Classical Political economists from Adam Smith to John Stuart Mill shared common threads. They all subscribed, in one form or another, to cost of production theories of value. Moreover, all of them focused on the process of accumulation, and all theories accounted for the end of accumulation and the arrival of the *stationary state*. For Smith the problem began with the division of labor. In rude and early societies a person's own labor could supply the bulk of his or her needs. But the introduction of the division of labor entails producing to fulfill the needs of others, which necessitates an accumulation of capital. The need for accumulation expands with the division of labor. An expanding market is necessary to absorb the output created by increased productivity. Virtuous individuals refrain from current consumption and use the savings to accumulate stock, thereby setting productive labor into action. By doing so, the private pursuit of profit is translated into public virtue. However Smith's vision also entails the end of accumulation. The increase in stock inevitably leads to a decrease in profits. The resulting decrease in profits leads to a secular reduction in investment and the onset of the stationary state.

The debates of years following the death of Smith (in 1790) earned economics the name "the dismal science." Thomas Robert Malthus, starting from the position that a fraction of the poorer classes will inevitably starve, warned that increasing the incomes of the poor would simply increase the demand for food and thereby drive up its price. Throughout the subsequent six editions of the *Essay on the Principles*

of Population Malthus extended this vision to argue that the prospects for growth would be limited by the growth of human population, unless somehow "checked" by measures that would either lower the birthrate, or alternate measures that would increase the death rate. What Malthus also took as given was the existing set of restrictions on imported food (the "Corn Laws") that ensured both higher food prices and a steady source of income in the form of rent for the landed gentry and aristocracy. These were restrictions he favored, as they increased rents relative to profits. Not only was Malthus a conservative spokesperson for landed interests, his theory of accumulation was grounded in the politics and economics of effective demand and a critique of Say's Law of Markets. Income that was not claimed as rent would accrue to capitalists who would save an inordinate share. These hordes of savings would reduce the overall demand and create the potential for a stagnant economy. Malthus' solution was the maintenance of the "Corn Laws" and the expenditure of the subsequent rents upon unproductive labor that did not add to the productive potential of the economy.

Opposing Malthus was David Ricardo, who argued for a repeal of the Corn Laws. Here Ricardo argued that since the most fertile land was used first an extension of domestic cultivation into parcels of inferior quality would produce both decreasing output per unit of input and increasing rents. Both rising costs and rising rents would squeeze profits. Eventually, as the margin of cultivation was extended, all of the social net product would be divided between wages and rents. But before the onset of this logical conclusion when profits fell below the rate of interest (Ricardo estimated 6-7%), capitalists would channel their funds into speculative pursuits. Consequently accumulation in the real economy would grind to a halt.

Ricardo argued that the elimination of the Corn Laws could postpone the onset of the stationary state. If cheap imported grains replaced those grown by the extension of British cultivation the onset of diminishing returns could be postponed. The resulting increase in profits would be reinvested expanding both the number of productive workers and their productivity. From Ricardo modern economics takes not only the principle of diminishing marginal returns but also the faith that these diminishing returns can be transcended by free trade and technological change.

The most unique approach to the stationary state among the Classical Political Economists was that of John Stuart Mill. Mill believed the onset of the stationary state was imminent, postponed only by government borrowing, capital export and waste. Furthermore, the end of the era of growth could be accompanied by a greater focus on improving the human condition by means of the combination of income redistribution and zero population growth. Once population growth among laborers ceased the pressure on the growth of capital formation to provide employment would be lessened and a superior society could ensue.

NEOCLASSICAL AND KEYNESIAN ECONOMICS

Notions of the stationary state came to a halt with the Jevonian Revolution and the rise of subjective utility theory. Not only were the determinates of price dematerialized and attributed to the final degree of

subjective utility, but the calculus of constrained optimization was adapted to quantify this subjective approach as science. The focus on accumulation was supplanted by static and comparative static determinants of price formation, and the tendency for the system to produce a stable equilibrium. All social policies and internal contradictions disappeared into the formalization of the self-regulating circular flow of material and money and the reassertion and mathematical formalization of Say's Law (Mirowski 1989). Adjustments in prices alone could equilibrate quantities in all markets, and efficient allocation by means of prices became the holy grail of economics--all in a world of certainty.

It was these propositions that Keynes opposed in his *General Theory of Employment, Interest and Money.* Keynes left the fundamentals of subjective theories of value and distribution in tact. His critiques focused on macroeconomic aggregates rather than price formation in individual markets. He focused one critique upon the labor market where he was primarily interested in refuting the idea that movements in relative prices would suffice to equilibrate output at a full employment level. Instead Keynes argued that wages comprised the major component of effective demand, and that a reduction in wages would reduce the capacity to consume the output.

It was in the area of savings and investment that Keynes made his biggest mark, arguing that changes in the relevant price, the interest rate, would not serve to equilibrate the market for loanable funds. Instead Keynes posited that savings depended upon the level of income. Investment was more complicated, as it depended not only upon the cost of acquiring funds, but also upon profit expectations. Profit expectations themselves depended upon animal spirits as well as income and the cost of capital. But these spirits themselves were fraught with uncertainty. It was simply impossible to predict the performance of finance markets over the lifetime of an investment. Moreover, liquidity preference became highly elastic at low interest rates, and a liquidity trap could reduce the effectiveness of monetary policy. Keynes put most of his faith and effort into governmental stimulation of investment. Investment was linked to the other component of final demand, namely consumption, through the marginal propensities to save and consume and by the multiplier. Keynes' theory of the causes of, and solutions to, the depression were squarely in the lineage of other underconsumption theories. This held out the possibility of effective demand would be chronic rather than acute, raising the need for long-term full-employment policy not simply short-term "pump priming." Theoretically, Keynes raised the possibility of secular stagnation, although he did not develop the details. Neither did he render his system dynamic. That task was left to his followers (Dobb 1973).

THE STAGNATION THESIS AND THE EMERGENCE OF MODERN GROWTH THEORY

An emphasis on economic growth emerged only in the latter stages of the New Deal. The early new Deal was about rescuing a failed growth economy and a collapsing financial sector within the confines of Franklin Roosevelt's fiscal orthodoxy. Spending programs were matched by a tax increase. FDR's commitment to a balanced budget did not free sufficient funds to "spend one's way out of the depression."

It was only during the Second World War, the largest public works program the United States has ever seen, that the needs of the emergency caused political leaders to put aside their fears of budget deficits. But during the second depression of the great depression (1937-39) the American Keynesian Alvin Hansen re-enunciated a theory of secular stagnation. Hansen argued that adequate investment outlays are needed to sustain full employment and adequate income levels. However these investment outlets are difficult to find in a mature economy. Mature economies exhibit slower rates of population growth, and population was the driving force behind investments in residential building, territorial expansion, and the expansion of railroads and public utilities. The 19th century gave birth to the industrial revolution with the use of fossil fuels to power machinery and the corresponding increase in investments in plant and equipment. Then came a period of "readjustment and relative stagnation" (Hansen 1938: 314). This period of stagnation was followed by the great expenditures of the railroad age, which themselves reached saturation. This was followed by the ages of electrification and the automobile which provided a whole host of opportunities for investment in related industries as well.

Hansen predicted that new innovations to come would use far less capital per unit of output After adjusting for new products and processes, as well as technological change, Hansen concluded that new opportunities sufficient to fully utilize the flow of savings would be difficult to find in the mature economy. Hence the dilemma was one of secular stagnation and the stability of the system, rather than a short-term imbalance of supply and demand. He advocated that the government become essentially an investment banker rather than simply a lender of last resort or a short-term pump primer.

Hansen's argument was supplemented in the late 1940s by the work of Evesy Domar in the "golden age" of growth theory. In 1947, fresh on the passage of the Employment Act of 1946 by congress (mandating the three goals of reasonably full employment, stable prices, and economic growth) In "Expansion and Employment" he posited that the percentage of the labor force employed depends upon the ratio of national income and productive capacity. His key insight was to point out what he called "the dual nature of investment." Investment creates income and employment, while simultaneously creating productive capacity that must be utilized to maintain employment. After formalizing the role of investment in the creation of national income, Domar turned to the determinants of investment.

(1) $Y = \Delta I \frac{1}{\alpha}$ where $1/\alpha$ is the multiplier, and I is Investment. Furthermore:

(2) $\Delta I \frac{1}{\alpha} = I\sigma$ where σ represents the potential social average productivity of investment as a measure of the change in productive capacity.

The left hand side of the equation stands for the annual incremental change in income (the demand side or the multiplier side) while the right hand side represents the annual increase in productive capacity (the supply or productive capacity side.) The dual nature of investment shows as investment appears on both sides of the equation. In other words investment generates an increase in income through a

multiplier effect but also generates an increase in productive capacity that must be employed. If income and employment are to rise a growing level of expenditures must be targeted towards investment. **The increase in income is not a function of the level of investment but of its rate of change.** This is the essence of the failed growth economy. Investment cannot grow forever at an increasing rate because of its role in augmenting productive capacity. At some point excess capacity appears and new investment is no longer profitable (Domar 1947).

The degree to which excess capacity affects investment depends upon the structure of industry. A competitive economy that is changing rapidly in terms of technology shows little positive correlation between capacity utilization and new investment. However unused capacity presents a serious threat to new investment in industries that are monopolized. Insufficient investment today creates unemployment today. If more is invested today even a greater amount will be needed tomorrow to maintain full employment. It is likely that the increase in income will more than compensate for the increase in capacity. However the increase in income is temporary while the increase in capacity is longer lived. If the dreaded stationary state is to be avoided humanity must arrive at a greater understanding of the relation between technical progress and spending patterns.

THE STAGNATION THESIS SYNTHESIZED

In 1966 Paul Baran and Paul Sweezy published *Monopoly Capital* in 1966 and enunciated the most complete modern statement of what has come to be called “the stagnation thesis.” They began their essay with the proposition that the economy had become dominated by the giant corporation. Technically Baran and Sweezy analyze an economy dominated by oligopolies

They then linked the rise of oligopoly with the decline of price competition. Oligopolies compete corespectively, to use Schumpeter’s term, by means of cutting costs and increasing market share, not by means of reducing prices. The new method of competition widened the difference between the value of the output and the cost of producing it, which Baran and Sweezy defined as economic surplus. They hypothesized that the fundamental economic problem in the era of monopoly capitalism is the inability to absorb, or find adequate spending outlets for, the rising economic surplus. This rising surplus leaves its statistical trace as excess capacity and unemployment. Baran and Sweezy then go on to explain the internal mechanisms by which surplus can be absorbed. Surplus can be invested, but here they utilize Domar’s theory of the dual nature of investment, and the greater impact of excess capacity upon new investment in a monopolized economy. In the end private investment alone proves inadequate to the task of fully absorbing surplus. Surplus can be consumed, but even with the development of the sales effort in the 1950s-from planned obsolescence to the explosion of advertising consumption cannot adequately absorb a rising surplus either. The ability of civilian government to consume adequately the surplus is constrained by the political process. The final option is to simply waste the surplus. Baran and Sweezy focused mainly on the military but in today’s climate one must also consider our fossil fuel-based

transportation system, our approach to health care and it's resulting enormous production of medical supplies and waste, and our crumbling yet expensive system of public schools. In the end the internal mechanisms are incapable of sustaining the absorption of economic surplus. In their sense of the term stagnation does not mean lack of production. Rather the macroeconomy generates actual income levels that are beneath potential income levels. This results in chronic excess capacity and unemployment. Consequently: **The natural tendency of a mature, monopoly capitalist economy is towards stagnation, not growth.**

The theoretical task then becomes one of explaining occasional periods of economic growth, not periodic downturns. This they do by relying on Hansen's and Joseph Schumpeter's notions of the *epoch-making innovation.* Such innovations not only absorb tremendous amounts of investment capital themselves but fundamentally transform the structure of industry by creating myriad ancillary industries to absorb more investment capital themselves. They utilize the three innovations cited in Hansen (1938): the steam engine; the railroad; and the automobile. From a biophysical point of view all these innovations transformed the structure of industry and society by requiring, and stimulating the search for, cheap fossil energy. The automobile helped create the demand for suburban housing, drive-ins, repair shops, the interstate highway system, and influenced the popular music, not to mention the sexual freedom, of the 1960s absorbing sufficient economic surplus to provide for a prolonged period of prosperity. The rise of the automobile utterly transformed the oil industry. The second factor that absorbed sufficient surplus was *war and its aftermath.* The Second World War was the largest public works program in American history. Unemployment, which stood at 17.9% in 1939, fell to 1% by 1944. The aftermath of the war led to the Bretton Woods Accords, the Marshall Plan and a long period of U.S. hegemony. The end of this hegemony, along with the collapse of the Bretton Woods Accords in the 1970 signaled a period of economic transformation.

THE RISE OF THE FINANCIAL ECONOMY

Throughout the 1980s Paul Sweezy and Harry Magdoff began to raise consistently questions of finance relative to production in the real economy. Moreover they treat the rise to prominence of the financial economy as a symptom of the overall secular stagnation of the real economy. Instead of being a parasitical usurper of funds that would naturally lead to productivity increases when invested in productive capacity, Magdoff and Sweezy assert that the rise of the financial services sector has been the primary means by which the stagnation of the real economy has been kept at bay. In addition they assert that the financial system has not simply grown, but has fundamentally transformed the economy. In traditional theory banks play the role of a benign vehicle by which the deposits of households and firms are transferred to investors by means of loans. Banking was reasonably safe, reasonably profitable, and downright boring. In the transformed state, termed "Monopoly Finance Capitalism" the expanded, and increasingly independent, financial services sector has come to claim the dominant share of the systems

profits, increase the offering of increasingly complex securities, and has become the nation's largest debt-holding sector. The financial sector exists to create ever more complex derivative securities in order to seek out potential opportunities for growth and profit, as well as reducing systemic risk. However, the rise of the financial economy is subject to its own limits, and in the end cannot overcome the inherent stagnationist tendency of mature capitalism. Limits are, once again, those internal to the investment process, biophysical, and political.

As the real economy has stagnated, financial activity and the expansion of debt have come to represent a growing share of total profits and a rising percent of the components of income. Moreover capacity utilization has fallen. It will help to recall that profit-making strategies in the monopoly capitalist era include cost reduction and the expansion of market share. But cost reduction has entailed the reduction in wage growth, as well as the reduction of materials cost. But how does one sell more output to a population with declining incomes? The answer is to be found in the expansion of debt. But at some point the limited income will also limit the expansion of debt, as will the rising inequality in the distribution of those incomes. This was manifest, over the course of the last year, especially in the housing sector. What productivity increases as have occurred in manufacturing have largely been the result of utilizing increasing quantities of cheap oil. (Cleveland, 1984)The advent of peak oil will raise these input prices with two potentially adverse results. On the supply side rising oil prices (and consequently rising prices for food, chemicals, transportation, etc) will eventually lead to the reemergence of cost-push inflation. It is very difficult to fight cost-push inflation by means of contractionary monetary policy, as interest charges are themselves components of costs. To the degree that oligopolies price by means of marking up prime costs, tight money can actually feed cost push inflation rather than remediate it (Wachtel and Adelsheim 1977). Furthermore, oil is denominated in dollars and any structural weakening of the U.S. economy, say by means of cost-push inflation, can further exacerbate the increase in oil prices.

However, the increase in oil prices can have demand-side effects as well. As oil prices increase the petrodollar stocks found in sovereign wealth funds will need to find outlets. If the country follows an accommodating monetary policy, these funds will look for higher rates of returns than can be found in Treasury Bonds. Either the giant global pool of money insatiably seeks out riskier but higher returns, for example mortgage backed securities and collateralized debt obligations or moves the money to financial systems promising higher rates of return on their securities (e.g. Iceland with its 11.5% rate of return). When these funds began to exit Iceland during the financial panic the country itself was near bankruptcy. So the coming of peak oil may have the same type of dual effect that Domar posed as a problem as regards investment. If a country adapts an easy money policy the risk of system-wide financial collapse increases as does the marketing of increasing quantities of lower quality securities spreads. If the nation runs a contractionary monetary policy the risk of stagflation becomes more manifest.

Climate change presents an entirely different set of biophysical limits. In absolute terms real per-capital GDP has increased, nearly exponentially since the middle of the 20th century. So too, has consumption, with the general exponential trend increasing despite the recurring recessions of this time

period. It is this increase in consumption, bigger cars, more gasoline, and more energy-intensive appliances that drives the increase in carbon emissions and carbon dioxide concentrations in the atmosphere. James Hansen of NASA Goddard Space Center, perhaps the nation's most prominent scientific voice concerning climate change estimates that the theoretical "tipping point" at which irreversible damage will occur at concentrations of 350 parts per million. The "Business as Usual" strategy may produce atmospheric CO_2 concentrations in the range of 1200-1400 ppm by 2100. The Intergovernmental Panel on Climate Change uses these data to predict a sea-level increase in the range of 18-59 centimeters, with a very likely of increase heat waves and heaving rains. In addition they predict the likely increase of droughts, tropical cyclones, and extreme high tides. If we already exceed the tipping point by nearly 30 parts per million and data-based projections indicate possible CO_2 concentrations that are four times the tipping point by the end of the century the indication is that we **simply cannot grow our way into environmental sustainability.** Unless you simply discount the validity of the scientific evidence, or believe we can instantaneously convert to a carbon free source of energy without fundamentally disrupting financial markets then climate change appears as an external limit to the potential for economic growth. Moreover, the advent of peak oil may likely occur in the same time frame (from now until mid-century) as do the impacts of climate change. In essence the next generation will have both problems to deal with at the same time with, as we shall see, little discretionary income with which to deal.

REFERENCES

Baran, Paul and Paul Sweezy 1966. *Monopoly Capital.* New York: Monthly Review Press.

Bowles, Samuel, David Gordon, and Thomas Weisskopf. 1990. *After the Wasteland.* Armonk, New York: M.E. Sharpe.

Cleveland, Cutler, Costanza, Robert, Hall, C.A.S., and Kaufmann, Robert. 1984. "Energy and the U.S. Economy: A Biophysical Approach." *Science* 211: 576-579.

Cobb, Clifford, and John Cobb. 1994 *The Green National Product.* Landham, MD: University Press of America.

Costanza, Robert, Ralph d'Arge, Rudolph deGroot, Stephen Farber, Monica Grasso, Bruce Hannon, Karin Limburg, Shahid Naeem, Robert O'Neill, Jose Paruelo, Robert Raskin, Paul Sutton and Marjan van den Belt. 1997. "The Value of Ecosystem Services and Natural Capital." *Nature* 387: 253-260.

Daily, Gretchen C.and Paul R. Ehrlich. 1992. "Population, Sustainability, and the Earth's Carrying Capacity." *BioScience* 42(10): 761-771.

Daly, Herman, and John Cobb. 1989. *For the Common Good.* Boston: Beacon Press.

Dobb, Maurice. 1973. *Theories of Value and Distribution Since Adam Smith.* Cambridge: Cambridge University Press

Domar, Evesy. 1947. "Expansion and Employment." *American Economic Review.* 37(1): 34-55.

Foster, John B, and Magdoff, Fred. 2009. *The Great Financial Crisis.* New York: Monthly Review Press.

Hansen, Alvin. 1938. *Full Recovery or Stagnation?* New York: W.W. Norton and Company.

Keynes, John Maynard. 1964. *The General Theory of Employment, Interest and Money.* New York: Harcourt, Brace and World.

Mirowski, Phillip. 1989. *More Heat than Light.* Cambridge: Cambridge University Press.

Postel, Sandra. 1994. "Carrying Capacity: The Earth's Bottom line. In *State of the World,* Lester Brown, Ed., 3-21. New York. W.W. Norton and Company.

Rees, William and Wackernagel, Mathis. 1992. Ecological Footprint. In *Investing in Natural Capital.* Ann Marie Jansson, Ed. St. Lucie, Florida: Island Press.

Vitousek, Peter, Paul R. Ehrlich, Anne H. Ehrlich, and Pamela A. Matson. 1986. "Human Appropriation of the Products of Photsynthesis." *BioScience.* 36(6): 368-373.

Wachtel, Howard, and Adelsheim, Peter. 1977."How Recession Feed Inflation: Price Mark-ups in a Concentrated Economy." *Challenge* 20 (4): 6-13.

Understanding Labor Flows in New York State Using Local Employment Dynamics Data

Arindam Mandal*

ABSTRACT

The purpose of this paper is to understand the labor market flows in the New York State using the Local Employment Dynamics (LED) data from the United States Census bureau. The LED data provides information regarding hires, separations, employment, job flows and wages at county and industry level for the United States. The study finds that in the period between 2001 and 2008 in the New York State, there have been substantial changes in the pattern of employment, labor turnover and wages based upon gender. Though female employability improved compared to male, but the gap between male-female wage widened during the period.

INTRODUCTION

The purpose of this paper is to understand the labor market flows in the New York State using the Local Employment Dynamics (LED) data from the United States Census bureau. Modern economies are characterized by extremely dynamic labor markets. There are large flows of jobs and workers between the states of activity and inactivity every month. For example, in the New York State, during the period between 2001 and 2008, in a typical month on average 0.49 million people are unemployed and actively looking for jobs. In the same time period, about 0.531 million vacancies are posted every quarter and about 1.2 million workers hired per quarter. To complement these labor market flows, average 1.6 million workers have been separated every quarter. All these flows are significant, when compare with the average 9.4 million labor force in the New York State. Though, the unemployment behavior has caught the interest of the economists for a long time, understanding the important labor market dynamics related to hiring, separations and vacancies are fairly new in the profession.

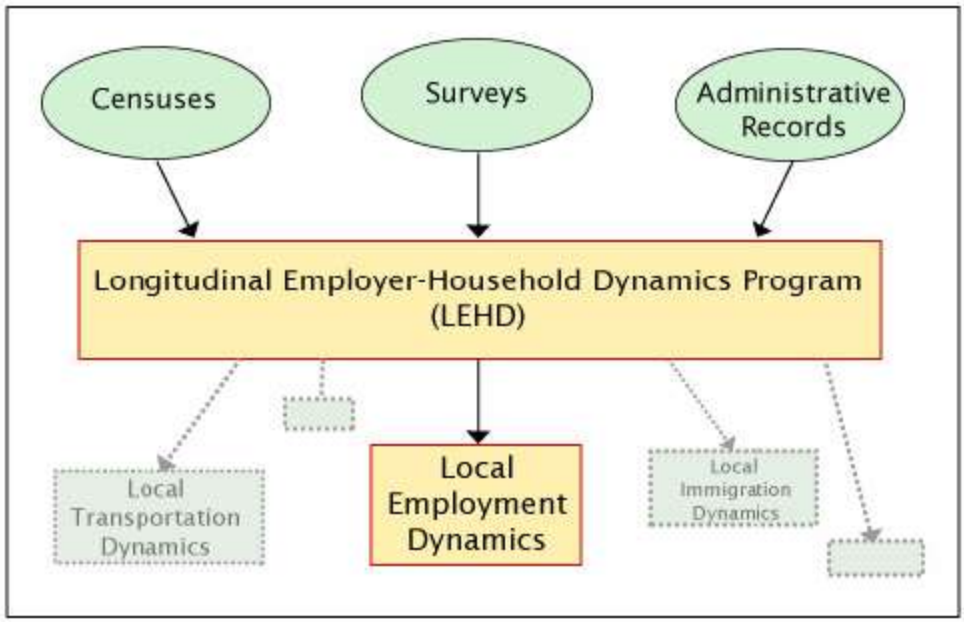
The paper details the characteristics of the LED data and provides descriptive evidence at the aggregate levels. Aggregate level relations between vacancies, labor turnover and unemployment are also analyzed. It also characterizes the data scope, measurement and the research potential these data have.

