Matthew Brigida  Volatility in Electricity Spread Options

Hillar Neumann
John Meyer  Quantitative Easing: Where is the Growth?
Where Did the Reserves Go?

Robert R. Dunn
Johnathan T. Williams  The Effect of Race on Work Commuting
Time in the Pittsburgh, PA Metropolitan Area

Kerry A. Adzima  Factors Contributing to Higher Student
Achievement in Pennsylvania Charter Schools
Table of Contents

Matthew Brigida
Volatility in Electricity Spread Options ................................................................. 1

Hilmar Neumann & John Meyer
Quantitative Easing: Where is the Growth? Where Did the Reserves Go? .................. 18

Robert R. Dunn & Johnathan T. Williams
The Effect of Race on Work Commuting Time in the Pittsburgh, PA Metropolitan Area .... 43

Kerry A. Adzima
Factors Contributing to Higher Student Achievement in Pennsylvania Charter Schools .......... 61
Initial submissions must be made in triplicate with a cover letter indicating that the manuscript is not currently being considered at another publication. Also include a complete electronic copy of the paper in MS Word or as a PDF file, or email file to the co-editors. If a PDF file is sent, remove any author identifying information from the file. The submission fee is $25. Send submissions to:

Drs. Thomas Tolin & Orhan Kara, Co-Editors
Pennsylvania Economic Review
Department of Economics and Finance
Anderson Hall, 309
West Chester University
West Chester, PA 19383
Email: ttolin@wcupa.edu & okara@wcupa.edu

Papers should have a cover page with the author's name, address, and affiliation. The second page should have centered in all caps and bolded the paper title, double space and in upper and lower case include the authors name followed by affiliation. Triple space and a fifty to one hundred word abstract should be included. The paper itself should be typed in times new roman 10 point. Do not number the pages electronically but instead write them in pencil on the back of each page. All margins should be 1.7 inches.

All papers accepted by the PER for publication must be submitted in Microsoft Word. Complete guidelines will be sent upon acceptance. Some basic guidelines include:
1. Equations: Default settings should be used. Set point size at 10 points; use Arabic numbers flush right in parenthesis.
2. References in the text should be parenthetical references with the author's last name and date of publication. Alphabetical reference list should be included. See reference lists in this issue for style format.
3. Tables should be numbered in Arabic numbers, e.g., Table 1. Columns of the tables should be set using tab stops, NOT SPACES, to align OR the table setting in Word can be used. Tables should be placed at the end of your article and their appropriate placement in the text noted.
4. Figures are numbered similarly to tables, e.g., Figure 1. The author will provide the figures in a Microsoft Word compatible program.
5. If you need to use footnotes, DO NOT use the endnote or footnote setting in Word. Use superscript numbers to refer to your endnotes, which should appear at the end of the article.
6. Sections of the paper, such as INTRODUCTION, should be flush left and bolded in all capital letters with one extra line space between section head and text. For example:

**INTRODUCTION**

Subheads should be in upper and lower case letters, flush left and italicized, with one extra line spacing above and below subhead. For subheads below the first level subhead, indent one tab for second subhead. Please limit subheads to no more than two.
VOLATILITY IN ELECTRICITY SPREAD OPTIONS

Matthew Brigida
Clarion University of Pennsylvania

ABSTRACT

It is widely held that an option’s time value is increasing in underlying asset volatility. In fact, this relationship does not hold for an important type of option, the spread option. In this paper we will show the conditions under which an increase in volatility can reduce the premium of a spread option. Moreover, we introduce, and show this result for spark and locational spread options. These options are heavily traded to hedge electricity generation and transmission assets, as well as being used to value power plants.

INTRODUCTION

This paper investigates the relationship between the price of a spread option, and the volatilities of the price processes underlying the spread. In particular we consider the relationship between a ‘spark spread’ option price and the volatility of natural gas and electricity prices. We do so because of the unique aspects of electricity and natural gas price processes, which will be explained later. Moreover, spark spread options have become widely used by electricity producers (mainly utilities) to hedge. The need to hedge has become increasingly important to electric utilities, because deregulation in the natural gas and wholesale electricity markets has exposed these utilities to substantial price risk.

With deregulation in these markets, and the increase in the trading of financial contracts to mitigate the resulting price risk, there has also been an increase in the amount of research on these markets and contracts. In Cartea and Villaplana (2008) and Bessembinder and Lemon (2002), fundamental models explaining electricity price behavior are constructed. Further, there has been substantial progress in creating econometric models which allow for particular stylized characteristics of electricity price processes. Notable among these papers are Mount, Ning, and Cai (2006), and Janczura and Weron (2010) which used regime switching models. Models of natural gas prices have often priced natural gas relative to crude oil. Examples are Hartley, Medlock, and Rosthal (2008), and Brigida (2014a) and