Existing research using LED is almost non-existent. Abowd et.al (2005) summarizes the technical details associated with LEHD and the creation of LED data set from the surveys.

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

DATA

Local Employment Dynamics (LED) is a part of the bigger program by United States Census Bureau called Local Employer-Household Dynamics (LEHD). LEHD use modern statistical and computing techniques to combine federal and state administrative data on employers and employees with core Census Bureau censuses and surveys. On the other hand, LED is a voluntary partnership between state labor market information agencies and the United States Census Bureau to develop new information about local labor market conditions at low cost, with no added respondent burden. The difference between LEHD and LED can be explained by the following diagram



Source: Quarterly Workforce Indicators, United States Census bureau

Quarterly Workforce Indicators (QWI) provide detailed local estimates of variety of employment and earnings indicators based upon information from LED. Employment, earnings, gross job creation and destruction, and worker turnover is available at different levels of geography, typically down to the county or metro area. At each level of geography, they are available by detailed industry (SIC and NAICS), sex, and age of workers. Currently, QWI is available for all the states of the United States, except for Connecticut, Massachusetts, New Hampshire, District of Columbia, Puerto Rico and Virgin Islands.

The Quarterly Workforce Indicators (QWI) are derived from state administrative records and basic demographic information from the Census Bureau. Employment totals from the QWI are not exactly comparable with those from other sources. Generally, coverage and definitions differ between the QWI and data about establishments from administrative records (e.g., the Quarterly Census of Employment and Wages or QCEW), and about workers from surveys (e.g., the decennial census, the American Community Survey, and the Current Population Survey or CPS.). Detailed information is available in a paper Stevens (2007).

In the paper, state level data for New York used for the sample period ranging from 2001 quarter 1 to 2008 quarter 3.

CONCEPTS AND DEFINITIONS

The definition of the measures in QWI is as follows: -

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

VACANCIES AND BEVERIDGE CURVE

The publicly available QWI estimates present a wealth of new evidence for the aggregate labor market. While the time series is short, it spans couple of recessions and slow labor market recovery, allowing researchers a glimpse of the cyclical behavior of vacancies and labor turnover. The National Bureau of Economic Research (NBER) dates the recessions during this period as starting in March and ending in November of 2001 and the second one starting in December 2007 and still ongoing. After the 2001 recession, losses in total employment continue through 2003 Q3 (Figure 1). Figure 5 illustrates the aggregate behavior of vacancies and unemployment between 2001 Q1 and 2008 Q3. The unemployment rate estimates come from the Local Area Unemployment Statistics (LAUS) and the vacancies based on the total job creation during a quarter. Though the unemployment rate shows a cyclical pattern, but vacancy rate shows a downward trend throughout the sample period. In 2001, because of recessions unemployment rises while vacancies fall. Both unemployment rate and the vacancy rate hover around 6 percent during this period. Beginning 2004 Q1, the unemployment rate begins to fall while the vacancy rate trended downward; these patterns continue until the end of 2006. From 2007 Q1, unemployment started going up, whereas vacancies show a steep decline from 2007 Q3.

An important relation in the theory of worker search and matching is the Beveridge Curve, which relates the cyclical movements of vacancies to those of unemployment. Figure 6 plots the aggregate Beveridge Curve with the vacancy rate on the vertical axis and the unemployment rate on the horizontal

axis. Search theory predicts a negative relationship between vacancies and unemployment (Pissarides, 2000). The expected negative relationship is not observed in the Beveridge curve for New York.

TRENDS IN TOTAL EMPLOYMENT

In LED, total employment is defined as the total number of workers who were employed by the same employer in both the current and previous quarter. By definition total employment reported by QWI measure is indicative of jobs, which are more stable in nature. In the period under study, in New York, total employment declined steadily from 2001 Q1 to 2003 Q3 and then started rising. It is indicative of the fact that employment takes time to recover after a recession (Figure 1). Since the beginning of 2001 recession, total employment in New York declines 4.6 percent by 2003 Q3. The recovery in total employment is slow and could only reach the pre 2001 recession only by the first quarter of 2008. But the first quarter of 2008 is marked by the start of another recession and the trends show a decline in total employment since the beginning of 2008 recession. It is interesting to note that, total employment rate defined as the total employment as proportion of the labor force declined steadily throughout the decade under consideration (Figure 2).

The QWI provides gender specific data. Figure 7 show trend in male-female total employment and it varied substantially over the period under study. In 2001 Q1, total employment among male and female were 4.31 million and 4.18 million respectively. Till 2003 Q3, both the male-female total employment shows declining trend but the rate of decline among males surpassed the decline rate among females. Though total employment started recovering since 2003 Q3, but the rate of recovery is faster among females than among males and by 2004 Q4, females surpassed males. By 2008 Q1, though female total employment reached the level in 2001 Q1 but the male numbers still remain well below the initial levels.

LED data can be used to measure the year-over-year annual net job gain or loss[1]. In turn from net job gains we can calculate net job gain rate[2]**.** Net job gain rate is shown in Figure 3 and 4. In Figure 3, during recession the net job growth declined, but it became positive only in the 2004 Q3. This indicates that job recoveries since recessions are slow. Figure 4 show the net job gain rate among male and female. Job gain rate among females are faster than males and also net job loss among females are less than for males.

LABOR TURNOVER

The LED data tracks labor turnover. The Figure 8 plots the time series of aggregate hires and separation rates over the sample period. Both hires and separations trended downward throughout the sample period. During recessions, both hires and recessions declined at a faster rate compare to other periods. After the 2001 recession, hires remained steady around 1.2 million jobs per quarter, but it never reached the pre 2001 levels of 1.4 million hires per quarter. Though separations declined throughout the period, but the faster rate of decline in separations during recessions is counter intuitive. Recessions must be marked by more separations because of increased layoffs, but the trend in data shows other way round. One possible reason for this trend is labor hoardings during recessions. During recessions workers

tend to hold on to a job and at the same time, often employers tend to hoard workers with the expectation that the market will improve in the near future.

The aggregate time series statistics of labor turnover are presented in Table 1. Separation rate is 17.4 % per quarter, which is higher than the hire rate of 13.5 %. Both hires and separations are positively correlated with unemployment. Vacancies, hires and separations are all highly positively correlated. Also, all the three variables show considerable persistence as shown by the autocorrelation.

The gender specific trends in labor turnover are shown in Figure 9. For both hires and separations, males surpassed females throughout the period. Male-female gap between hires and separations reduced over the sample period.

WAGE TRENDS

In the period under study, both new hire earnings and average monthly earnings trended upwards. But the rate of increase of average monthly earnings is around 1% per quarter, which is higher than the rate of increase of average new hire earnings of 0.7% per quarter. During recessions, both the earnings of new hires and the average earnings showed a downward trend as expected.

Figure 10 and Figure 11 show the average monthly earnings and average monthly earnings of new hires for male and female employees respectively. Female earnings remained well below the male earnings throughout the period. Over the decade, the gap between average monthly earnings for male and female widened from $1,912 in 2001 Q1 to $ 2,751 in 2008 Q3, whereas the gap between the male and female new hire earnings remained constant at around $1200 per month.

CONCLUSION AND FURTHER RESEARCH POTENTIAL

The LED data provide a wealth of labor market information at both the aggregate and micro levels. This is the only labor market flow data, which is available at the state and the county levels. Since the data is also available at the industry level, hence it can be used for various policy related studies. For example, LED has been used for the targeted response to economic shocks (Saleh, 2009).In the state of New Jersey, information from LED has been used to provide relief to dislocated workers in the wake of financial crisis. LED data can also be used for economic assessment and industry targeting. The gender and age specific information can be used for better targeting of the groups in need. The LED based QWI is still evolving. Much better and detailed economic analysis can be performed when relatively long-term information are available overtime.

REFERENCES

Abowd, John M. et.al. 2006. "The LEHD Infrastructure Files and the Creation of Quarterly Workforce Indicators," *United States Census Bureau Technical Paper TP-2006-01*.

Hall, Robert E., 2006. "Job Loss, Job Finding, and Unemployment in the U.S. Economy over the Past Fifty Years," *2005 NBER Macroeconomics Annual. Cambridge, MA: National Bureau for Economic Research.*

Pissarides, Christopher 2000. Equilibrium Unemployment Theory, Second Edition, *The MIT Press.*

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

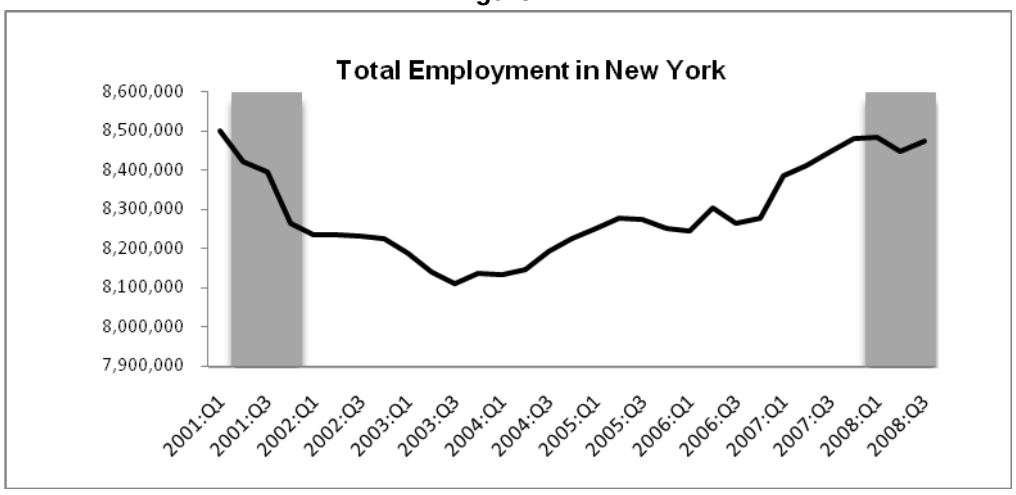
Shimer, Robert, 2005. "Reassessing the Ins and Outs of Unemployment," mimeo, University of Chicago.

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

ENDNOTES

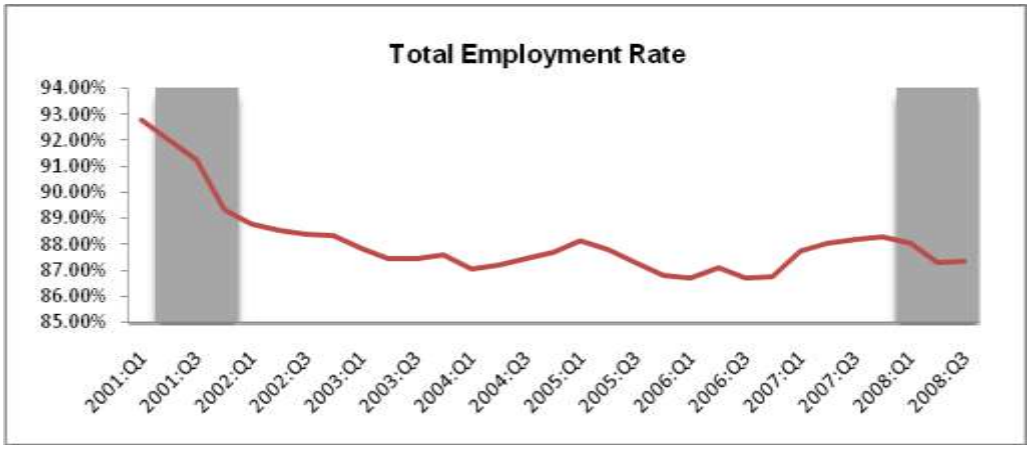
1. Net Job Gain = $E_t - E_{t-4}$, where E_t is employment in year t.
2. Net Job Gain Rate = $\frac{Net\ Job\ Gain}{\frac{(E_t + E_{t-4})}{2}}$ x 100

Figure: 1



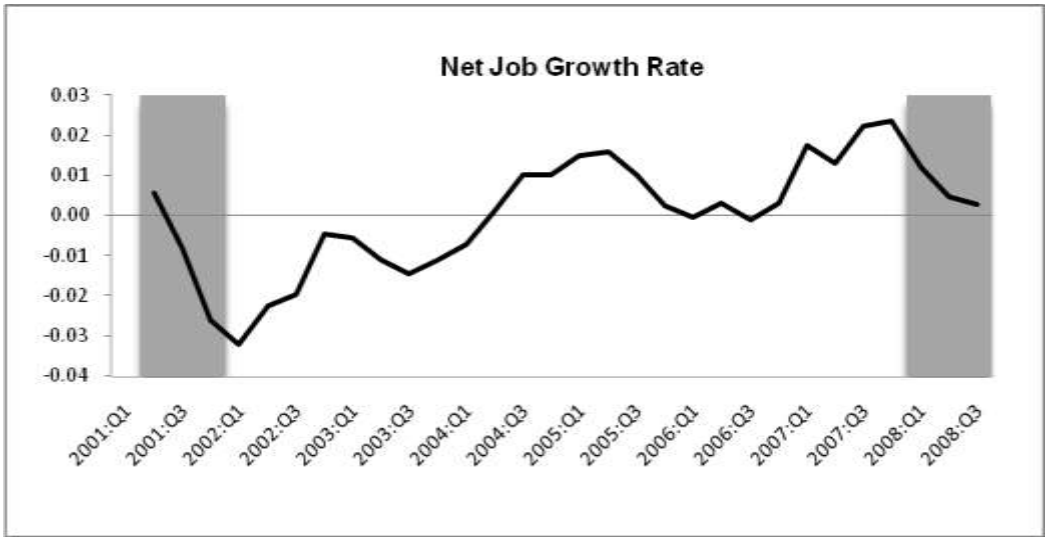
Source: Quarterly Workforce Indicators, United States Census bureau

Figure: 2



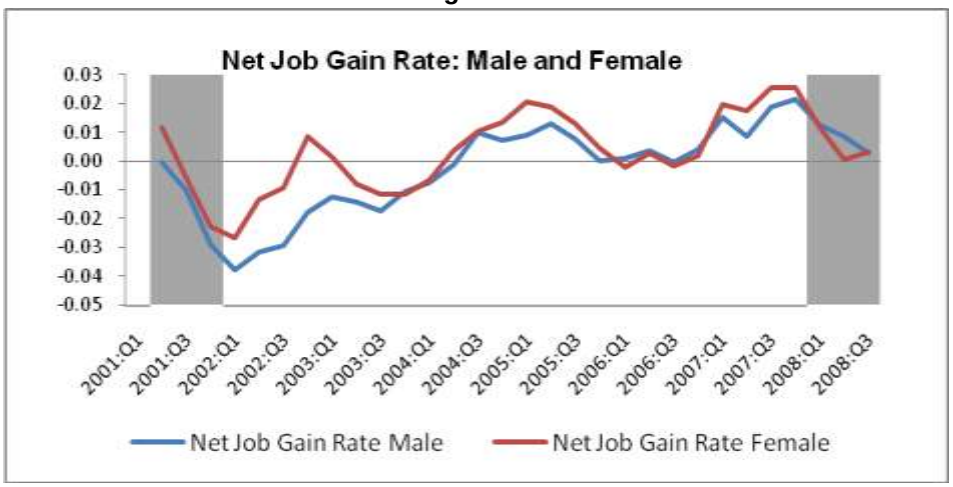
Source: Quarterly Workforce Indicators, United States Census bureau

Figure: 3



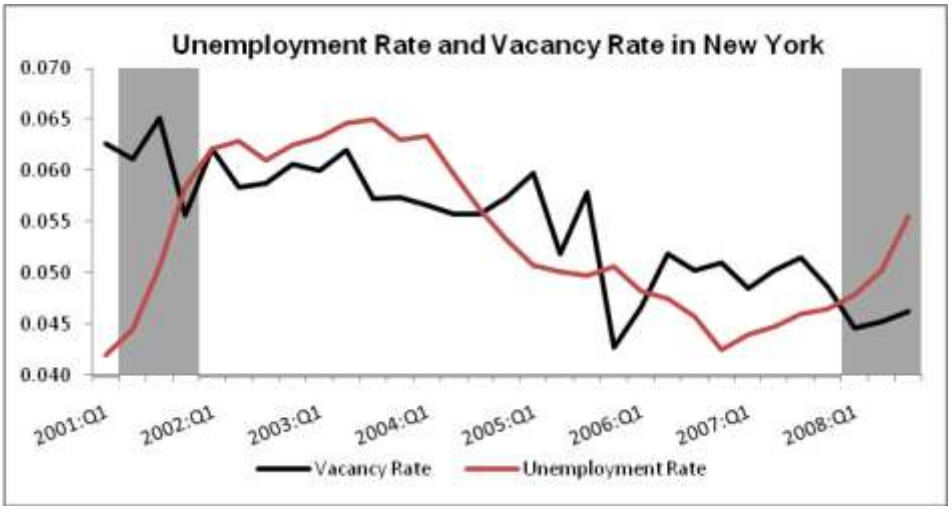
Source: Quarterly Workforce Indicators, United States Census bureau

Figure: 4



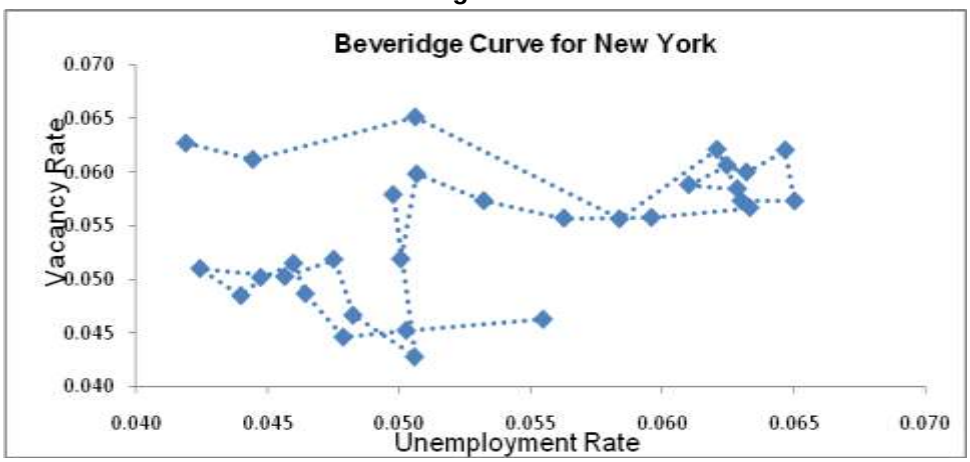
Source: Quarterly Workforce Indicators, United States Census bureau

Figure: 5



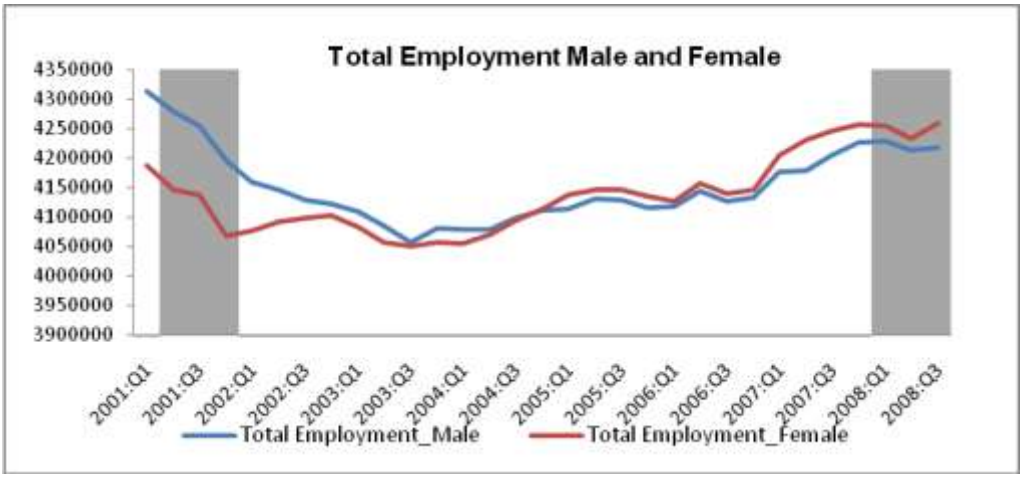
Source: Quarterly Workforce Indicators, United States Census bureau

Figure: 6



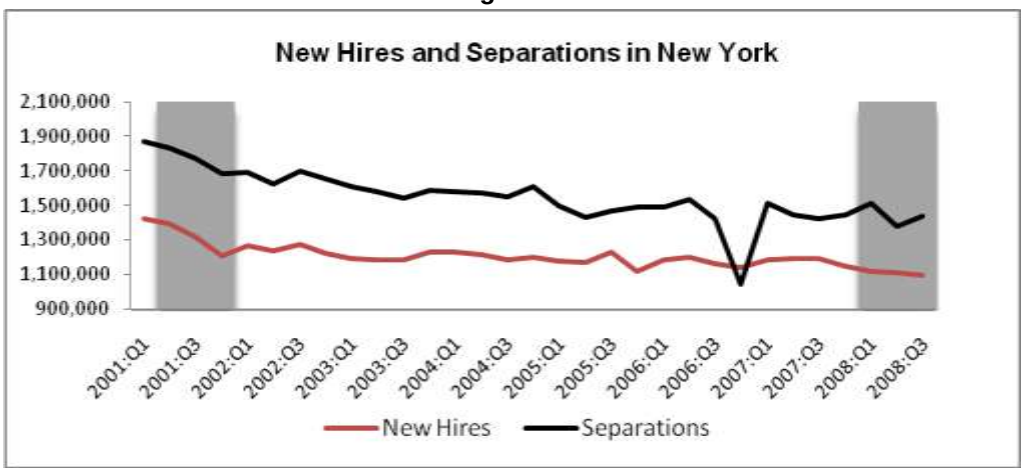
Source: Quarterly Workforce Indicators, United States Census bureau

Figure: 7



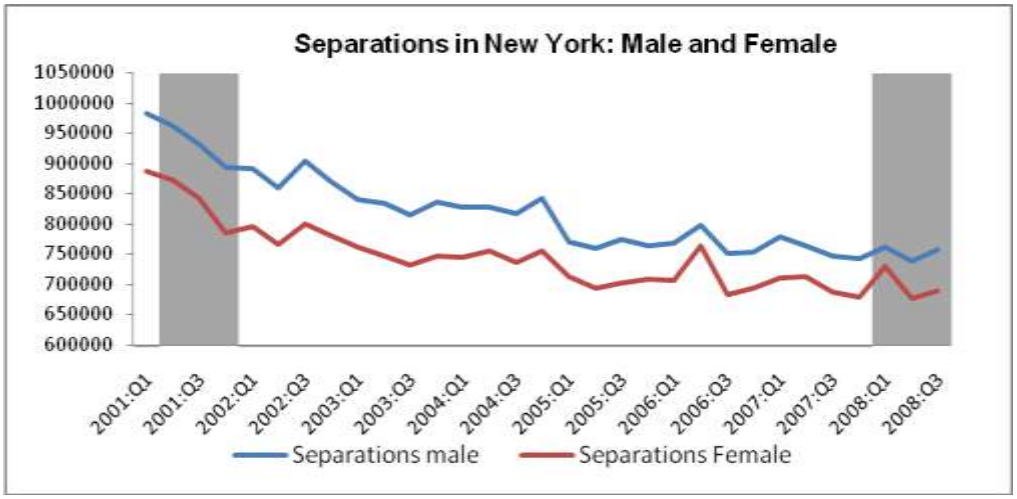
Source: Quarterly Workforce Indicators, United States Census bureau

Figure: 8



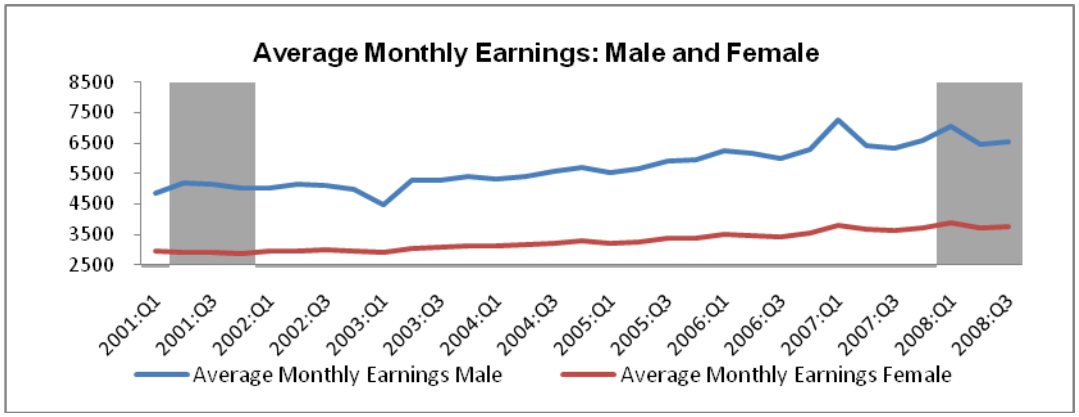
Source: Quarterly Workforce Indicators, United States Census bureau

Figure: 9



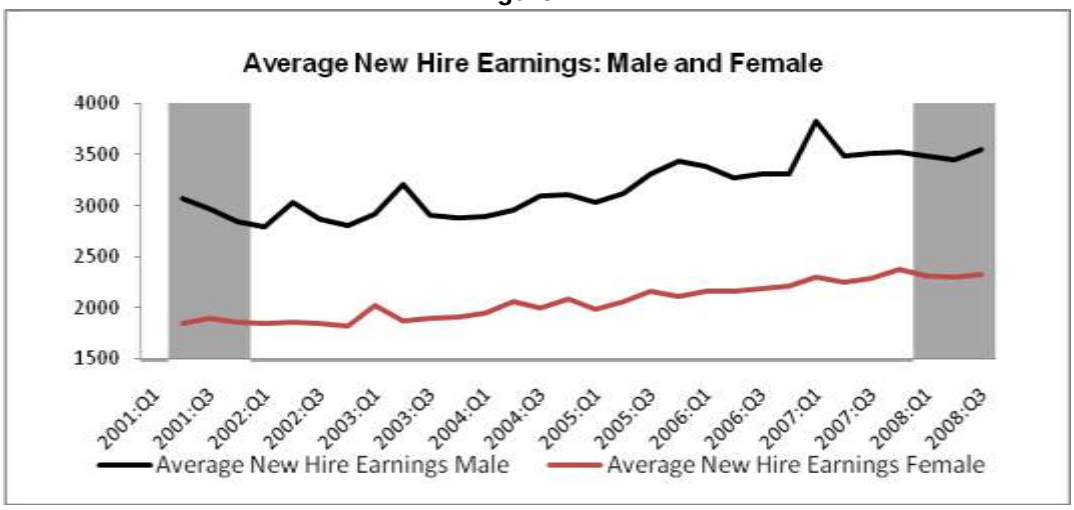
Source: Quarterly Workforce Indicators, United States Census bureau

Figure: 10



Source: Quarterly Workforce Indicators, United States Census bureau

Figure: 11



Source: Quarterly Workforce Indicators, United States Census bureau

Table: 1

Vacancy and Labor Turnover Aggregate Summary Statistics

	Vacancies	Hires	Separations	Unemployment
Mean	0.055	0.135	0.174	0.053
[Standard Deviation]	[0.006]	[0.01]	[0.019]	[0.007]
Correlation with...				
Vacancies	1	0.837079811	0.723719493	0.440879196
Hires		1	0.860771175	0.190190631
Separations			1	0.392727796
Unemployment				1
Autocorrelations				
AR(1)	0.586*	0.648*	0.575*	
AR(2)	0.522*	0.403*	0.438*	
AR(3)	0.348*	0.280*	0.363*	

Source: Author's calculations based on QWI and LAUS aggregate data (seasonally adjusted). Statistics are based on data from 2001 Q1 through 2008 Q3. Asterisks (*) denote significance at the 5 percent level.

The Basic Economics of Health Care Reform

William P. O'Dea*

I. INTRODUCTION

In this brief paper, I will not offer a plan to reform the delivery of health care in the United States. The purpose of this paper is more modest. It is an attempt to highlight the contribution that a grasp of basic economic theory can make to the current debate on health care reform.

Before we consider the lessons that economic theory can teach, it is important to put the health care debate in its proper context. This is something that this is not often done. Currently, the United States devotes sixteen percent of its GDP, or approximately $2300 billion, to the provision of health care. The average for all countries belonging to the Organization for Economic Cooperation and Development (OECD) is 8.9 percent. Per family, the United States is spending $19,000 on health care. If our expenditures on health care were to equal to the OECD average, expenditures per family would be approximately $10,000. The $9000 difference represents a sizable percentage of real median family income, which is slightly over $50,000. It is also worth noting that while other advanced industrial economies provide health coverage to nearly all of their citizens, fifteen percent of Americans, or 46.3 million people, lack health insurance (US Bureau of the Census, 2009).

The rate at which health care costs are growing is alarming. In 1970, the U.S. devoted seven percent of GDP to health care. By 2015, the President's Council of Economic Advisers projects that health care spending will exceed 20 percent of GDP (Rivlin and Antos, 2007).

Some economists have attributed the stagnation in median real family income to the rapid growth in medical costs. The majority of Americans receive health insurance coverage through their employers. Over the last decade, the cost of providing medical coverage to a family has increased at an annual rate of 8.9 percent (Kaiser Family Foundation and Health Research and Education Trust, 2009). The need to cover higher insurance costs leaves employers with less money to fund wage increases. The burden of funding higher insurance has doubtless contributed to the decrease in the percentage of employers offering health insurance from 69 percent in 2000 to 60 percent in 2009 percent (Kaiser Family Foundation and Health Research and Education Trust, 2009).

Rapidly growing health care costs are a budgetary challenge for all levels of government. Currently, Medicare is the third largest line item in the federal government budget after national defense and Social Security. In 2008, federal government spending on health care exceeded $752 billion. This amount represented 25.2 percent of total federal expenditures or 5.3 percent of GDP. This amount does not include the cost of excluding the portion of health insurance premiums paid by employers from the definition of gross income in the personal income tax. The Congressional Budget Office predicts that by

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

2045 federal spending on health care will be approximately 18.4 percent of GDP, which is the long-term average of federal revenues as a percentage of GDP (Rivlin and Antos, 2007). No money would be left to cover spending on such things as defense, Social Security, agricultural price supports and interest on the public debt. Leaving the structure of the Medicare and Medicaid programs unchanged would present federal budget makers with an unpalatable set of choices: massive cuts in other spending programs, massive tax increases or borrowing on the grand scale.

Jagadeesh Gokhale and Kent Smetters (2003) approach the question from a different angle. They calculated the present value of federal revenues from here to etermity and the present value of future federal expenditures. They found that the present value of federal expenditures exceeded the present value of revenues by approximately $44 trillion. The bulk of this gap, $36.6 trillion, was due to Medicare. By contrast, Social Security accounted for only $7 trillion of the fiscal gap. Eliminating a fiscal gap of this magnitude would require policy makers to take drastic action. For example, an immediate and permanent doubling of payroll taxes that fund Medicare and Social Security would eliminate the imbalance. An alternative would be an immediate and permanent 45 percent reduction in the level of Medicare and Social Security benefits. Calculating an economy's fiscal gap requires selecting the proper discount rate and predicting future demographic trends, the rate of productivity growth, the rate at which health care spending will grow. Their estimate is particularly sensitive to the rate of growth of health care spending per capita. Their baseline computations assume that health care expenditures will grow by 1 percent more than the growth of real GDP. If health care expenditures grow by.5 percent lmore than the increase of real GDP, the fiscal imbalance is reduced to $29.5 trillion. If health care expenditures grow by 1.5 percent more than the increase of real GDP, the fiscal imbalance increases to $64 trillion. In Laurence Kotlikoff's (2006) judgment, the latter is the more likely outcome. Gokhale and Smetter's sensitivity analysis emphasizes the importance of controlling the rate of growth of health care expenditures. Finally, their analysis highlights the importance of dealing with the fiscal imbalance sooner rather than later. Assuming that no changes are made in the major entitlement programs, they show the size of the fiscal imbalance growing from $44 trillion in 2002 to $54 trillion in 2008, with most of the growth being attributable to the Medicare program. The clear message is that if federal expenditures are to be brought into alignment with federal revenue/es the need to reform the Medicare program cannot be avoided and should not be postponed.

II. LESSON 1: THE FINAL GOOD IS HEALTH

One possible explanation for the large percentage of U.S. GDP that is devoted to health care is that it simply reflects the decision of Americans to consume more medical services as they have become wealthier. It is undeniable that advances in medical technology have led to dramatic improvements in life expectancy and the quality of life.

However, a basic lesson that we teach our students in principles of macroeconomics is that GDP is intended to measure the rate at which the economy is producing final goods and services. We

emphasize the importance of distinguishing between finals goods such as automobiles and the intermediate goods such as steel and glass that go into their production. We teach our students that including both final goods and intermediate goods in the computation of GDP would be double counting and would result in an overestimate of the economy's production of final goods.

The computation of GDP includes expenditures for medical goods and services, which are in the nature of intermediate goods. Very few consumers undergo open heart surgery or have MRI scans performed, because they enjoy the experience. Rather, medical goods and services are inputs in the production of the final good, which is health. In the production of health, health care is not the only input. Also important are lifestyle choices regarding alcohol consumption, diet and smoking and genetic inheritance.

Clearly, measurement issues preclude including the value of health in the computation of GDP. It is then reasonable to ask whether a higher level of consumption of medical care leads to better health outcomes. International and domestic comparisons indicate that the higher level of expenditures on health care in the United States has not lead to a correspondingly higher level of health for Americans.

Two common metrics that are used to assess the performance of a country's health care system are longevity at birth and infant mortality. In both cases, the United States does not fare well in comparison with other advanced economies. Life expectancy at birth in the US is 78.1 years. (OECD, 2009a) The average for all industrial countries is 79.1 years. The highest life expectancy at birth is 82.6 years in Japan. Infant mortality in the United States (per thousand live births) is 6.7. The average for all industrial countries is 3.9. Sweden's infant mortality rate of 2.5 is the lowest.

As is the case with any complex system, the health care sector of the United States has its strengths and weaknesses. OECD (2009b) data indicate that the US health care system is the most effective in treating cancer. However, the data also indicate that it does a poor job of managing chronic diseases such as asthma, diabetes and hypertension. The general thrust of an international comparison of health care systems is that the dramatically higher level of health care spending in the United States has not produced dramatically better health care outcomes for Americans. Indeed, it appears that in general we are spending more and getting worse outcomes.

Since 1992, the Dartmouth Atlas Project has been engaged in a study of Medicare spending in 3436 hospital service areas and 306 hospital referral areas in the United States (Fisher et al., 2009a). The project reveals startling differences in Medicare spending within the United States. In 2006, Medicare spending per enrollee ranged from a high of $9564 in New York to a low of $5311 in Hawaii. The expenditure variations between the hospital referral areas are even more dramatic. Expenditures per enrollee in Miami, the referral area with the highest costs, were $16,351, which was more than three times the expenditure level of $5311 per enrollee in Honolulu, the referral area with the lowest costs. Not only do the levels of spending vary widely but so also do the growth rates. Between 1992 and 2006, the annual growth rates of Medicare spending per enrollee were 8.31 percent in McAllen, Texas, the referral area with the highest growth rate, and only 1.63 percent in Honolulu, the referral area with the lowest

growth rate. Due to the power of compound growth, these differences in growth have important implication for the Medicare budget. Fisher et al. (2009a) find that if the growth of Medicare spending could be reduced from 3.5 percent, the average annual rate of growth between 1992 and 2006, to 2.4 percent, the average annual rate of growth in San Francisco, the cumulative savings to the Medicare program by 2023 would be $1.42 trillion. Instead of a $660 billion deficit in 2023 Medicare would have surplus of $758 billion (Fisher et al., 2009b).

The regional differences cannot be explained by differences in the health of the populations in the different referral areas. Fisher et al. (2009b) note that "marked regional differences remain after careful adjustments for health and there is no evidence that health is decaying more rapidly in Miami that in" Honolulu (849). The differences also cannot be explained by technology since all referral areas have access to the same medical technology or by the payment system since most medical care providers in the United States are compensated on a fee-for-service basis. Fisher et al. found that the differences were due to the fact that physicians in the higher spending regions "were much more likely than those in the lower spending regions to recommend discretionary services, such as referral to a subspecialist for typical gastroesophacal reflux or stable angina (850)." Importantly, the provision of more medical services in the higher spending regions did not result in improved medical outcomes. Indeed, Fisher et al. report that "the quality of care and health outcomes are better in lower-spending regions and that there have been no greater gains in survival in regions with greater spending growth (850)."

Taken together, the international data and the work of the Dartmouth Health Atlas suggest that reductions in expenditures for medical care need not result in deterioration in the health status of a population. The key point is that our focus should be on the final good, health, rather than on the intermediate good, health care services.

III. LESSON 2: THIRD PARTY PAYERS ARE AN ESSENTIAL ELEMENT OF THE HEALTH CARE SYSTEM

Given the cost of medical procedures, the overwhelming majority of American families are in no position to self-insure. Most American families do not have the income, the liquid assets or the borrowing capacity to pay for a cancer treatment or open heart surgery. Indeed, the debts incurred to finance medical emergencies are a leading cause of bankruptcy filings in the United States (Himmelstein et al., 2009).

Medical expenditures have been characterized as "large, lumpy and uncertain." Consequently, some mechanism is needed to spread the costs incurred by the unfortunate few over a larger population. Some sort of third party payer is essential.

Unfortunately, the analysis in George Akerlof's (1970) classic paper, "The Market for Lemons", suggests that a purely private system in which the decision to purchase health insurance is left to individuals is unlikely to function optimally. There is an informational asymmetry between insurance

buyers and sellers. Even if individuals purchasing health insurance were forced to undergo a stringent physical examination, the buyer still possesses two important pieces of information not easily accessible by the seller: his or her health history and that of his or her family. In addition, buyers would have an incentive to behave strategically by putting off the purchase of health insurance until they need it. Knowing these things, insurance sellers will assume that the typical purchaser of insurance will have higher health care costs than the average person in the general population and set their premiums accordingly. High premiums would drive out the healthiest members of the group considering the purchase of health insurance leading to a further increase in premiums followed by a further reduction in the size of the market for health insurance and so on. In a limiting case, the private market for health insurance could disappear. In addition to setting high premiums, insurance sellers would have an incentive to refuse coverage for preexisting conditions. The result is that individual health insurance policies, assuming they are available, would be expensive and limited in their coverage.

When insurance is provided to a large enough population, the problem of asymmetric information disappears. Predicting the number of claims that a large population will generate in a year and their cost is an actuarial problem. The key issue is how to create a sufficiently large risk pool so that the health status of any particular person becomes a non-issue. Internationally, a variety of approaches have been used. In the United Kingdom, the National Health Service manages the delivery of health care. The nearest analogue in the United States would be the Veterans Health Administration system. In France and Canada, medical care is privately provided but the entire population belongs to a single-payer national health insurance system. In the United State, Medicare would be a close approximation. In Switzerland, health insurance is privately provided. However, the system is heavily regulated by the government. All citizens are required to purchase health insurance with subsidies being provided to lower income individuals. Insurance companies are required to offer a basic plan, whose elements are set by the government, on a not-for-profit basis. They are allowed to earn a profit on more elaborate plans. For a more complete description of the Swiss system, see Nelson Schwartz (2009). In the United States, the Wyden-Bennett proposal is a close approximation of the Swiss system.

The essential point is that government has to be involved in the provision of health insurance. And this is true in the United States. Most Americans receive health insurance through their employers. The portion of the health insurance premium paid by the employer is not part of the employer's taxable income. The Office of Management and Budget (2009) estimates that in 2008 this exclusion cost the federal government $142 billion in lost tax revenues. This is the single largest tax expenditure. In addition to being expensive, this exclusion is regressive since its value depends on an individual's marginal tax rate. The exclusion is also uncapped. The health insurance plan provided to the 400 managing directors at Goldman Sachs, which costs $40,000, receives the same tax treatment as the average policy which costs $13,375. Finally, individuals who purchase health insurance must pay their premiums out of their after tax income and thus receive no tax benefit.

However, the preferential treatment of employer provided health does have an important social benefit: the administrative costs of providing insurance to large groups is much smaller than the administrative cost of providing individual coverage. It is generally estimated that ten percent of the premiums paid for group policies are used to cover administrative costs versus thirty percent for individual policies. For the average family policy, the employer contribution is $10,000. For a family in the 25 percent marginal tax bracket, the revenue loss to the federal government would be $2500. However, this loss would be balanced by an administrative cost savings of $2700. A major drawback of employer provided insurance is that since individuals only pay attention to their direct costs they do not develop a proper appreciation of the cost of the U.S. health care system.

IV. LESSON 3: INCENTIVES MATTER (AND SO DOES CULTURE)

Since the prices that individual consumers pay for medical services are generally less than marginal cost, it is easy to demonstrate that the condition for Pareto efficiency in production and exchange will not be met. More precisely, the marginal rate of product transformation between a numeraire good and medical care will be less than the marginal rate of substitution between the goods. The implication is that medical care will be overproduced and other goods and services will be underproduced.

However, the reality is more complicated than this blackboard exercise. The opportunity to purchase an MRI with a list price of $1700 for an out-of-pocket payment of $35 is not the same as being offered the opportunity to purchase a ticket to “Wicked” or a Bruce Springsteen concert for $35. Medical procedures tend to be unpleasant and carry an element of risk. Very importantly, most people lack the technical expertise to judge whether a drug, test or procedure is appropriate for their condition. Most of us rely on our physicians to identify the medical services we need.

Given that the consumers of medical care are not the sole, or even the primary, decision makers, it is important to consider the incentives that face medical care providers. The majority of doctors and hospitals in the Unites States are compensated on a fee-for-service basis. The more medical services they provide the more income they receive. When prescribing treatments, physicians are well aware of the insurance coverage of their patients. It is standing joke in my community that the first piece of information collected when a patient shows up at the local emergency room is the name of the patient’s insurance carrier.

The danger of combining medical care providers who work on a “fee for service” basis with patients who are only directly responsible for a small percentage of the cost of their medical care is obvious. Physicians, knowing that their patients are only responsible for a small fraction of the cost of the health care services they consume, have no incentive to look for the most cost effective treatment options. Fisher et al. (2009b) find that in the higher-spending regions doctors are more likely to recommend discretionary services and less likely to recommend palliative care.

The incentives to overprescribe medical care are compounded when physicians have an ownership stake in scanning centers, laboratories and hospitals. A frequent complaint is that doctors are more generously compensated when they prescribe expensive tests or procedures than when they take the time to take a detailed medical history or to help a patient manage a chronic condition. Fisher et al. (2009b) note that existing compensation schemes reduce the incentives to health care providers to improve the quality of outcomes. They point out that: "hospitals lose money when they improve care in ways that reduce admissions, and they lose market share when they don't keep pace in the local medical arms race (850)."

As a case study, Atul Gawande (2009) examined the culture of the health care communities in to two hospital-referral regions in Texas: McAllen and El Paso. In 2006, the Dartmouth Atlas reported that inflation-adjusted total Medicare spending per enrollee was $14, 946 in McAllen and $7504 in El Paso (Fisher et al., 2009a). Gawande reports that both regions had similar demographics and health statistics and comparable medical technologies were employed. In both regions, medical care providers worked on a fee-for-service basis. The major difference between the regions was that in McAllen the medical care community was more entrepreneurial and focused on income maximization. In El Paso, the focus was on health. Importantly, the provision of more medical services in McAllen did not lead to better health outcomes. There was no apparent difference in the quality of service in the two regions.

High quality, low cost health care providers such as Kaiser Permanente, the Mayo Clinic, and the Veterans Health Administration have incentive structures that are focused on the provision of health rather than health care services. Physicians in these systems are on salary and thus derive no additional income from prescribing more services. These systems provide primary care physicians, specialists, pharmacies, hospitals and clinics under one roof. Thus, the provision of health care is coordinated. Because these systems have a long-term relation with their patients they have an incentive to manage chronic conditions such as diabetes and hypertension effectively. Because there is a constant churn in the population covered by private insurance providers, these companies lack the incentive to manage chromic conditions. While managing diabetes is much less expensive than the cost of dealing with the complications of the disease if left unchecked, an insurance company knowing that one its competitors might realize the future benefit might rationally choose to underinvest in prevention. Finally, the high quality, low cost providers have invested in electronic medical records and have analyzed the information collected to refine their delivery procedures to provide higher quality care. [See Phillip Longman (2005) for a detailed analysis of the Veterans Health Administration.]

V. LESSON 4: RATIONING IS UNAVOIDABLE

Our last lesson is the most basic. Rationing is unavoidable. If we define an absence of rationing as a situation where every medical service with a positive marginal benefit (or expected marginal benefit) is provided, then no economy can afford to provide this level of care. In some way, any economy is going to

have to deny patients access to medical services whose marginal benefits are positive but are less than the marginal cost of provision. In the United States, the uninsured or under-insured are rationed out of the market for health services.

Since we cannot do everything that everyone would like to see done, it is important that the resources devoted to the provision of medical care be used as well as possible. To an economist, it is obvious that the use of public resources to identify which medical procedures and drugs are the most cost effective is money well spent. Opponents argue that such studies might lead to an outcome in which government tells physicians how they will practice medicine and patients are denied access to certain drugs and procedures. The recent controversy that resulted from the United States Preventive Services Task Force's recommendation that mammograms to detect breast cancer be performed every other year starting at age 50 is a good example. But, for a woman without health insurance, the debate over whether cancer screening should begin at age 40 or 50 and mammograms should be performed every year or every two years is a matter of purely academic interest.

VI. CONCLUDING COMMENTS

The rapidly growing share of American GDP devoted to the provision of medical care is a threat to the health of the economy that must be confronted sooner rather than latter. The health care delivery system must be reformed. We need to devise a system of incentives that encourage health care providers to focus on the quality of health care outcomes rather than the delivery of health care services. Finally, if rationing is inevitable, we have to face this reality and design a rationing scheme that is equitable and cost effective.

REFERENCES

Akerlof, George A. 1970. "The Market for 'Lemon': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics*, 84(3): 488-500.

Fisher, Elliott S., Julie Bynum and Jonathan Skinner. 2009a. *The Policy Implications of Variations in Medicare Spending Growth: A Dartmouth Atlas Project Brief Report*. The Dartmouth Institute for Health Policy and Clinical Practice.

Fisher, Elliott S., Julie Bynum and Jonathan Skinner. 2009b. "Slowing the Growth of Health Costs-Lessons from Regional Variation." *The New England Journal of Medicine*, 360(9): 849-852.

Gokhale, Jagadeesh and Kent Smetters. 2003. "Fiscal and Generational Imbalances: New Budget Measures for New Budget Priorities." *Federal Reserve Bank of Cleveland Policy Discussion Papers*, no. 5.

Gawande, Atul. 2009. "The Cost Conundrum: Expensive Health Care Can is Harmful." *The New Yorker*, 89(16): 36-44.

Himmelstein, David U., Deborah Thorne, Elizabeth Warren, and Steffie Woolhandler. 2009. "Medical Bankruptcy in the United States, 2007: Results of a National Study." *The American Journal of Medicine*, 122(8): 741-746.

Kaiser Family Foundation and Health Research & Educational Trust. 2009. *Employer Health Benefits: 2009 Annual Survey*. Menlo Park, CA and Chicago, Il.

Kotlikoff, Laurence J. 2006. "Is the United States Bankrupt?" *Federal Reserve Bank of St. Luis Review*, 88(4): 235-249.

Longman, Phillip. 2005. "The Best Care Anywhere." *Washington Monthly*, January/February. www.washingtonmonthly.com/features/2005/0501.longman.html.

Office of Management and Budget. 2009. *Analytical Perspectives: Budget of the U.S. Government, Fiscal Year 2010.* Washington, D.C.: U.S. Government Printing Office.

Organization for Economic Co-operation and Development. 2009a. *OECD in Figures2009: Statistics of the Member Countries*. Paris: OECD.

Organization for Economic Co-operation and Development. 2009b. *OECD Health Data 2009*. www.oecd.org/health/healthdata.

Rivlin, Alice M. and Joseph R. Antos, eds. 2007. *Restoring Fiscal Sanity 2007: The Health Spending Challenge*. Washington, D.C.: Brookings.

Schwartz, Nelson D. 2009. "Swiss Health Care Thrives Without Public Option." *The New York Times*. September 30, A1, A3.

U.S. Bureau of the Census. 2009. *Income Poverty and Health Insurance Coverage in the United States: 2008*. Washington, D.C.: U.S. Government Printing Office.

Japanese versus American Financial Crises: Are There Any Lessons to Learn from the Japanese Experience?

Arina Shnaider*

ABSTRACT

The purpose of this paper is to analyze the origins of the Japanese crisis of the 1990s and determine the ways the Japanese economy was recovered. Having done that, the Japanese crisis will be compared to the American crisis of 2008-2009 in terms of the causes of the economic downturn. Finally, after considering the similarities between the two crises, the conclusion will be made about what policies that the Japanese government implemented are applicable to the current situation in the United States.

INTRODUCTION

The current economic situation not only in the United States but around the world as well has received great attention for the past year-and-a-half. There is at least one article in *The New York Times* every day on how the government is trying to stimulate the economy with hundred-billions-dollar stimulus packages or by lowering interest rates and stimulating consumption. In addition, there have been a number of articles in *The New York Times* and *The Economist* that draw a comparison between the financial crisis we face today and that faced by the Japanese in the 1990s. Moreover, the authors of these articles are trying to extract some lessons from the Japanese and make them applicable to the American crisis of 2008-2009. For instance, in his *New York Times* article "Japan's Big-Works Stimulus is a Lesson," Martin Fackler writes about Japan's numerous infrastructure works done during the late 1990s that had little long-term economic benefit to the society. He also argues that, while keeping in mind Keynesian philosophy on government spending role during the economic downturn, the government should increase spending wisely on infrastructure and not simply be "digging out holes and filling them back up." In *The Economist* article "Big Government Fights Back," the author compares the size of spending by the Japanese government in the 1990s and the American government these days and concludes that taking into account the sizes of the two economies the U.S. government would need to spend far more in stimulus packages than it has as of the end of 2008.

Having come across several articles mentioned above, I decided to concentrate my research on comparing the two financial crises: the Japanese crisis of the 1990s and the American crisis of 2008-2009. First, I shall present background information on some of the causes of the Japanese financial crisis. Second, I shall study the Japanese financial system and determine the changes that were implemented

* SUNY-Oneonta, Oneonta, NY. 2009 Winning paper from the NYSEA Undergraduate Paper/Presentation Contest.

after the crisis occurred. In addition to analyzing the financial system, I am going to examine the monetary and fiscal policies Japan adopted during the financial crisis. Following that, I shall present my analysis of how applicable the changes to its financial system and monetary and fiscal policies that Japan had to make in response to the crisis are to the crisis taking place in the United States at this moment. After studying the nature of the Japanese economy and culture, I shall be able to say what some of the lessons that the United States should take out of the Japanese crisis of the 1990s are. Certainly not all policies implemented by the Japanese government are suitable for the U.S. economy simply because of the difference in the size of the two economies and the time period during which the current financial crisis is taking place. However, given that there are similarities in the nature and causes of the two crises, certain strategies can be and should be taken by the American government as lessons from Japan.

LITERATURE REVIEW

Several previous papers have done a comparison of Japan and the U.S. including an article from the *New England Economic Review* (2001) by Lynn E. Browne "Does Japan Offer Any Lessons for the United States?" In her article, Browne compares Japan's rapid growth in the 1980s to the United States' economic growth in the 1990s. In both situations stock market prices were rapidly rising which resulted in great economic growth. However, the only difference between the two countries' economic growth rates was that the growth in Japan, in addition to rapidly increasing stock prices, was accompanied by rapidly rising real estate values and consequently the value of land, which caused the so-called housing bubble. Despite this difference in the two economic situations in Japan and the United States, Browne concludes that American policymakers should take some important lessons from the Japanese crisis of the 1990s and avoid making the same mistakes such as discontinuing stimulus packages too soon and lowering interest rates that brought few positive results.

Another paper that focuses on comparing Japan to the United States in terms of economic growth and slowdown was written by Federal Reserve economists Alan Ahearne, Joseph Gagnon, Jane Haltmaier, and Steve Kamin in June 2002. It is called "Preventing Deflation: Lessons from Japan's Experience in the 1990s." The authors of this work draw parallels between Japan in the 1990s and the United States in 2001 in terms of monetary policymaking in a situation of a zero nominal interest rate. As in the previous paper mentioned above, these authors tried to prevent the same situation as in Japan in the 1990s from happening in the United States in 2001. In other words, they described what the Japanese government did in order to pull the economy back to normal growth and pin pointed the mistakes that the government made. The paper then concluded with what the U.S. government needed to avoid and what to adopt in terms of monetary and fiscal policy making in order to avoid making the same mistakes as the Japanese government.

As described above, previous papers were written on comparing the Japanese and the United States' economic growth in the 1980s and 1990s respectively and comparing the monetary and fiscal policies of

the two countries in the 1990s (Japan) and 2001 (the U.S.). My work will concentrate on answering the question "Can Japan's Financial Crisis of the 1990s Serve as a Lesson for the United States' Macroeconomic Policymakers in the Current Financial Crisis?" I shall base my idea on Browne's work in comparing the two economies. However, the time period considered will differ from that described in her work. I shall also take the idea of comparing Japan's and the United States' monetary and fiscal policies from the authors of "Preventing Deflation: Lessons from Japan's Experience in the 1990s" and apply it to a different time frame. As a result, I keep the Japanese economic crisis of the 1990s as it has been done in previous works but compare it to the financial crisis that is currently taking place in the United States.

HYPOTHESIS AND METHODOLOGY

Given the question posed throughout this research work, "Can Japan's Financial Crisis of the 1990s Serve as a Lesson for the United States' Macroeconomic Policymakers?", my hypothesis for this paper is that there are certain strategies that the United States should keep in mind and possibly even implement based on the Japanese experience in the 1990s to resolve the current economic crisis. I am going to use comparative and logical reasoning in order to answer my research question. I shall compare the two financial crises – the U.S. and Japanese – then study the policies and actions of both the Bank of Japan and Japanese government that were undertaken to resolve the crisis as quickly as possible. Finally, taking into consideration the difference in sizes of the American and Japanese economies, different cultures and different time periods the crises occur in, I shall attempt to determine whether the strategies used by the Japanese in the 1990s provide any lessons to American macroeconomic policymakers about what they should do in response to the current crisis.

THE ORIGIN OF THE CRISIS AND JAPAN'S FINANCIAL SYSTEM

According to Takeo Hoshi and Anil K. Kashyap (2004), the origin of the Japanese financial crisis of the 1990s goes back to rapidly rising stock market prices in the 1980s (see Appendix 1) accompanied by similarly rapidly rising land prices (see Appendix 2). Economic growth was steadily rising in the middle of the 1980s and at the end of 1980s the growth rate dropped significantly by nearly 7 percentage points starting from 1988 through 1993 (see Appendix 3). As a result of rising real estate values the value of land also increased rapidly. Because of the rapid economic expansion, lending and investment were growing. However, when the consumer confidence started to fall and there were more than one default on a mortgage and a loan on property, the value of the land and stock prices began to fall. From 1990 through 1992 the stock prices dropped by almost 60% (from 38,916 down to 15,910 percentage points) and the growth of land prices fell by nearly 17% (from 15% down to -2%). Stock prices have not returned to the 1990 peak even today. Land prices began to rise in 1993, however the prices did not reach the 1984 level even after 6 years.

During the 1970s and 1980s, Japanese economic growth was superior to the other major Western economies.[1] Although, Japan's ability to manufacture high - quality products at low cost was one of the reasons for the rapid growth, government guidance was another factor that resulted in the Japan's superior economy. The Japanese government made strategic moves in determining which industries had high growth potential and therefore needed financial support. In addition to government support, major Japanese companies were isolated from short-term financial pressures because they would form alliances among each other and around a main bank, thus making their ownership shares the majority and providing management independence from outside stockholders. Major companies in Japan primarily received their financing from banks and rarely relied on sales of stock or bonds. However, the emergence of the crisis in the 1990s proved the system to have many weaknesses, which were mainly due to the bank's dependence on land prices since land was used as primary collateral for mortgages and other property loans. Therefore, when the housing market bubble burst, the decline in land value resulted in many losses for banks and, thus led to decreased lending that companies heavily depended on (Krugman, 1999; 63-64).

Although the banking industry was regulated by the government in the 1980s, banks started to break more rules in terms of determining ways to direct their capital. Due to the rapid economic growth and times being good, banks began to extend credits to businesses without making sure that those businesses had good repayment potential. According to Paul Krugman, "Japan's banks lent more, with less regard for quality of the borrower, than anyone else's; and in so doing they helped inflate the bubble economy to grotesque proportions." (Krugman, 1999; 69) This situation is commonly classified as moral hazard because while having government support and knowing that it would step in and help to finance them, banks were prone to taking more risk than they usually would. This problem of moral hazard resulted in many debtors defaulting on payment. The default on payment forced debtors to sell their assets and use the proceeds to pay back the loan. However, selling their assets drove the prices of those assets further down as there were more and more defaults on debts. That is how the housing market collapse occurred. When businesses in the construction industry were unable to repay their debts, they were forced to sell their buildings, particularly residential buildings. As a result, prices of houses that reached their peak only months ago dropped substantially. As the prices of houses dropped, the value of land decreased as well. And since land was used as collateral for most of the loans, its declining value resulted in banks being in big trouble because of the inability to receive the amount lent even after selling all the assets (Hoshi, 2004). This resulted in banks' liability to pay their depositors exceeding their assets side of the balance sheet. As a result, Japan went from an economic crisis to a financial one.

It was clear that along with a number of stimulus packages aimed at curing the financial system and banking industry in particular there was a need for a reform that would change the way the system operated. The major goal was to restore investors' confidence in the financial system and establish stability of the system as a whole. One such reform was the new Bank of Japan Law that became effective on April 1, 1998 and it was an event of major significance in 115-year history of the bank. The

two major principles that the new law established was “independence” of the monetary policy-setting by a politically neutral Policy Board and “transparency” – proper disclosure of policy-making processes. The primary goal of the new law was “to pursue price stability, thereby contributing to the sound development of the national economy, and to ensure smooth settlement of funds among financial institutions, thereby contributing to the maintenance of an orderly financial system” (Bank of Japan, 1997). Another reform of the Japanese financial system was the “Big Bang” deregulation package that mainly aimed at revival of the Tokyo market as a free and active global market (Matsushita, 1997).

JAPAN’S MONETARY AND FISCAL POLICIES

Given the collapse of the financial system and effect of that on consumer confidence, a standard response to this situation would be to cut interest rates in order to restore borrowing and spending and help to stimulate economic growth. At first the Japanese government was slow in cutting short-term nominal interest rates after the bubble burst, but eventually interest rates were cut to zero (Krugman, 1999; 74). According to Table 1 in Appendix 4, it took a year-and-a-half (1991) since the beginning of the crisis for the Bank of Japan to start slowly cutting interest rates. It took another five years, when the interest rates reached almost zero in 1996 as indicated by the graph in Appendix 4.1. However, “perhaps because of its aging population, perhaps also because of a general nervousness about the future, the Japanese public does not appear willing to spend enough to use the economy’s capacity, even at a zero interest rate” (Krugman, 1999; 73). This situation is referred to as a liquidity trap by economists. The liquidity trap is the situation when monetary policy becomes ineffective in helping the economy recover from a recession through cutting interest rates. The monetary policy ineffectiveness is due primarily to consumers, banks and firms becoming more risk averse and giving more preference to liquidity of cash instead of using the credit offered (*The Economist*). The question then becomes what is the next step the government or the central bank can take to help the economy recover.

The next option the Japanese government chose was to increase government spending and thus create jobs. Since the beginning of the 1990s, the government introduced a number of stimulus packages that were directed to building roads and bridges whether the country needed them or not. To demonstrate the amount of government spending in stimulus packages, Paul Krugman presents the statistics that “in 1991 Japan’s government was running a fairly hefty budget surplus (2.9 percent of GDP); by 1996 it was running a quite nasty deficit of 4.3 percent of GDP” (Krugman, 1999; 74). However, the fear of long-term impact of the budget deficit and the responsibility of the government to provide the aging population of Japan with proper pension and health care compensation made Japan’s Ministry of Finance increase taxes to reduce the budget deficit in 1997. This move resulted in even deeper recession and in 1998 the government was forced to return to massive government spending policy (Krugman, 1999; 75).

As discussed above, neither the cut of short-term interest rates nor the government's massive spending helped the economy recover from the crisis because neither of the two forced an increase in spending and borrowing, thus leading Japan into a deflationary era (see Appendix 5). The problem, according to Paul Krugman, was that "an economy which is in a liquidity trap needs expected inflation – that is, it needs to convince people that the yen they are tempted to hoard will buy less a month or a year from now than they do today" (Krugman, 1999; 78). In order to fight deflation, the Bank of Japan started large purchases of government debt and setting certain levels for expected inflation to foster consumer spending. Creating expected inflation reduces the real interest rates (nominal interest rates less inflation) at zero nominal interest rates. The real interest rates reflect the true expected return on an asset. Therefore, with falling real interest rates consumers are willing to borrow today because they would have to repay in less valuable yen when their debt matures.

ANALYSIS

BACKGROUND ON AMERICAN CRISIS OF 2008

In 2006 and 2007, the United States experienced rapidly growing stock prices as indicated by the S&P and NASDAQ indices (see Appendix 6) accompanied by the rapid growth of the housing prices (see Appendix 7). Home prices, which peaked in the middle of 2006, slowly started to decline thereafter. An even sharper decline was seen at the end of 2007 and the beginning of 2008 when the price index dropped from 174 down to 145. The home price index has continued to fall even through the first quarter of 2009 as indicated by the graph presented in Appendix 7. The stock market was volatile from September 2007 until September 2008. After September 2008, the market collapsed as illustrated by the graph in Appendix 6. In addition to the housing and stock market collapses, several financial institutions failed during 2008 mainly because they held large amounts of mortgage-backed securities that lost their value once the housing market collapsed. For instance, Lehman Brothers Holdings Inc. filed for bankruptcy at the end of 2008 and Merrill Lynch was acquired by the Bank of America around the same time in order to save the institution from going bankrupt (Yahoo Finance). As illustrated in the graph in Appendix 8, the United States real GDP growth has been falling since 2004. However, there has been an even sharper decline starting from the end of 2006 and continuing to the present.

APPLICABILITY OF THE JAPANESE STRATEGIES TO RESOLVING AMERICAN CRISIS

In order to determine the applicability of the Japanese tactics and strategies to resolving the American crisis these days, the similarities and differences between the two economic downturn situations have to be assessed first. In the next paragraph, I will present the analysis of the causes of the crises and the subsequent situations.

Although there are no quantitative similarities between the Japanese and American crises, it is still possible to draw qualitative parallels between the two. First, the situation with the housing and stock markets seem to be similar in the sense that both experienced rapid growth followed by a drastic collapse at some point. In Japan, after the stock prices tripled in the period of five years (from 1985 through 1990), they dropped down by nearly 60% within the next two years (see Appendix 1). Also, land prices were steadily rising from 1985 until 1990 and had increased by almost 13% when there was a sudden drop in land prices by nearly 20% in the subsequent three years (see Appendix 2). In the United States, stock prices fell by nearly 32% in one month (from September 2008 through October 2008). Constantly rising home prices in the U.S. dropped by almost 36 percent from 2006 through the first quarter of 2009. Second, the effect of the housing market collapse on the financial institutions is similar because in both countries land and residential property were heavily used as collateral for mortgages and other property loans and, thus, when the value of the collateral dropped, the value of banks' mortgages and loans on property followed the same path. Similar to Japan's lending practices, American banks, due to the rapid economic growth, began to give out loans and mortgages without hardly checking for debtors' creditworthiness, which led to more defaults on payments then it otherwise would had been. Although the effect of the housing bubble on the financial system seems to be similar, there is one significant difference in this factor. The Japanese firms were more highly dependent primarily on banks and less so on bond and stock markets for raising capital, when compared to the United States. The United States companies rely both on debt as well as on equity financing. Therefore, some of the policies the Japanese government directed at curing the country's financial system will need a closer observation in terms of their applicability to the American financial and economic situation.

Having assessed the similarities and differences on the major causes of the two crises, the applicability of the policies can now be discussed. The first monetary policy tool the Japanese central bank used was to lower short-term nominal interest rates. Although this strategy proved to be ineffective in Japan because of its aging population and the nature of the society to save rather than spend, cutting interest rates might be appropriate for the United States situation because American do not save as much as the Japanese. The mistake the American government should learn from the Japanese is not to be slow in cutting interest rates as the Japanese government was. However, the government should also be aware of the aging population of baby boomers whose retirement is approaching and, thus pushing the government into paying pension and health care compensations. Because of the aging population and consumer uncertainty about the future, the effectiveness of nominal interest rates cuts in the United States might be less as well. Therefore, the next step would be to stimulate government spending through stimulus packages. This government strategy has been in the news since the time the official recession was announced. The U.S. government has already spent hundreds of billions of dollars in stimulus packages injecting funds in troubled companies through a number of bailout programs and in banks through the process of purchasing their troubled assets in order to stimulate banks' lending and normal course of operation. Specifically, the U.S. government spent $168 billion on tax cuts and rebates in 2008

and $789 billion on another stimulus package in 2009 ("Economic Stimulus," The New York Times). The amount of stimulus packages cannot and should not be compared mainly because the U.S. economy and population is much greater than those of Japan, thus it implies only that the U.S government will have to spend significantly more funds in stimulus packages than Japan did. Also, there is a mistake the American policymakers can learn from the Japanese in terms of when to stop government aid. As mentioned above, the Japanese government, in fear of large budget deficits and their impact in the long-run, stopped stimulus financing in 1997 which led to worsening the recession in 1998 and forced the government to resume its aid to the public as well as private sector. Therefore, the lesson to be learned is that the government should not eliminate stimulus package financing too soon, though, there is a fear of large budget deficit that will have a negative impact in the long-run. However, in the case with the United States, the danger of the budget deficit might not be as great as that of Japan's because the reputation of the dollar and the U.S. government might result in other nations pumping investments in the country which will allow the government to use these funds for stimulus packages.

In addition to the monetary and fiscal policy tools that the U.S. government can use to help the economy recover, the government should consider ways to confront falling expected inflation which is due mainly to consumers' fear and unwillingness to spend. As illustrated by the graph in Appendix 9, from the middle of 2006 through July of 2007, the rate of inflation dropped nearly one percentage point. After that inflation rate started to increase, however, there was another sharp decline in July 2008 of almost one percent in a period of five months. Although inflation rose in the late 2008 and early 2009, as illustrated by the graph, it has been falling since April 2009. Although the United States has not been experiencing deflation yet, certain government actions should be taken to prevent it from happening.

In order to understand which policies are suitable for the current situation, two models of the Phillips curve that differ in how inflation expectations are formed should be examined. The first, "unanchored" Phillips curve is based on the assumption that inflation rates depend mainly on past inflation rates. The logic behind this model is that in the beginning of a recession rising unemployment causes inflation rate to fall. As a result, people expect future inflation rate to follow the same pattern and, thus a deflation spiral forms. The deflation is stabilized only when unemployment starts to fall. The second model is where inflation expectations are "well anchored." This means that expectations about inflation are consistent with the goals and policies of the central bank. Even during severe recessions, people expect the central bank to keep inflation rates positive. The deflationary spiral can form when there is high unemployment and when monetary policy actions are ineffective at curing the economy. The two models of the Phillips curve have different implications for the possibility of deflation to occur. According to John Williams, due to rising unemployment "the estimated probability of deflation is about 30% for 2009 and 85% for 2010 for the Phillips curve model which is based on the "average" behavior of inflation over the past five decades… A Phillips curve model estimate using data since 1993 is consistent with well-anchored inflation expectations and precludes emergence of a deflationary spiral." (Williams, 2009) As Williams suggests, the Federal Reserve should employ all available tools to preserve price stability and keep

inflation expectations well anchored. Taking the lesson from the Japanese, the practice of quantitative easing should be implemented by the Federal Reserve along with actions aimed at reducing unemployment to prevent expectations of deflation from developing in the United States.

According to the data extracted from the St. Louis Federal Reserve Bank website, expected inflation experienced a sharp decline in the middle of 2008 but since December 2008 it has been rising (see Appendix 10 and 11). This fact implies that the Federal Reserve has been able to keep inflation expectations well anchored by affectively implementing the practice of quantitative easing and, therefore, increasing consumers' confidence.

CONCLUSION

At times of economic downturns like the one taking place in the United States today, it is useful to look back at similar situations and extract some lessons in terms of policy-making in order to avoid similar mistakes and take advantage of already established ways for recovering from a recession.

After observing the causes of the two crises, drawing some parallels between the two, and looking at the monetary and fiscal policies as well as some of the financial system reforms made in Japan, I made an analysis of the policies that could be applied to the United States scenario. I found that such policies as increasing government spending through a number of stimulus packages and increasing the central bank's purchases of government debt and setting certain levels for expected inflation to stimulate lending and consumer spending could and should be applied to the current economic downturn in the U.S. However, the strategy of cutting interest rates to nearly zero might not prove to be effective because of the aging population of baby boomers and fear of uncertain future which all contribute to saving more even at a zero nominal interest rate.

In conclusion, there are lessons to be learned from the Japanese experience in the 1990s for the American macroeconomic policymakers. The first lesson discussed in the paper was lowering short-term interest rates that proved to be ineffective in Japan but might be applicable to the American situation. Cutting interest rates in the United States might be less effective as well as in Japan because of the uncertainty in the future and increasing aging population. There is one mistake to learn from the Japanese crisis and that is not to be slow in cutting interest rates which might reduce its effectiveness. The second lesson was to stimulate government spending through the use of a number of stimulus packages. The biggest mistake the American policymakers should keep in mind is not to stop stimulating spending too soon which proved to worsen the situation in Japan. The final lesson to be taken out of the Japanese crisis of the 1990s is to confront falling expected inflation. As discussed in this paper, the Federal Reserve should keep inflation expectations well anchored by implementing the practice of quantitative easing and taking actions aimed at reducing unemployment.

ENDNOTES

1. Japan's real GDP growth over that period of time reached 100.7%. At the same time, real GDP growth in UK was 48.7%, France 63.4%, Germany 55.7%, Italy 72.1%, Canada 78.8% and US 73.3%. The data was found from Robert J. Gordon "Macroeconomics" textbook (eleventh edition), Appendices A and B.
2. The information on Japanese stock performance was found on the Yahoo Finance website <www.finance.yahoo.com>.
3. These statistics were extracted from the IMF Working Paper (2000) "The Japanese Banking Crisis of the 1990s: Sources and Lessons" written by Akihiro Kanaya and David Woo.
4. The statistics were extracted from Robert Gordon's "Macroeconomics" textbook, Appendix B.
5. These data were obtained from the Bank of Japan official website that is available at <http://www.stat-search.boj.or.jp/ssi/mtshtml/m_en.html>.
6. These data were obtained from the Bank of Japan official website that is available at <http://www.stat-search.boj.or.jp/ssi/mtshtml/m_en.html>.
7. The statistics were extracted from Robert Gordon's "Macroeconomics" textbook, Appendix B.
8. The information on the U.S. stock performance was found on the Yahoo Finance website <www.finance.yahoo.com>.
9. The data are available from the online source by Robert Shiller.
10. The data are available on the St. Louis Federal Reserve Bank website at <http://research.stlouisfed.org/fred2/series/GDPCA?cid=106>.
11. The data are available on the St. Louis Federal Reserve Bank website at <http://research.stlouisfed.org/fred2/series/CPILFESL/downloaddata?cid=9http://research. stlouisfed.org/fred2/series/CPILFESL/downloaddata?cid=9>.
12. The data are available on the St. Louis Federal Reserve Bank website at <http://research.stlouisfed.org/fred2/series/MICH?cid=98>.
13. The data are available on the St. Louis Federal Reserve Bank website at <http://research.stlouisfed.org/fred2/categories/115>.

REFERENCES

Ahearne, Alan, Joseph Gagnon, Jane Haltimaier, Steve Kamin, Christopher Erceg, Jon Faust, Luca Guerrieri, Carter Hemphill, Linda Kole, Jennifer Roush, John Rogers, Nathan Sheets, and Jonathan Wright. 2002. "Preventing Deflation: Lessons from Japan's Experience in the 1990s." *International Finance Discussion Papers*, 2002 – 729. Washington: Board of Governors of the Federal Reserve System, June.

Bank of Japan. 1997. “Forward by the Governor.” *Annual Review,1997.* August 29; available at <http://www.boj.or.jp/en/type/press/koen/kako/ko9708a.htm>.

Browne, Lynn E. 2001. “Does Japan Offer Any Lessons for the United States?” *New England Economic Review*. Issue #3.

The Economist. 2009. “Big Government Fights Back.” *The Economist***.** 29 January, 2009. Available at <http://www.economist.com/finance/displaystory.cfm?story_id=13035552#top>.

The Economist. 2009. Economics A-Z. *The Economist*. 2009.; available at <http://www.economist.com/research/Economics/>

“Economic Stimulus**.”** *The New York Times* 6 May, 2009. Available at <http://topics.nytimes.com/topics/reference/timestopics/subjects/u/united_states_economy/economic_stimulus/index.html?scp=1&sq=stimulus%20plan%202009&st=cse>

Fackler Martin. “Japan’s Big-Works Stimulus is a Lesson.” *The New York Times.* 5 February, 2009. Available at <http://www.nytimes.com/2009/02/06/world/asia/06japan.html?_r=2&ref=todayspaper>.

Federal Reserve Bank of St. Louis, Economic Research. Online data is available at <http://research.stlouisfed.org/>

Gordon, Robert. 2009. *Macroeconomics*. 11th edition; Pearson Education, Inc.

Hoshi, Takeo and Anil K. Kashyap. 2004. “Japan’s Financial Crisis and Economic Stagnation.” *Journal of Economic Perspectives.* Winter, 18(1): 3-26.

Kanaya, Akihiro and David Woo**.** 2000. “The Japanese Banking Crisis of the 1990s: Sources and Lessons.” IMF Working Paper.

Krugman, Paul. 1999. *The Return of Depression Economics*. New York: W.W. Norton & Company.

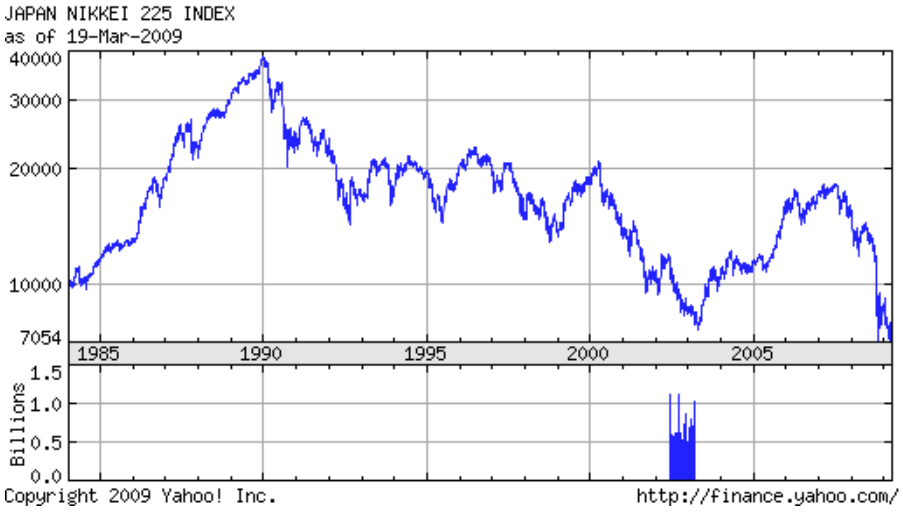
Matsushita, Yasuo. 1997. “Recent Monetary and Economic Conditions in Japan and the Reform of the Financial Markets.” April 14; available at <http://www.boj.or.jp/en/theme/public/koen/tainin/governor/matsushita/index.htm>.

Shiller, Robert. Online Data available at <http://www.econ.yale.edu/~shiller/data.htm>.

Williams, John C. 2009. “The Risk of Deflation.” *FRBSF Economic Letter*. Number 2009-12.

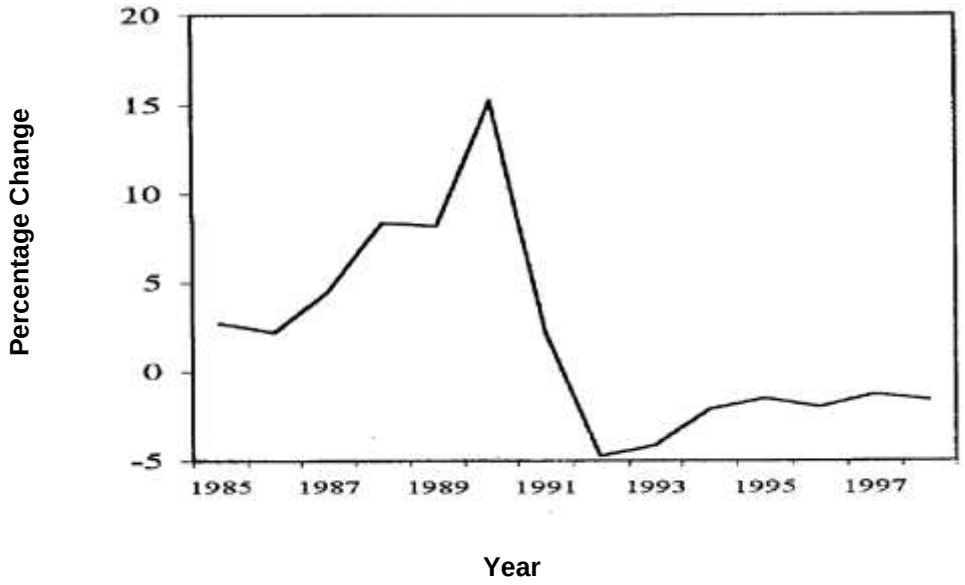
APPENDIX 1[2]

Japan Nikkei 225 Index

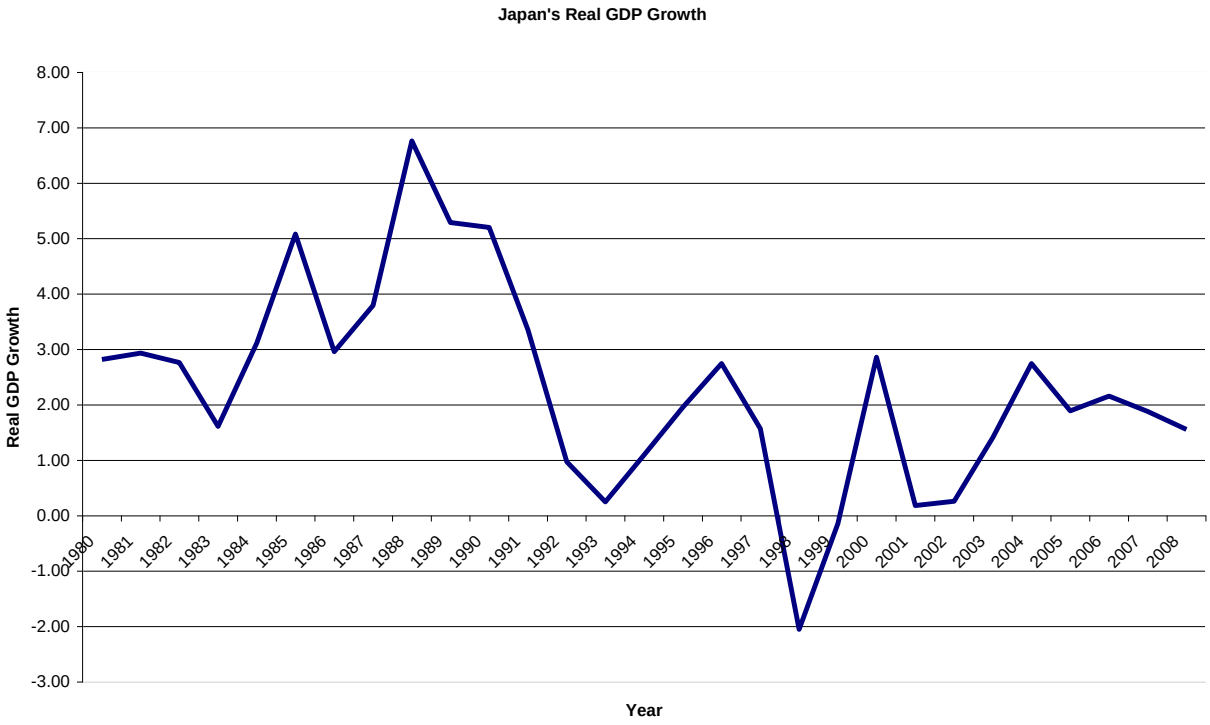


APPENDIX 2[3]

Japan's Growth in Residential Land Prices



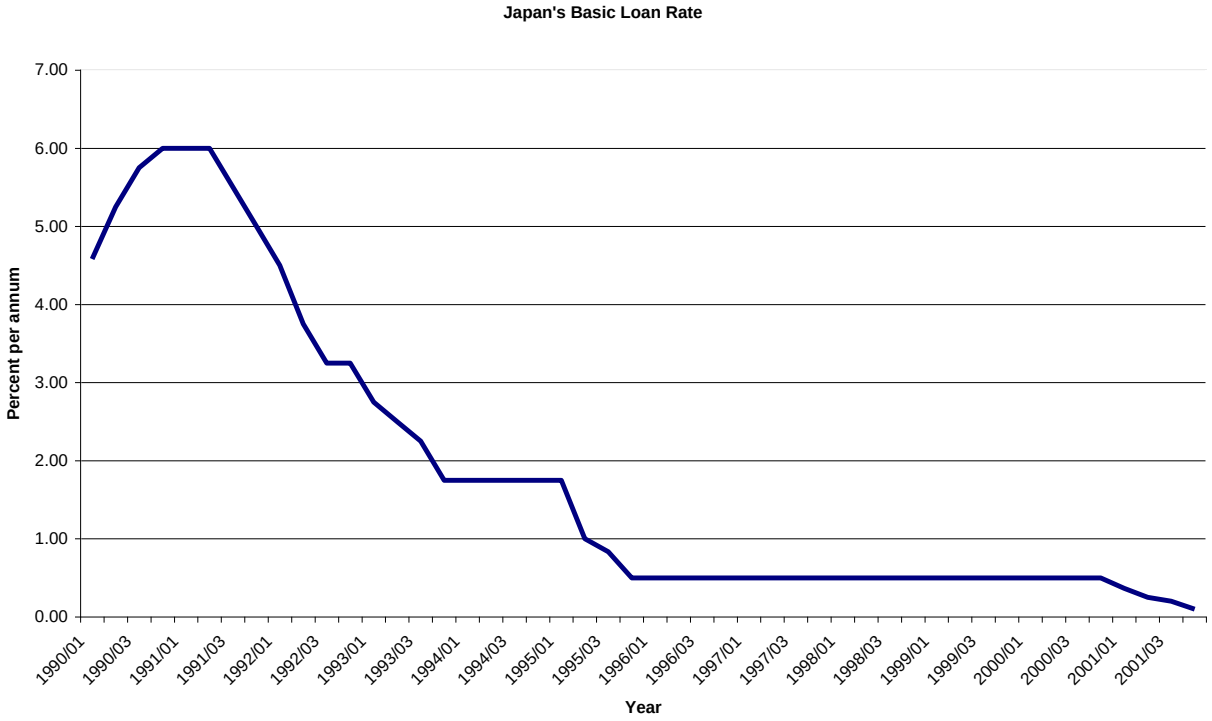
APPENDIX 3[4]


Japan's Real GDP Growth
Real GDP Growth
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
Year

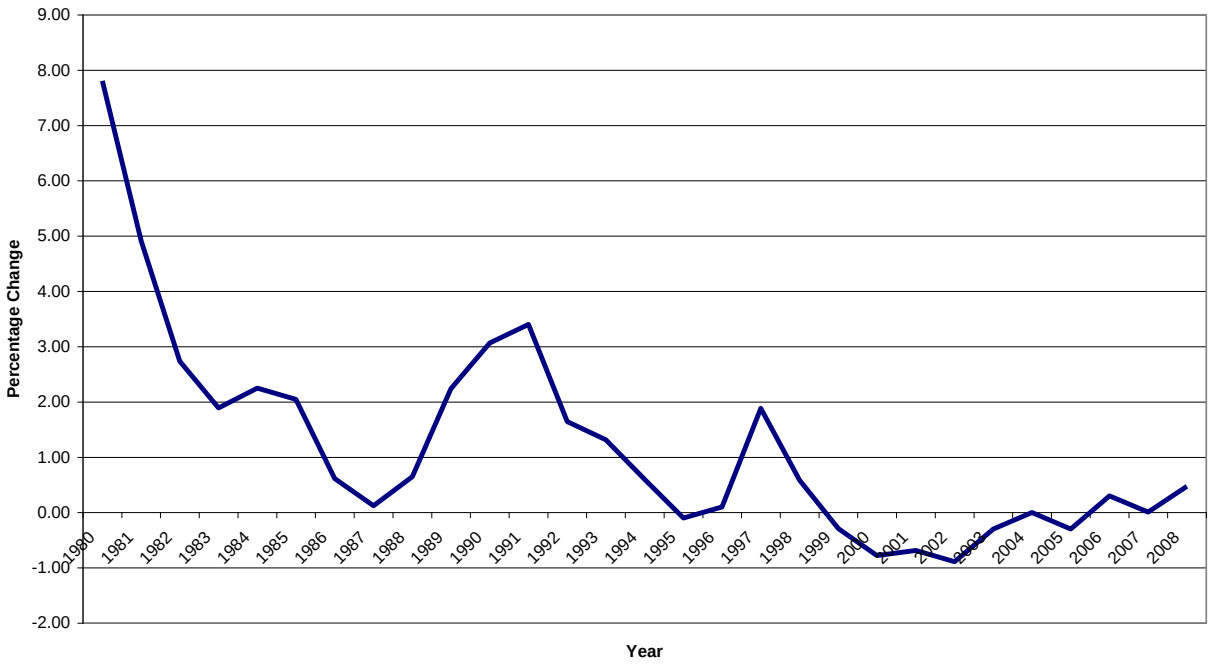
APPENDIX 4[5]
TABLE 1
Japan's Basic Loan Rate, 1990-2001

Year	Basic Loan Rate	Year	Basic Loan Rate
1990/01	4.58	1996/01	0.50
1990/02	5.25	1996/02	0.50
1990/03	5.75	1996/03	0.50
1990/04	6.00	1996/04	0.50
1991/01	6.00	1997/01	0.50
1991/02	6.00	1997/02	0.50
1991/03	5.50	1997/03	0.50
1991/04	5.00	1997/04	0.50
1992/01	4.50	1998/01	0.50
1992/02	3.75	1998/02	0.50
1992/03	3.25	1998/03	0.50
1992/04	3.25	1998/04	0.50
1993/01	2.75	1999/01	0.50
1993/02	2.50	1999/02	0.50
1993/03	2.25	1999/03	0.50
1993/04	1.75	1999/04	0.50
1994/01	1.75	2000/01	0.50
1994/02	1.75	2000/02	0.50
1994/03	1.75	2000/03	0.50
1994/04	1.75	2000/04	0.50
1995/01	1.75	2001/01	0.37
1995/02	1.00	2001/02	0.25
1995/03	0.83	2001/03	0.20
1995/04	0.50	2001/04	0.10

APPENDIX 4.1.[6]

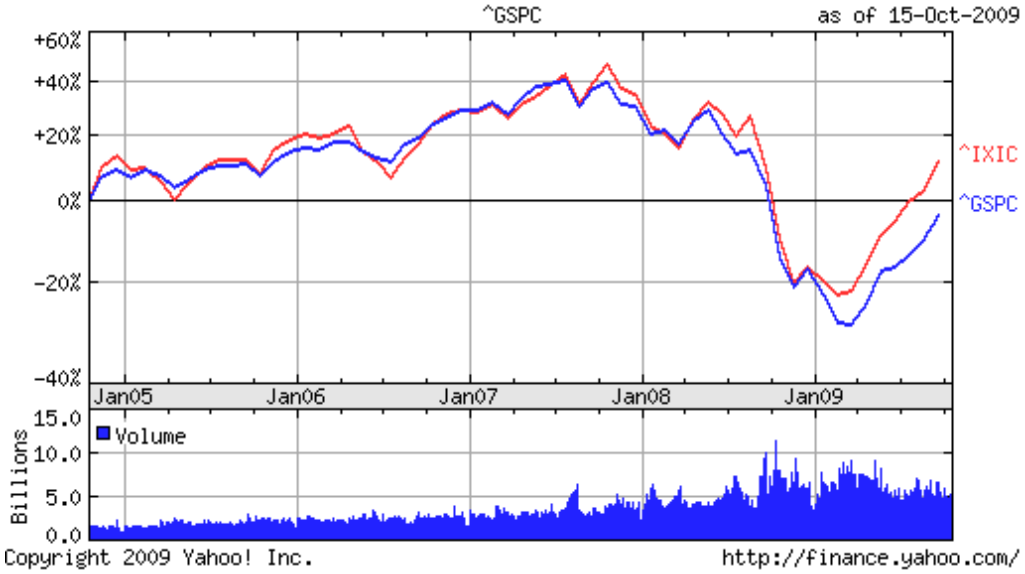


APPENDIX 5[7]



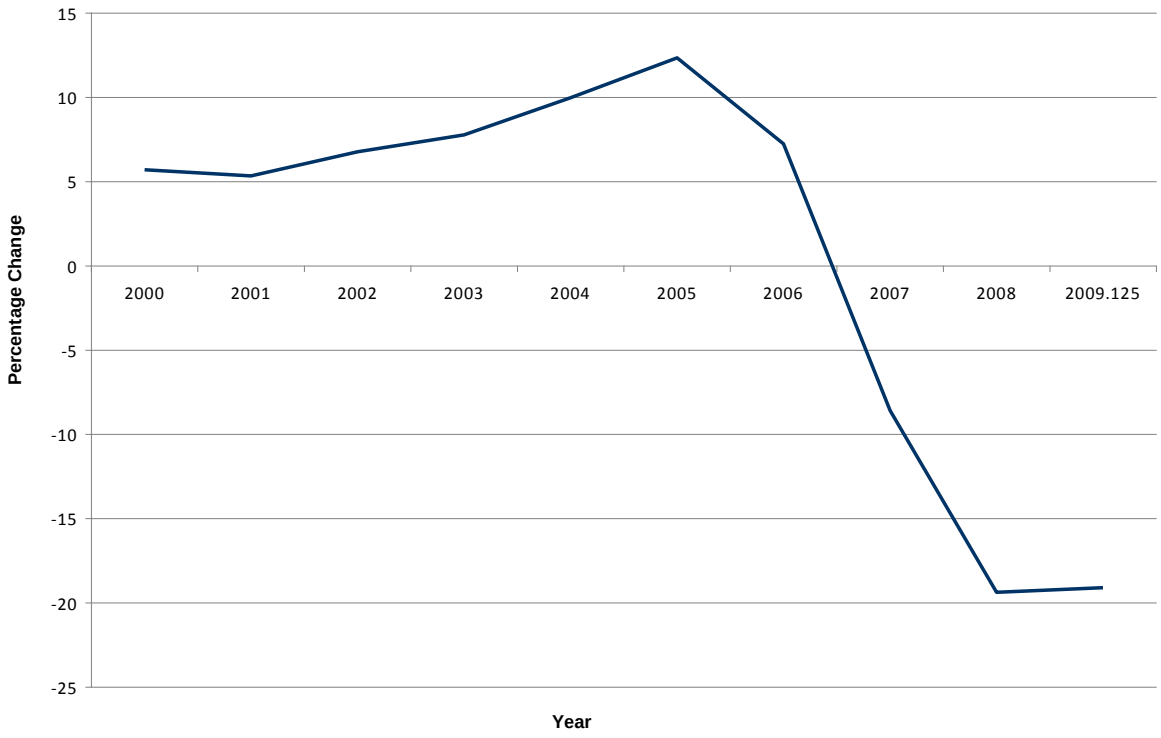
APPENDIX 6[8]

Comparison of the U.S. S&P 500 (GSPC) and NASDAQ (IXIC) Indices



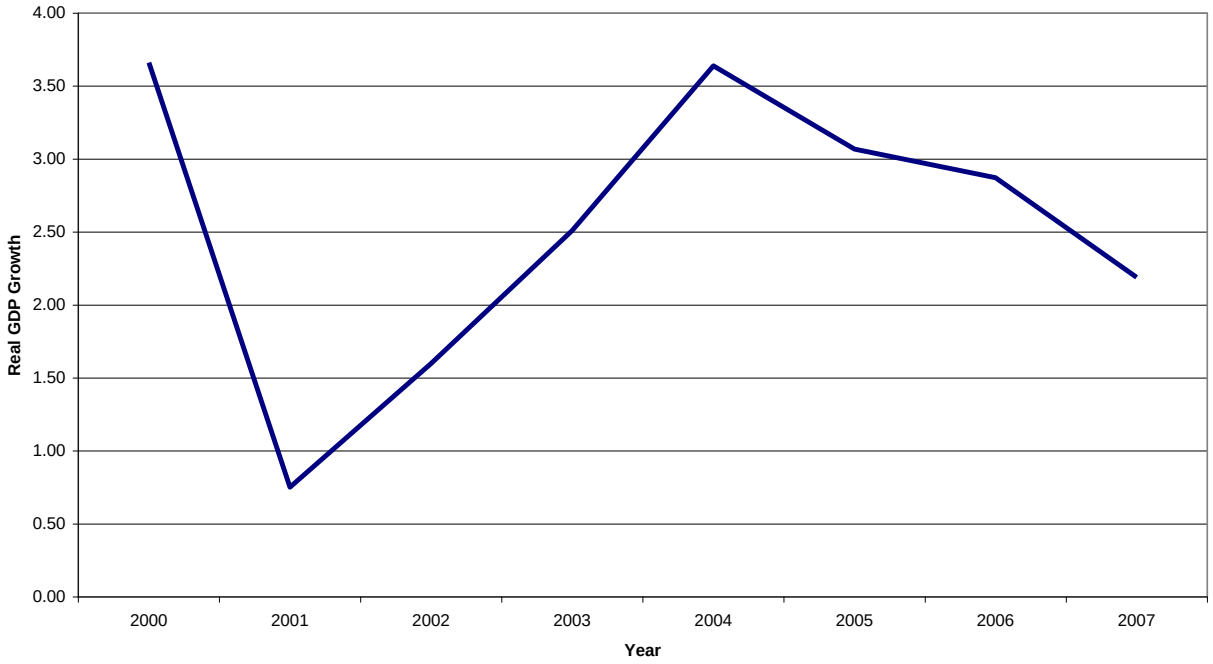
APPENDIX 7[9]

The U.S. Real Home Price Index Percentage Change



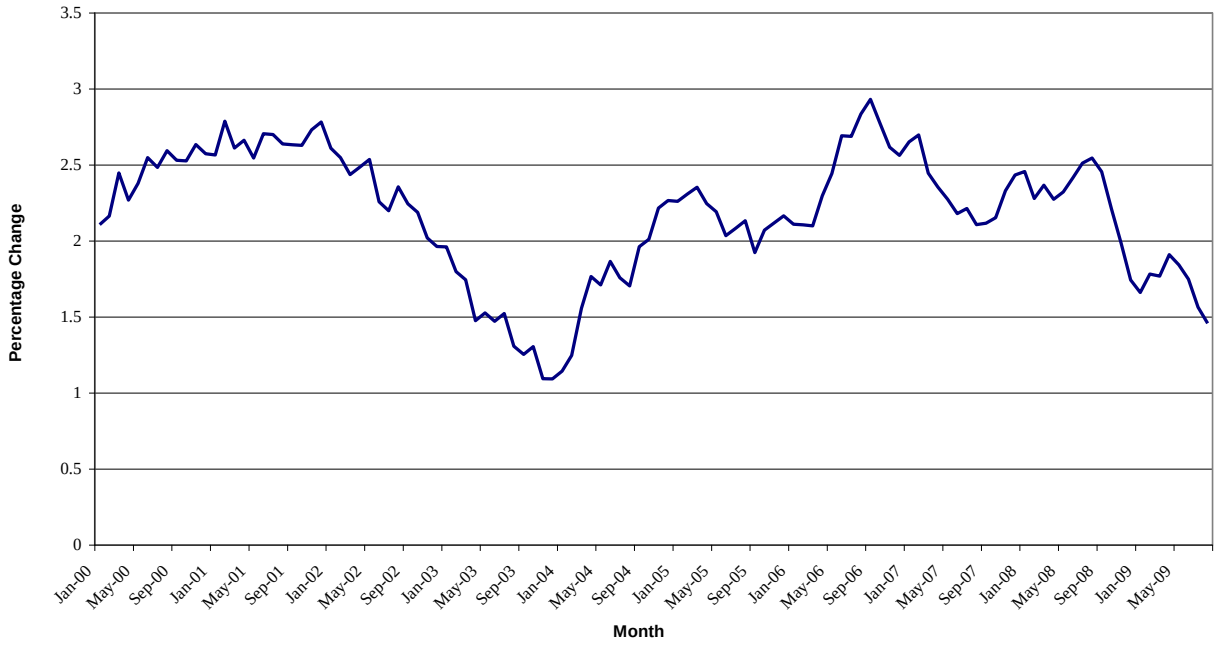
APPENDIX 8[9]

The U.S. Real GDP Growth

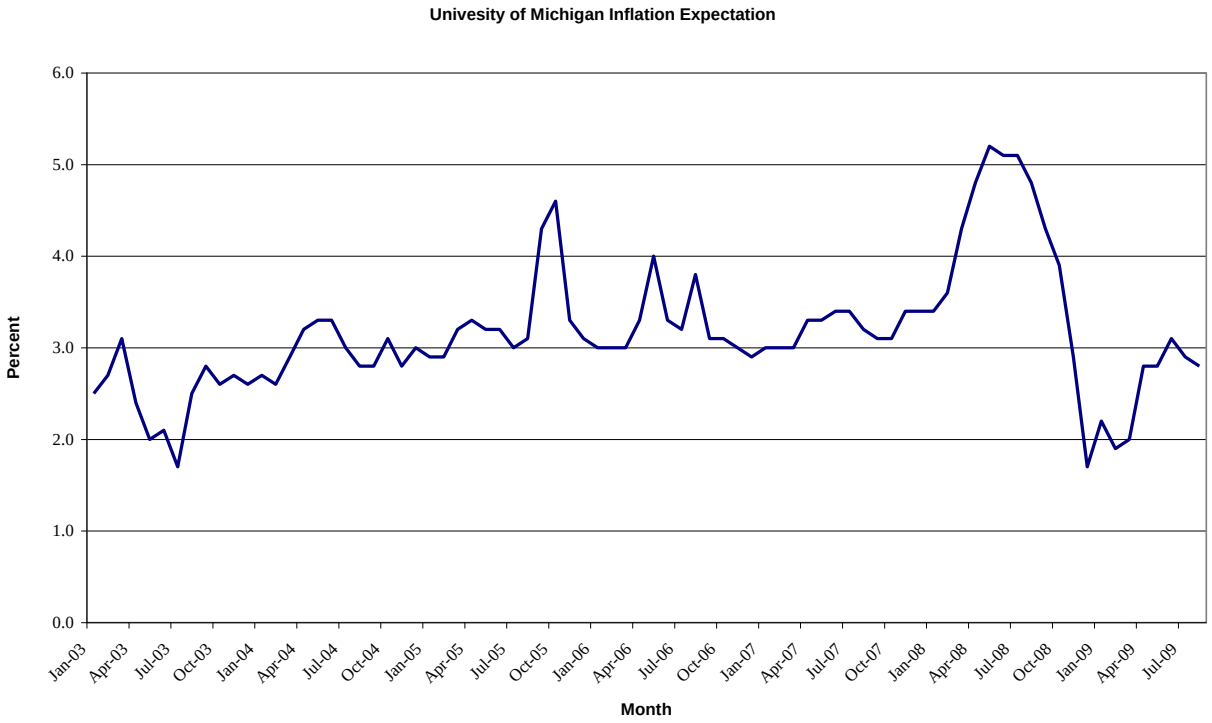


APPENDIX 9[11]

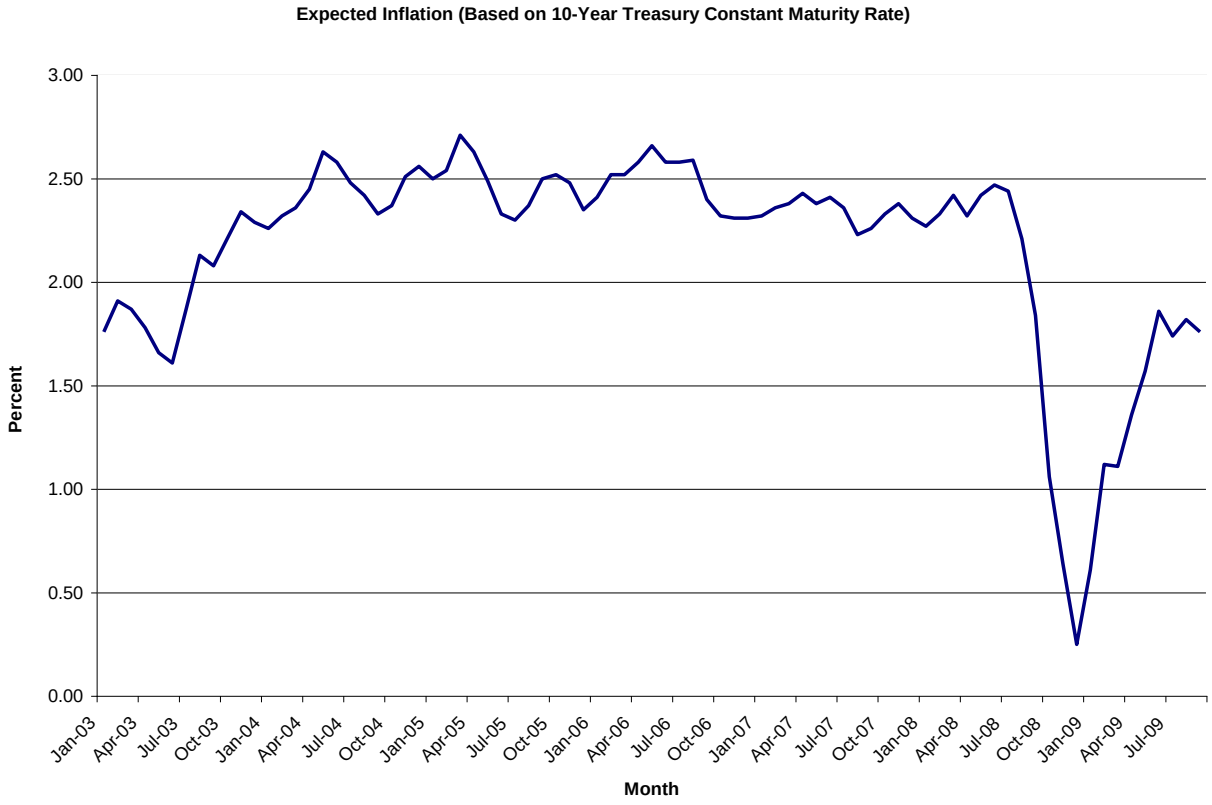
The U.S. CPI (All Items Less Food and Energy) Percentage Change



APPENDIX 10[12]



APPENDIX 11[13]



Fecundity and Husband-Wife Age Gap At First Marriage: Cross Country Analysis

Xu Zhang*

ABSTRACT

One explanation given for the gender wage gap is the division of labor in the home. On the marriage outset, averagely speaking, husbands who are older invest more human capital in market activities than wives. This paper concentrates on the age gap at first marriage. It hypothesizes that the demand for children is one important reason for the husband-wife age gap. To show this, the paper extends a two-sided matching model of marriage (originally introduced by Eugenio Giolito (2003)) based on the biological fact that men have a longer fecundity horizon than women and examines cross-country data from World Marriage Patterns 2000 to show a higher marital age gap the greater the role children play in economic activities, the greater the importance of agriculture in an economy, the greater a country's rural compared to total population, and the greater the population's proportion of pro-natal religions.

1. INTRODUCTION

The prevalent age structure when a marriage union forms is that older men tend to marry younger women. However, the positive difference of SMAM (Singulate Mean Age at Marriage)[1] between husband and wife varies across countries. According to World Marriage Pattern Report 2000, released by United Nations, the significant age difference of SMAM varies between countries and regions. In less developed countries and regions, the age difference of SMAM tends to be larger, while in relatively more developed areas, the age difference of SMAM tends to be smaller.

A large body of research on the timing of marriage suggests economic growth, availability of potential candidates for marriage, culture differences and social norms can all be possible explanations for the rising age at marriage for both males and females. However, most of the previous studies on timing of marriage pattern concentrate on one or two countries or part of regions. Saardchom and Lemaire (2005) perform an international regression analysis on the timing of marriage pattern, but their explanation is based on several theories and is not attributable to a specific universal reason. In this paper, I concentrate on the age gap at first marriage. Similar to the model in Giolito (2003), my model also assumes asymmetric fecundity horizon between male and female. But instead of a constant utility multiplier k, I assume the utility from having children within marriage captures information about personal preferences and country-specific characteristics. I perform cross-country regression analysis and conclude the value of children in the marriage is the main driving force for the age gap at first marriage between husbands and wives. In another word, the more children within marriage are desired, the larger

* Department of History, Economics & Politics, SUNY-Farmingdale, 2350 Broadhollow Road, Farmingdale, NY 11735. Xu.Zhang@farmingdale.edu

age gap at first marriage between male and female. Any country-specific characteristic affecting people's value for children within the country will drive the age gap at first marriage flowing in magnitude.

2. MARRIAGE AND FERTILITY PATTERNS AROUND THE WORLD

Africa is a continent with the highest average age gap at first marriage. It also has one of the highest fertility levels in the world. The average SMAM difference between males and females is around 5 years, ranging from 9.2 years in Gambia to 1.8 years in South Africa. Also, females in Africa tend to marry early. Of the 20 countries with the highest age gaps in mean age at first marriage, 16 countries are in Africa. Though the overall fertility level is decreasing worldwide, there are a number of countries experiencing meager declines. These latter countries constitute 21 developing countries where fertility rates declined by less than one child since 1970. Of these 13 countries are from sub-Saharan Africa. The large age gap at first marriage and the high fertility level in Africa are basically associated with low healthcare conditions and fewer family planning policy controls.

The average SMAM difference between males and females in Asia is around 3 years. The age gap at first marriage varies dramatically with a minimum of around 1 year in Myannar and a maximum of around 7 years in Afghanistan and Bangladesh. Developed countries such as Japan, Korea and Singapore are experiencing low fertility level due to rapid growth in their economies and religion beliefs that don't promote fertility. China has reduced her fertility level by 4 children per women since 1970 by taking on family planning policies.

Europe is the continent with lowest SMAM difference between males and females. The data shows a minimum of -0.2 years in San Marino (meaning that in San Marino wives are actually on average 0.2 years older than their husbands) and a maximum of 4.9 in Greece. North Europe is an area with a low age gap at first marriage and high social welfare.

The SMAM difference is smaller in Latin America and the Caribbean than in Africa or Asia. Similar to the North European countries, there are many consensual unions in this area, therefore, the data may not exactly reflect the age at first marriage.

3. DATA SOURCES AND DEFINITIONS

The age gap at first marriage is measured by the difference of *singulate mean age at marriage* (SMAM) between males and females. SMAM is a formulation developed by Hajnal (1953) to indirectly compute a population's mean age at first marriage from census type data on the proportion of a population's single people at each age level (assuming all first marriages have taken place by age 50). It is expressed as

$$SMAM = \frac{1}{1 - p_{50}} \sum_{x=0}^{50} (p_x - p_{50}) .$$

Thus SMAM is the sum, up to age 50, of the difference between the proportion single at age x and the proportion single at age 50 divided by 1.0 minus the proportion single at age 50. Intuitively, this is the weighted average of the ages at which individuals get married up to age 50.

Fertility level is a proxy for the value of children. A greater number of children in a family indicate children are relatively more desired and valuable. In our model more children are associated with a larger age gap at marriage. Thus, I expect a positive correlation between fertility level and the age gap at first marriage.

Child labor also denotes the value of children. According to a report from UNICEF, an estimated 146 million children are engaged in child labor. The Asian and Pacific Regions have the largest number of child workers in the 5-14 age group, with over 19% of child labor in the world, 127.3 million in total. Sub-Saharan Africa has about 48 million child workers. Latin America and the Caribbean countries have almost 17.4 million child workers. Other regions with large volumes of child labor include the Middle East, North Africa and some areas in industrialized and transition economics. In these regions children are desired in the sense that they either earn money to increase family income or provide paid or unpaid housework help at home or in agriculture. Therefore, I expect countries with prevalent child labor to have a larger age gap at marriage.

The difference in the value of children is also reflected by area of residence. Rural areas often motivate families to have more children since rural areas are usually agriculture-related so that parents can get help from children. Families in urban areas have fewer children. Thus I expect smaller marital age gap in urban compared to rural areas.

Also, with advancing medical technology, having children early in life has become less crucial in countries where better health care increased the female fecundity horizon. In such countries, the asymmetric gap in male and female fecundity horizons has becomes smaller, thereby narrowing the relative biological advantage young women have compared to older women. This narrowing of the gender fecundity horizon causes men to be less choosy, and as a result, my model suggests the age gap at marriage to be smaller.

Education usually indicates the development of a country. Becker argues highly educated people positively sort due to more mutual understanding and compatibility. Highly educated spouses are usually associated with a lower desire for children. As a result, my model indicates a negative correlation between the prevalence of high education and the marital age gap.

At this point, I simply use ordinary least squares regression to examine how much of age gap at first marriage between couples is explained by the value of children in the family. All data are extracted from the World Bank's *World Development Indicators 2006*, the World Bank's, *World Marriage Pattern 2000*, from United Nations' *World Fact Book*, from the UNESCO Institute for Statistics, and the US CIA website.[2] The variable names and definitions are following:

Child Labor is defined as percentage of children from age 10 to 14 who are involved in economic activities out of children in that age group.

Urban Population is defined as percentage urban population out of total population.

Employment in Agriculture measures percentage of employment in agriculture relative to total employment.

Tertiary Education of Females refers to post-secondary education, including universities, colleges, technical training institutes, community colleges, nursing schools, etc. The data on female tertiary education describes the percentage of females at age 15 or over who complete tertiary education out of total females.

Literacy Rate is based on the UNESCO definition: "Literacy is the ability to identify, understand, interpret, create, communicate and compute, using printed and written materials associated with varying contexts." Higher literacy rates are associated with having gotten more education.

African Country Dummy is designed to display African regional effects. It is defined as one if data are from an African country, zero if not.

Muslim is defined as percentage of Muslim in the total population.

A summary of these variable definitions is listed in Table 2-1 (available upon request).

The data on age gap at first marriage come from World Marriage Pattern 2000. The age gap at first marriage is calculated as the male-female difference in SMAM for 179 countries. Unfortunately the data are not uniformly available for the same year, but instead varies from 1970-1998. However, this is not a major problem since the age gap at first marriage changes slowly from year-to-year. Thus I match each country's age gap at first marriage with corresponding explanatory variables for the appropriate year, and treat the data as a single cross-section. Thus each observation defines the characteristics for one country. All the data are extracted from World Development Indicators 2006 CD-ROM with the exception of the *Tertiary Education of Female* which we obtain from the Institute for Statistics of UNESCO and the data for the *Muslim* variable which we get from the *World Facts Book*. Summary statistics of all variables in the data are shown in Table 2-2 (available upon request).

4. EMPIRICAL RESULTS

The econometric specification for regression is

$$Agegap = \beta_0 + \beta_1 Childlabor + \beta_2 Year + \beta_3 Urban + \beta_4 Employ\_agr + \beta_5 D\_africa + \beta_6 Mus\lim + \beta_7 Tertiary(/\ Literacy) + \nu$$

Or $Agegap_i = \beta_0 + \beta_1 demand\_for\_children + \beta_2 X_i + v_i$

My theory implies in a country where children are more valuable and desired, the age gap at first marriage between male and female tends to be larger. And any factor leading to children will tend to positively affect the age gap while factors decreasing the quantity demand for children will have negative impact on age gap. Therefore, holding other variables constant, I am expecting $\beta_1 > 0$ as the higher proportional child labor exists, the more important role children play in economic activities. Similarly, β_4 is expected to be positive in the sense that inhabitants of African countries tend to have more children

than any other regions in the world. The intuition behind this is the high death rate in Africa due to bad medical condition forces people to have more children to maintain some level of population growth and to deal with activities for a living. I take variable "*Year of census*" into account for the various years associated with each country's data, therefore, β_2 could be negative or positive. $\beta_3 < 0$ and $\beta_4 > 0$ are expected because agriculture industry is labor intensive, people living in urban area are less likely to have as many children as people living in rural area have. Children are more valuable in rural areas. The higher population density in an urban area or the lower the employment coming from agriculture the less the demand for children, and the smaller the age gap at first marriage. Therefore, these two variables somehow capture similar information so I may see multicollinearity in the regression. β_6 is estimated to display the relationship between the prevalence of Muslim Religion and the age gap at first marriage. As we know, people who are Muslim tend to have more children so we expect positive β_6. As for the impact from level of education on the age gap at first marriage, we expect $\beta_7 < 0$ because high education usually is related to fewer children. "Tertiary education of female" and "literacy rate" both measure the average education attainment of a country, so we need to choose one from of two variables.

As shown in Figure 2-1, the *Age Gap at First Marriage* is positively related to *Child Labor* and *Employment in Agriculture* while negatively related to the *Urban Population* and *Literacy Rate*. In order to verify the relationship between *Tertiary* and *Age Gap,* I examine the scatter plot shown in figure 2-2. The graph shows a slightly positive pattern which contradicts with our expectation. However, Figure 2-3 shows a negative relationship between *Age Gap* and *Literacy Rate,* as I expect. By running a simple regression of Age Gap on Tertiary and Literacy Rate separately, as shown in Table 2-3, I find variable *Tertiary* is insignificant while variable *Literacy Rate* is significant. Therefore, I pick Literacy Rate as the variable to measure the average education attainment.

Since the *Urban Population* and *Employment in Agriculture* capture similar information regarding the importance of agriculture and hence fertility, I drop the variable *Employment in Agriculture* to avoid multi-collinearity. The resulting regression results are shown in Table 2-4.

As expected, *Childlabor* and *D_Africa* have significantly positive effects on the *Age Gap*; *Literacy* shows a negative significant coefficient which means a country with higher literacy rate tends to have smaller age gap at first marriage. However, *Year*, *Urban* and *Muslim* are insignificant though of the correct sign. Diagnostics indicates no serious collinearity among the predictors. All the coefficients of VIF are less than 10 and all the coefficients of Tolerance are greater than 0.1. Also the student – t normalized residuals displayed in figure 2-4 do not show much deviation from the normal distribution thereby ensuring the validity of t and F tests.

Figure 2-5 checks the linearity between explanatory variables and the dependent variable *Age Gap*. Here, I find the *Literacy Rate* to show some degree of non-linearity. Therefore, I try *Lsquare* as square of the *Literacy Rate*. We also drop *Urban* from the regression and the result is shown as Table 2-5.

Table 2-4.OLS Regression Results and Collinearity Diagnostics(Model 1)

Age Gap	Coef. (Std. Err.)	p-value	VIF	Tolerance
Childlabor	.0311315 (.0135897)	0.024	4.90	0.2039
Year	-.0383998 (.0240396)	0.113	1.12	0.8890
Urban	.0058541 (.0058972)	0.323	2.25	0.4449
D_Africa	.9619853 (.2758416)	0.001	1.89	0.5283
Muslim	.4576667 (.2870273)	0.113	1.38	0.7229
Literacy	-.0253539 (.0076413)	0.001	3.90	0.2567
_cons	81.01474 (47.65364)	0.092		
R-squared	0.6231			
Adj R-squared	0.6050			

Model 2 has slightly smaller R-squared due to fewer explanatory variables. *Child Labor* is still significant with a coefficient .032143 meaning that one percentage increase of *Child labor* will increase the *Age Gap* by .032143 years. The impact from African Countries is also still significant with a coefficient indicating a .9848378 year difference the in marital age gap between African Countries and the other regions. The effect from *Lsquare* on *Age Gap* is negative implying the marginal effect to be decreasing as the Literacy Rate increases. I interpret this to indicate that the more literate the population, the greater a country's socio-economic development, the more the non-agriculture economic activity, and the lower the desire for children, resulting in a smaller age gap at first marriage. It is more likely that marital partners tend to find a partner with similar education background when the level of literacy rate approaches high levels. As Becker (1991) claims highly educated people tend to be positively sort. Also they are usually associated with a lower desire for children so that according to our model the age gap at first marriage tends to be smaller. The *Muslim* variable still has a positive sign thought he coefficient becomes statistically insignificant. Figure 2-6 plots the regression residuals of Model 2. Again, there is little deviation from the normal distribution.

One of main assumption of OLS regression is homoskedasticity of the variance of residuals. If the model fits the data well, there should be no explicit patterns to the residuals plotted against the fitted values. Figure 2-7 does not show very strong evidence for heteroskedasticity. However, it shows there are likely some influential outliers in the data. I can plot the square of residuals and look for observations that are jointly high on both of the two measures. As shown in Figure 2-8, Afghanistan easily catches our attention.

Table 2-5.OLS Regression Results and Collinearity Diagnostics (Model2)

Age Gap	**Coef. (Std. Err.)**	**p-value**	**VIF**	**Tolerance**
Childlabor	.032143 (.011845)	0.008	3.62	0.2763
D_Africa	.9848378 (.2952317)	0.001	1.91	0.5245
Muslim	.5390995 (.3041157)	0.079	1.38	0.7224
Lsquare	-.0001448 (.0000599)	0.017	3.77	0.2653
_cons	3.848283 (.5518538)	0.000		
R-squared	0.5930			
Adj R-squared	0.5803			

I can check the overall measure of influence of this outlier by looking at the Cook's D statistic which combines information on both leverage and residuals. The higher the Cook's d is, the more influential the data point. I do find the Cook's d for Afghanistan is the larger than for any other observation at .1852593 whereas the cut-off point is .03007519(4/133). For this reason, I run a regression omitting "Afghanistan." The results in Table 2-6 show a significant change in the coefficients.

Table2-6.Regression Results of Model 3&4 (Omitting "Afghanistan")

Age Gap	**Model 3**		**Model 4**	
	Coef. (Std. Err.)	p-value	Coef. (Std. Err.)	p-value
Childlabor	.0333763 (.0115621)	0.005	.0254655 (.0118889)	0.034
D_Africa	1.15948 (.2949033)	0.000	1.293173 (.2959466)	0.000
Muslim	.6983815 (.3022475)	0.022	.6105411 (.2998096)	0.044
Literacy			-.0594371 (.0259942)	0.024
Lsquare	-.0001006 (.0000606)	0.099	.0003156 (.0001915)	0.102
_cons	3.433575 (.5590522)	0.000	5.399208 (1.020521)	0.000
R-squared	0.6014		0.6173	
Adj R-squared	0.5888		0.6021	

In Model 3, *Child Labor* , *D_Africa* and *Muslim* are all significant with correct signs as we predict, while *Lsquare* becomes insignificant. We add Literacy to the regression as shown in Model 4, *Child Labor* , *D_Africa* and *Muslim* are still significant with correct signs. Literacy shows UP negatively and the coefficient is significant, but Lsquare is still insignificant and with a positive though very small coefficient. The marginal effect from Literacy Rate on Age Gap at first marriage becomes:
-0.0594371+2*.0003156*Literacy, which means as literacy increases, the age gap is decreasing, and the marginal effect from Literacy is increasing until literacy rate gets to one. Obviously, there is collinearity problem in Model 4. However, I find no matter how we change the other explanatory variables in the model besides Child Labor, Child Labor always shows strong positive effect on the Age Gap, which is consistent to my prediction that the more valuable of children or the higher demand for children, the larger age gap at first marriage.

ENDNOTES

1. As will be explained later, the SMAM is a population's mean age at first marriage computed via a demographic technique first developed by Hajnal (1953).
2. www.cia.gov

REFERENCES

Becker, Gary S. 1973. "A Theory of Marriage: Part I ", Journal of Political Economy, 81(4): 813-846.
Becker, Gary S. 1991. *A Treatise on the Family*. Cambridge, MA, Harvard University Press.

Bergstrom, Theodore and Bagnoli, Mark. 1993 “Courtship as a Waiting game”, *Journal of Political Economy*, 101(1).

Brien, Michael. 1997. “Racial Differences in Marriage and the Role of Marriage Markets ”, *Journal of Human Resource*.

Burdett, Kenneth and Wright, Randall. 1998. “Two-sided Search with Nontransferable Utility”, *Review of Economic Dynamics*, 1(1).

Gale, David and Shapley, Lloyd. 1962. “College Admission and the Stability of Marriage”, *American Mathematical Monthly*, 69.

Giolito, Eugenio P. 2003. “A Search Model of Marriage with Differential Fecundity”, IZA Bonn Discussion Paper No.1082.

Hajnal, J. 1953. “Age at Marriage and Proportions Marrying”, *Journal of Population Studies*, 17: 115-136.

P.M. Kulkarni. 1995. Asia-Pacific Population Journal, 10(4): pp.53-74

Mortensen, Dale. 1988. “Matching: Finding a Partner for Life or Otherwise”, *American Journal of Sociology* 94.

Saardchom, Narumon and Lemaire, Jean. 2005. “Causes of Increasing Age at Marriage: An International Regression Study”, *Marriage and Family Review*, Vol.3.

Unesco. *World Fact Book.* Unesco Institute of Statistics.

United Nations. 2000. *World Marriage Pattern.*

United States Bureau of the Census. The United States Median Age at First Marriage. Data from 1890-2002.

World Bank. 2006. *World Development Indicators.*

Cumulative Voting and the Tension between Board and Minority Shareholders

Aiwu Zhao and Alex Brehm*

ABSTRACT

The separation of management and ownership has created various agency problems and long-lasting conflicts between the board and minority shareholders. Our empirical study looks into whether cumulative voting can help ease the tensions between these two groups of investors. We use shareholder proposals as an indication of the nature and amount of tensions between the board and minority shareholders. Our research shows that cumulative voting can help ease the tension between the board of directors and minority shareholders and reduce investors' monitoring costs. In some cases, however, the impact of cumulative voting can be constrained by other corporate governance mechanisms.

INTRODUCTION

In the corporate world, people try to maximize their own benefits. Under the principal–agent relationship framework, executive groups should work their best to benefit shareholders. But executives usually hold shares of the company as well and are often members of the board of directors, so it is likely that executives and directors, who are shareholders with direct control of the company, may make decisions that will benefit themselves but not necessarily other shareholders who do not have direct control of the company.

Shareholders who do not have direct control of the company are usually referred to as minority shareholders. Though directors are elected to represent the interest of shareholders, there always seems to be some discordance between minority shareholders and the board of directors. One extreme example is the "just vote no" campaign initiated by some shareholder activists in the 1990s. The idea is that because it is not very cost-benefit efficient to campaign issue-oriented shareholder proposals, these shareholder activists simply urge "just vote no", meaning to withhold votes towards a director's election to express dissatisfaction with management performance or the firm's corporate governance structure, thinking this should be more effective because directors want to avoid the public embarrassment and the associated damage to their individual reputations (Del Guercio, et al. 2008).

Facing the long lasting conflicts between the board and minority shareholders, corporations actually have an alternative choice, which is cumulative voting, to ease the tension between the board and minority shareholders. Cumulative voting allows stockholders to cast all of their available votes in director elections for a single director nominee. Therefore it will enhance the voting power of minority

* Department of Management and Business, Skidmore College, Saratoga Springs, NY 12866. We would like to thank the support of the Skidmore College Summer Faculty/Student Scribner-Mellon Scholar Research Program.

shareholders. Cumulative voting is also more cost-benefit efficient to minority shareholders compared to an issue-oriented campaign as discussed by some legal literature (Pozen, 2003).

However, even though cumulative voting was popular at the beginning of the 20th century, it has experienced a steady decline since the 1950s. In 1984, cumulative voting was mandatory in 18 states. The number reduced to six in 1992 (Maassen and Brown, 2006). According to the Investor Responsibility Research Center (IRRC), in 2003, only 9.2% of the Standard & Poor's Super 1500 have cumulative voting (Gordon, 1994).

The main reason for many states to abandon mandatory cumulative voting in the '80s was because there were numerous hostile takeovers happening during this period of time. Nowadays, U.S. corporations are operating in a much less hostile environment because of the development of investment funds in the past decades. Investment funds have become the dominant shareholders these days. The investment goals of most institutional shareholders are long term and they are not likely to takeover a corporation. However, there was little responding change in the corporate governance structures of the majority of U.S. companies. The corporate bylaws are very protective to controlling shareholders as a result of the hostile takeover wave in the '80s. So we still see many conflicts between the boards and minority shareholders. Adopting cumulative voting, repealing a classified board, and executive compensation issues, have consistently been the top three institutional shareholder proposals (Campbell et al., 1999; Gillan et al., 2007). All these shareholder proposals are associated with a costly solicitation process. Since cumulative voting will increase the chance for minority shareholders to elect some board members who can better represent their interests, with the change of minority shareholder structure, it is necessary to examine whether the corporate world needs to reconsider the adoption of cumulative voting.

Our paper tries to examine whether cumulative voting plays a role to ease the tension between board of directors and minority shareholders. Previous research on the impact of cumulative voting usually focused on the market reaction to the changes of adoption of cumulative voting. For example, Bhagat and Brickley (1984) test the stock returns around the time of the announcement of management-sponsored proposals that include amendments of cumulative voting. They find that the announcement of proposed amendments that reduce the effect of cumulative voting is associated with abnormally negative returns. In this paper we use the voting results of investor proposals to measure the conflict between board of directors and minority shareholders. First, we argue that market movements can be affected by multiple factors and it is not easy to ascertain the impact of a single factor. For example, in Bhagat and Brickley's (1984) research, they have not observed any significant market reaction when proposals that enhance cumulative voting are announced. So we consider voting results for investor proposals a better and more direct measurement to reflect the conflicts between controlling and non-controlling shareholders. Second, we consider the costs related to investor campaigns and proxy solicitation processes a reduction to investors' wealth as well and these costs should be counted when examine the value impact of cumulative voting. Previous empirical studies have not paid enough attention to these costs, so our study provides new evidence on the value impact of cumulative voting.

We look at 201 companies that have investor proposals in 2008 and find that companies with a classified board and more block holders tend to show more conflicts between controlling and non-controlling shareholders. A regression test shows a negative relationship between the presence of cumulative voting and the conflicts between the board and minority shareholders. These results indicate that cumulative voting is an efficient approach to ease the tension between the board of directors and minority shareholders and can help reduce investors' monitoring costs on a corporation.

This paper proceeds as follows. In the next section, we review the related literature on the value impact of cumulative voting. In Section III, we present the methodology applied in our study, and report the empirical evidence as well. Section IV concludes the paper.

LITERATURE REVIEW

In the corporate world, if shareholders are not satisfied with companies' performance, they have several choices. First, they can just trade their shares – a practice of voting by feet (this is known as the "Wall Street Rule"). Second, they can conduct takeovers and LBOs. Through this choice, investors could fundamentally change corporate control. Third, besides the former two extreme choices, shareholders can vote to influence managers' decisions.

Shareholders have the right to elect directors who will then make major decisions on behalf of investors' interests. Shareholders also have the right to vote on other major issues such as merger decisions, liquidations, and charter amendments. However, in the real world, power of shareholders is seldom equal. For example, Johnson et al (2000) find that tunneling, a phenomenon where controlling shareholders transfer assets or profits out of firms for the benefit of themselves, is very common in both developing and developed countries.

Theoretical discussion from the 1950s (Williams, 1955) suggests that cumulative voting is basically fair, because minority representation under cumulative voting does not break the majority rule and minority groups have the potential power to influence directors by having representation on the board.

But cumulative voting has not been a favored choice of boards and managers. It was seldom used in the greater part of U.S. corporations even several decades ago. The arguments regarding the disadvantages of cumulative voting that are frequently cited by corporate boards usually include: (1) A good director should not be captured by any special interest group; (2) The board of directors should possess mutual confidence and respect; (3) Disharmony could harm the energy of management; (4) Confidential information could be leaked; (5) Shareholders with narrow, selfish interests could abuse cumulative voting.

Bhagat and Brickley (1984) tested the stock returns around the time of the announcement of management-sponsored proposals that include amendments of cumulative voting. They analyze two kinds of amendments. One is amendments that reduce the impact of cumulative voting; the other is amendments that enhance the impact of cumulative voting. They collect samples from 1962 to 1982 and find that the announcement of proposed amendments that reduce the effect of cumulative voting is

associated with abnormally negative returns. The authors find on average that proposals that enhance cumulative voting do not affect firms' values.

Dodd and Warner (1983) examine a sample of firms experiencing proxy contests for seats on their board of directors. They find that dissident shareholders usually fail to obtain a majority of board seats. However, some minority shareholders obtain some seats, via the cumulative voting mechanism, in over half of the sample contests.

The capital market has radically changed for the past several decades after the hostile takeover waves in the 1980s. One of the significant changes is that market funds have grown at an incredible pace. Institutional investors have come to own substantial common stocks of most publicly traded companies. According to a 2008 Institutional Investment Report issued by the Conference Board, in the largest 1,000 U.S. corporations, institutional investors held 76.4% of total stocks by the year end of 2007. In 1987, they only controlled 46.6%. These institutional investors include mutual funds, pension plans, banks and foundations, and insurance companies. Institutional investors controlled $27.1 trillion assets in 2007, but the number was only $2.7 trillion in 1980.

This significant growth in institutional ownership of U.S. publicly traded companies has broad implications for the corporate governance. The main purpose of these investors is to diversify risks and gain steady returns. Fund managers pay more attention to long-term consideration instead of short-term speculation. Moreover, as managers handle huge amounts of investments for each fund, it is not easy for them using the Wall Street Rule, meaning vote with their feet, to deal with their stocks. For example, fund managers usually will not sell stocks when stock price is temporarily undesirable.

As institutional investors hold an increasing share of equity of publicly traded corporations, they recognize their potential ability to influence companies. Moreover, they are becoming activists to submit their own corporate governance proposals at corporations' annual meetings. There is evidence showing that the increase of shareholders' proposals is related to the growth of institutional investors. The Investor Responsibility Research Center has tracked the shareholder proposals filed at U.S. companies since 1973. From 1973 through 2004, more than 15,000 shareholder proposals were filed at U.S. firms on a wide range of social responsibility and governance subjects during the 32 years. Remarkably, approximately 25% were filed just in the last four calendar years, 2001 through 2004.

Gillan and Starks (2007) sum up the number of corporate governance proposals submitted over two periods, one from 1987 to 1994 and the other from 2001 to 2005. They find that there are substantial changes in the sponsors of the corporate governance proposals between the two periods. Specifically, union funds had become a larger force in submitting proposals. They submitted 10% of the proposals during the 1987 to 1994 period but over 60% between 2001 and 2005. In contrast, unaffiliated individuals accounted for 70.4% in earlier period, but only 44.5% in the later period.

With the changes of minority shareholder structure, it is necessary to reconsider the appropriateness of keeping the unequal rights that exist between the board and minority shareholders.

Cumulative voting, as one of the major mechanisms that can help equalize shareholder rights, deserves more research that is based on current empirical evidence.

EMPIRICAL TESTS

We use voting results on shareholder proposals as a proxy for the conflicts between the board and minority shareholders. Data on shareholder proposal voting results are obtained from the US Annual Corporate Governance Review issued by Georgeson Inc., a Computershare company that provides strategic shareholder consulting services. The Georgeson annual review tracks the shareholder proposals voted on in the annual meetings of the Super S&P 1500 firms.

Based on Georgeson's report, 201 companies voted on shareholder proposals in their 2008 annual meeting. Non-parametric Kruskal-Wallis tests on the differences between companies that voted on shareholder proposals and other S&P1500 companies indicate that companies with bigger size, reflected in total assets and total revenue, are more likely to receive shareholder proposals. But there is no difference in performance, in terms of profit margin and yearly stock return, between companies receiving shareholder proposals and companies not receiving shareholder proposals. The average G-index of companies with shareholder proposals is higher than that of companies without shareholder proposals.

Based on the voting results on shareholder proposals tracked by Georgeson Inc. for 2008, some of the proposals received "For" votes for as low as 0.9%, whereas some receive "For" votes as high as 92%. So whether a company receives shareholder proposals is not a very reliable indicator on the conflicts between minority shareholders and the board. Companies with bigger size tend to have more shareholders, which increases the possibility of having a few unsatisfied activist investors who will bring up shareholder proposals. But not every proposal brought up in the annual meeting reflects the consensus of minority shareholders, so we consider the percentage "For" votes received for shareholder proposals a more appropriate measure on the discordance between the minority shareholders and the board. Boards of directors always recommend an "Against" vote for shareholder proposals. So if a shareholder proposal receives a high percentage of "For" votes in annual meeting, it reflects that there is some tension between the minority shareholders and the board.

We consider two major groups of factors that will influence the relationship between the minority shareholders and the board: corporate governance factors and performance factors. For corporate governance factors, besides the adoption of cumulative voting, which is the focus of our study, we also include number of board classes, number of block holders, percentage of stock held by block holders, percentage of stock held by directors and executive as a group, and G-index. For performance factors, we include total assets, total revenue, profit margin, and yearly stock return.

Among all the companies that received shareholder proposals in 2008, we first compare the differences between those companies who use cumulative voting in their director election process and those who use straight voting.

The non-parametric tests do not show statistically significant differences between the two groups of firms; especially we do not observe that companies that adopted cumulative voting tend to have less tension between the minority shareholders and the board. The correlation tests show similar results that the adoption of cumulative voting is not significantly related to any corporate governance choice or any performance feature.

Cumulative voting is just one of the corporate governance mechanisms that affect the relationship between minority shareholders and the board. A company that has adopted cumulative voting still has other options, such as classified board, to restrain the power of minority shareholder. A regression test is then used to determine the contribution of each corporate governance and performance factor. The result is reported in Table 1.

Table 1: OLS Regression testing the impacts of governance and performance factors

Variable	Parameter Estimate	Standard Error	t-Value	Pr > |t|
Intercept	0.207	0.095	2.17	0.0319
cv_code	-0.160**	0.075	-2.15	0.0334
BoardClass	0.028*	0.015	1.9	0.0589
Qblock	0.024**	0.012	1.99	0.0485
Holding_pblock	0.002	0.002	0.86	0.3923
HoldingDirExc	-0.005**	0.003	-1.98	0.0493
at	-1.53E-08	1.20E-07	-0.13	0.8992
revt	-5.60E-07	8.13E-07	-0.69	0.4919
pm	0.252**	0.126	2.01	0.0466
ret_y	-0.085	0.062	-1.37	0.1724
G	0.013	0.009	1.49	0.1384
N	153			
Adj. Rsq.	0.149			

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

The regression results indicate that after controlling the impacts of other factors, the adoption of cumulative voting now displays a negative relationship with the percentage of “For” votes received for shareholder proposals, meaning cumulative voting does have the influence of easing the tension between minority shareholders and the board. At the same time, there are other governance factors showing significant impacts on the shareholder proposal voting results as well. For example, there is a positive relationship between the number of board classes and the proposal voting result, indicating the more classified the board is, the more likely the company is to get more “For” votes for shareholder proposals.

Companies with more block holders tend to have more issues that can not be agreed upon between minority shareholders and the board. On the other hand, increased share holding of directors and executives has the impact of decreasing the discord between the two groups of investors.

On the performance factor side, only profit margin shows a statistically significant influence on shareholder proposal voting results. This result is different from evidence identified in the 1980 and early 1990s that underperforming companies are more likely to receive corporate governance proposals from investors (Romano, 2001).

In the next step, instead of focusing on the explanatory power of each single factor, we use an F-test to look at the explanatory power of governance and performance factors as groups.

Table 2: Explanatory power of governance and performance factors

Reduced Model	F-test	
	F-value	**Pr>F**
Reduced model for governance factors	4.55	0.0003
Reduced model for performance factors	1.59	0.18

When we drop the governance factors, the reduced model has an F=4.55, Pr.>F=0.0003; when we drop the governance factors, the reduced model has an F=1.59, Pr.>F=0.18, indicating the level of unsatisfactory of minority shareholders are more driven by the corporate governance structure of the company rather than its performance.

CONCLUSION

The main question we want to address through this research is whether cumulative voting can ease the tension between minority shareholders and the board of directors. Previous studies usually evaluate the impacts of the corporate agency problems by examining how firm value is influenced by various corporate governance structures. Our study, on the other hand, pays more attention to minority shareholders' monitoring costs.

Our research results show that discord between minority shareholders and the board of directors is mostly affected by corporate governance structure rather than company performance. This empirical evidence indicates that minority shareholders not only care about business performance of the corporation, they care about the fairness in benefit distribution as well. This is especially true for companies with more long-term minority shareholders, i.e. block holders.

We have also looked whether cumulative voting still plays an important role in corporate governance in today's business world. Cumulative voting attracted much attention in both business and academia in the '80s and early '90s. It became a less popular topic in the latest decade when more and more companies abandoned cumulative voting.

The recent credit crunch has demonstrated that minority shareholders may need to pay more attention to the policies adopted by the board of directors. The SEC is considering enforcing a proxy access rule, a process that could let certain dissenting shareholders nominate directors opposed to management without paying for the mailing and publicity costs from their own pocket as occurs in the current system. Our study finds that cumulative voting does have an impact in alleviating the confrontation between dissenting shareholders and the board of directors. This result indicates that cumulative voting is still an effective mechanism that can lower investors' monitoring cost on boards of directors.

Looking back to history, most of the companies chose to abandon cumulative voting where minority shareholders were mainly individuals. Now with the change of minority shareholder structure, it is time for regulators and companies with major minority shareholders to reconsider the adoption of cumulative voting. Most of the U.S. companies abandoned cumulative voting during the 1980s. There are few cases of adoption of cumulative voting after that period. Considering the unbalanced power between the boards and minority shareholders, it is hard to re-enact a mechanism that is not welcomed by the boards. As a result, there is no recent empirical evidence on the impact of adopting cumulative voting on stock prices. Bhagat and Brickley's positive evidence in 1984 can not be used as a sufficient evidence to claim that adopting cumulative voting will increase firm value under in the current and future business world. But considering the huge amount of resources used in shareholder campaigns, our research results show that it is an efficient choice if some companies could adopt cumulative voting. It will not only reduce investment costs for many shareholders, but will reduce the agency costs in the corporate world as well.

REFERENCE

Bhagat, Sanjai, Brickley, James A. 1984. "Cumulative Voting: The Value of Minority Shareholder Voting Rights." *Journal of Law and Economics*, 27(2): 339-365.

Campbell, Cynthia J., Gillan, Stuart L., and Niden, Cathy M. 1999. "Current perspectives on shareholder proposals: Lessons from the 1997 proxy season." *Financial Management*, 28(1): 89-98.

Del Guercio, D., Seery, L., & Woidtke, T. 2008. "Do boards pay attention when institutional investor activists "just vote no"?". *Journal of Financial Economics*, *90*(1): 84-103.

Dodd, Peter, Warner, Jerold B. 1983. "On corporate governance: A study of proxy contests." *Journal of Financial Economics*, 11: 401-438.

Gillan, Sturart L, Starks, Laur T. 2007."The evolution of shareholder activism in the United States." *Journal of Applied Corporate Finance*. 19(1): 55-62.

Gordon, Jeffrey N. 1994. "Institutions as Relational Investors: A New Look at Cumulative Voting." *Columbia Law Review*, 94(1): 124-192.

Maassen, G.F., and Brown D. 2006. *The Effectiveness of Shareholders' Meetings: An Overview of Recent Developments, in International Corporate Governance After Sarbanes-Oxley,* Paul Ali and Greg N. Gregoriou (Editors).

Pozen, Robert C. 2003. “Institutional Perspective on Shareholder Nominations of Corporate Director.” *The Business Lawyer*, 59 (1): 95-118.

Romano, R., 2001. “Less is more: making institutional investor activism a valuable mechanism of corporate governance.” *Yale Journal on Regulation* 18: 174–251.

Williams, Charles M. 1955. “Cumulative Voting.” *Harvard Business Review*, May/Jun55, Vol. 33(3): 108-114.

An In-class Pricing Game

J. Patrick Meister*

ABSTRACT

An in-class price competition is discussed and analyzed. Students are distributed into five firms that choose their prices in a framework in which they know the demand and cost conditions facing them and their rivals. Students learn about trying to maximize profit in an oligopolistic environment with differentiated products. The game, and subsequent discussion and analysis can be used for reference when covering standard theoretical principles of price competition. Furthermore, the instructor can use insights from the game to help illustrate pricing practices seen in real industry.

I. INTRODUCTION

Many Microeconomics and Industrial Organization students learn about price competition models such as the classic Bertrand duopoly model, the Hotelling linear city model, and other models of price competition in the course work. For examples in Industrial Organization, see Pepall, Richards and Norman (2008). However, students can gain deeper understanding and ability to apply the models if they experience price competition and analyze their experience in a simulated market setting. The game in this paper was developed to help students achieve this. After students have experience with this game, the instructor can readily compare course concepts to aspects of the game that students have played. When students play this game and analyze it, they tend to have a better understanding of pricing principles in general. For a game somewhat related to this, see Eckalbar (2002). However, Eckalbar only has two firms, and players choose prices based on a payoff matrix. For a quantity competition game similar to the one in this paper, see Meister (1999).

II. SETUP AND INSTRUCTIONS

Students are to read the following set of instructions for the game.

* Department of Economics, 319 Muller, Ithaca College, Ithaca, NY, 14850. Ph: 607-274-3883, E-mail: pmeister@ithaca.edu

◊ I would like to thank Richard Vogel for his careful reading and thoughtful comments on this work.

OLIGOPOLY PRICING GAME (INSTRUCTIONS)

You and your teammates are going to run a firm in an oligopolistic industry with five firms. In your industry, you produce a product differentiated from those of your competitors. Your decision is what price to charge for your product during each time period. The price you charge, the prices your competitors charge, and the demand curve facing your firm determines how much output you sell. Thus, your profit will depend not only on your own price, but also on the price levels chosen by your rivals.

You have the same production technology as your rivals (i.e., the same constant Marginal Cost and the same level of Total Fixed Cost). This information will be common knowledge among all firms in the industry. You will not know the price levels chosen by your competitors in a given round until you see all the results from that round. However, you will be able to see your competitors' pricing histories as the game progresses. Assuming you are firm *i*, you have the following things to consider.

Your price:	p_i
Rivals' price levels:	$p_j \qquad i \neq j$
Demand facing firm *i*:	$q_i = \max \{230 - 8p_i + \sum_{j \neq i} p_j \ , 0 \}$
Constant MC:	$c = \$20$
Total Variable Cost for *i* :	$TVC = c\, q_i = 20q_i$
Total Fixed Cost for all *i* :	$TFC = \$300$
Profit for firm *i* :	$p_i = TR_i - TVC_i - TFC_i$
	$= P_i q_i - 20q_i - 300 = (p_i - 20)q_i - 300$

Additional Comments

Using the information above, one should note that if all five firms charge a price of $22, all firms would have negative profit. Also, if all firms charge a price of $56, all firms would have a negative profit.

I will divide each team's profit by a pre-determined number, and that is how many extra credit points each of your team members will earn! Extra credit will be added after the grade scale is determined. You will not be told when the game ends.

If a firm (or firms) price such that quantity demanded is zero, I will impute price(s) to the firm(s) selling 0 such that the imputed price would cause the firm's (or firms') quantity (or quantities) demanded to be exactly 0. Thus, if Z is the set of firms pricing such that they sell 0, I will impute price(s) to the firm(s) of

$$\bar{p} = \frac{230 + \sum_{j \neq i} p_j}{8 - (z - 1)},$$ where *z* is the number of firms in Z.

END OF INSTRUCTIONS

The instructor should inform the students that the "max" operator on the demand function, $q_i = \max \{230 - 8p_i + \sum_{j \neq i} p_j \ , 0 \}$, is there in case some firm charges such a high price relative to the others that $q_i = 230 - 8p_i + \sum_{j \neq i} p_j < 0$. In this case, the quantity will be truncated to be zero.

In past trials, firms have typically been comprised of three to seven students each, depending on the size of the class. Students in Principles of Microeconomics, Industrial Organization, and Managerial Economics have participated in this game. It is a good idea to give students ample time to work through examples on their own before they play. Thus, it is advisable to have students confer in part of one class period and then actually play the game in the next class period. Outside of perhaps spending more time on examples, no modifications of the game are necessary for Principles' students (as long as they have covered things like marginal revenue, marginal cost, monopoly, and profit calculations).

III. PRE-GAME ANALYSIS

A. NUMERICAL EXAMPLE

The Instructor may decide to give students a numerical example to help them get accustomed to working the details of the game. If all firms were to choose a price of \$50, then firm *i* would sell

$q_i = \max \{230 - 8p_i + \sum_{j \neq i} p_j \ , 0 \}$

$= \max \{230 - 8p_i + \sum_{j \neq i} (50), 0 \}$

$= \max \{230 - 8(50) + 200 , 0 \}$, because $\sum_{j \neq i} (50) = 200$, due to there being four other firms;

$= \max \{30 , 0 \} = 30.$

Then, we can substitute this into firm *i*'s profit equation to obtain:

$p_i = (p_i - 20)q_i - 300$

$= (50 - 20)(30) - 300$

$= 900 - 300 = 600.$

B. RESIDUAL DEMAND DISCUSSION

To conduct more pre-game analysis, the instructor may wish to introduce the students to the concept of the residual demand function. This is the demand facing firm *i* once all other firms' prices (or predicted

prices) are taken into account. In this game, we obtain firm *i*'s residual demand function the following way. First, invert the positive part of firm *i*'s demand function, $q_i = 230 - 8p_i + \sum_{j \neq i} p_j$, and obtain

$$p_i = \frac{230 + \sum_{j \neq i} p_j}{8} - \frac{1}{8} q_i .$$

The instructor can graph this noting that the intercept term is the first term on the right-hand-side of the equation above. Then note that as other firms increase their prices, $\sum_{j \neq i} p_j$ increases. Therefore, the residual demand facing firm *i* increases. This is analogous to a case in which Coke sees an increase in its demand when Pepsi increases its price, ceteris paribus.

Next, note that for linear demand, marginal revenue has the same intercept, but is twice as steeply sloped. Thus,

$$MR_i = \frac{230 + \sum_{j \neq i} p_j}{8} - \frac{1}{4} q_i .$$

It may be useful to remind students that in a one-period framework, profit for firm *i* is maximized at the quantity for which MR = MC, and then the optimal price would be read off of the demand curve. The instructor may go into more detailed examples if desired.

IV. RESULTS FROM A "TYPICAL" GAME

Next, results from a "typical" game are given. The word "typical" is used loosely, here, but most games have ended with prices near the range they would be in a one-period, simultaneous choice Nash equilibrium (which by the way would be all firms charging $32.50 – more on this later). Students are also instructed that admissible prices must be multiples of $0.25. This seems to make adjustments in the game move at a faster pace than if they can change prices by a penny. The Spreadsheet in Table 1 below lists the round of the game in the leftmost column, entitled "Rnd#1." The second column, "p(i)," gives the five firms' prices for each round. Note below that in round 1, firm 1 charged a price of $24, firm 2 charged a price of $25, firm 3 charged a price of $40, and firms 4 and 5 charged a price of $30. The third column, q(i), tells the quantity each firm sold under these prices. The entry uses the demand equation. For firm 1 in this case, $q_i = \max \{230 - 8p_i + \sum_{j \neq i} p_j , 0 \} = \max \{230 - 8 (24) + (25+40+30+30) , 0 \} =$ 163. The fourth column, prof(i), gives each firm's profit and uses the profit equation given in the instructions. Note that in the first round, firm 1 had a markup of $4 over its constant marginal cost of $20, and sold 163 units. This gives total markup of $652, but then the fixed cost of $300 must be subtracted, to arrive at the profit of $352. The fifth column, avg proft(i), simply keeps track of a given firm's average profit per round as the game progresses. The sixth column, e.c., gives the current status of a firm's extra

credit (which is a firm's average profit divided by 300 – the instructor can choose different numbers, of course.). Please note that the level of extra credit can fluctuate like a thermometer during the game. Students should be made aware of this so that they do not think extra credit on the spreadsheet readout is cumulative. Also, a spreadsheet is used by the instructor and projected onto a screen so that students can view what is happening during the game. Students will not know what prices their rivals are choosing in a given round, but they will see the prices chosen when all are entered for that round. Therefore, students will see the pricing histories of their rivals as the game progresses.

Table 1: Spreadsheet of a Past In-class Game

Rnd #1	p(i)	q(i)	prof(i)	avg prof(i)	e.c.
1	24	163	352	352	1.173333333
	25	154	470	470	1.566666667
	40	19	80	80	0.266666667
	30	109	790	790	2.633333333
	30	109	790	790	2.633333333
AvgP,1	**29.8**				
2	28	122.5	680	516	1.72
	27.5	127	652.5	561.25	1.870833333
	28	122.5	680	380	1.266666667
	32.5	82	725	757.5	2.525
	28.5	118	703	746.5	2.488333333
AvgP,2	**28.9**				
3	31.99	99.78	896.3622	642.7874	2.142624667
	33.7	84.39	856.143	659.5476667	2.198492222
	27.75	137.94	769.035	509.6783333	1.698927778
	33.5	86.19	863.565	792.855	2.64285
	30.75	110.94	892.605	795.2016667	2.650672222
AvgP,3	**31.538**				
4	32.5	98.75	934.375	715.6843	2.385614333
	32.75	96.5	930.375	727.2545	2.424181667
	30.75	114.5	930.875	614.9775	2.049925
	33.5	89.75	911.625	822.5475	2.741825
	31.75	105.5	939.625	831.3075	2.771025
AvgP,4	**32.25**				

Note from Table 1 that in the first round of the game, the average price was $29.80. The low price was firm 1 at $24 and the high price was firm 3 at $40. Firm 1 sold the most, but had a meager $4 markup. It had the second-lowest profit. Firm 3 had a $20 markup, but only sold 19 units and only had profit of $80. Both of these firms could do after-the-fact analysis and see that they could have done better. Firm 1 could have increased its profit by increasing its price by a dollar (for example). If it had charged $25, its markup would have been $5, it would have sold 8 fewer units (recall the own-price coefficient on

a firm's demand curve is "8"), which would have been 155. Thus, firm 1's profit would have been $p_i = (p_i - 20)q_i - 300 =$ \$5*155 – \$300 = \$475, which is noticeably higher than \$352 it earned in round 1. In fact, if firm 1 were to set *MR* = *MC* as discussed in a previous section, it could have solved for its optimal quantity, then plugged that back into its residual demand curve to get its optimal price. Of course, this is all after-the-fact analysis, but it can give a firm an idea of where things stood in the previous round, and may help the firm figure out what it wants to do in the next round. Firm 3 could do similar calculations to determine it could have made more profit by charging a price lower than the \$40 price it chose in round 1.

Note in round 2 that the low price firms (firms 1 and 2, charging \$24 and \$24, respectively) decided to increase their prices (to \$28 and \$27.50, respectively). Note also, that the highest-priced firm (firm 3) decided to lower its price from \$40 to \$28. The two middle-priced firms (4 and 5) kept their prices about the same. The more extreme priced firms (1, 2, and 3) saw their profits increase, but the middle-priced firms saw their profits decrease somewhat primarily because the average price came down from \$29.80 to \$28.90. After this, it appears that firms may have been figuring out what was best given the other firms' choices, and that other firms were thinking this way to. It seems to be this way because the price drifted toward what would be the one-period, simultaneous-choice Nash equilibrium of this game (i.e., all firms pricing at \$32.50). In the last round of this game, the average price was \$32.25, without much variation. This average happens to be equal to the one-period, simultaneous-choice Nash equilibrium of this game.

Although repeated games with unknown numbers of rounds may have multiple equilibria, this game has usually ended up relatively close to the one-period, simultaneous-choice Nash equilibrium. Perhaps this is due to the fact that the students attach a higher probability to the rounds that are close to the end of the class period, and effectively treat it like the last round. Regardless, this does give the instructor opportunity to discuss the concept of Nash equilibrium in the context of this game.

V. POST-GAME ANALYSIS

As mentioned in the previous section, the instructor can illustrate that all firms charging \$32.50 is a one-period, simultaneous-choice Nash equilibrium of this game. One could do this mathematically, but immediately following the game, it may be more instructive to illustrate it on computer via overhead projection. The instructor can type in all five prices at 32.50, and students will see that each firm sells 100 units and earns profit of \$950. Then the instructor can change one firm's price (unilaterally) to \$32.25 and see that profit for that firm falls below \$950. Similarly, the instructor can change that firm's price to \$32.75 and see that profit for that firm falls below \$950 as well. Thus, if the other firms are charging \$32.50, the best the remaining firm can do is to charge \$32.50 as well. This will hold for the other firms as well by the symmetry of the setup. (The instructor can discuss how things might be different if one firm, for instance, had a lower constant marginal cost than the other firms.) This can be used as a springboard for

discussion about where prices come from in oligopolistic markets. Other aspects of the game could be discussed here as well.

VI. COOPERATION

An interesting experiment would be to allow one round of cooperation by allowing students to talk across teams and make agreements if they like. At some point, it is useful to tell them they can talk no longer across teams, and subsequently make their decisions and hand in their prices independently. If an agreement is made to charge higher prices than have prevailed during the non-cooperative game, many teams will tend to undercut the agreement at least to some degree (at least they often have in past trials). Then it would be a good idea to discuss what is going on in this case.

After this experiment, the instructor may tell the students that if all choose a price of $38.75, the sum of the five firms' profits (joint profit) will be $1106.25. The quantity that each firm will sell in this case is 75. Then, we can use the residual demand and its associated marginal revenue to explain the incentive the individual firm has to undercut the agreed-upon, higher price. Start by looking at the one-period, simultaneous-choice Nash equilibrium of all firms charging a price of $32.50, and each firm selling 100 units and earning profit of $950. Now suppose all five firms agree to charge a price of $38.75. As stated before, this will cause each firm to sell 75 and earn profit of $1106.25. From firm *i*'s point of view, however, the increase in the price of the other firms' prices (from $32.50 to $38.75) will increase its residual demand (and its associated marginal revenue). Thus, firm *i* has individual incentive to sell more than it did in the one-period, simultaneous-choice Nash equilibrium (which was 100 units). However, we said that firm *i* would only sell 75 if it increases its price to $38.75 along with the other four firms. Therefore, if firm *i* were to undercut the agreed-upon price of $38.75, it could sell more than 75 (and in fact more than 100 if it were to cut its price enough). The instructor could discuss the idea of the "Best Response" function here. Unfortunately for all of the firms, they have this same individual incentive to undercut the agreed-upon price of $38.75. Therefore, it should not be surprising that players often price at a lower level than they agreed to in market experiments like this one. The instructor might find it valuable to discuss cases of such pricing agreements and their stability. (If the instructor decides to run the non-cooperative part of the game for extra credit, it may be better not to do the cooperative part for extra credit because of the individual incentive to cheat on the collusive agreement.)

VII. CONCLUSION

I have found this game to be very useful in helping students understand a variety of principles for pricing. As an instructor progresses through the course and introduces topics such as best-response functions, Nash equilibrium, residual demand functions, joint profit maximization, and undercutting collusive pricing agreements, (s)he can refer to elements of this game and students' experience with it to

help students gain a deeper understanding for the application of the analytical models they are learning. Furthermore, students can gain more facility in analyzing real markets by comparing aspects thereof to what they have learned from this game and beyond.

REFERENCES

Eckalbar, John, C. 2002. “An Extended Duopoly Game.” *The Journal of Economic Education* 33(1): 41 – 52.

Meister, J. Patrick. 1999. “Oligopoly: An In-class Economic Game.” *The Journal of Economic Education*, 30(4): 383 – 391.

Pepall, Lynne, Dan Richards, and George Norman. 2008. *Industrial Organization: Contemporary Theory and Empirical Applications*. Fourth Edition. Massachusetts: Blackwell Publishing.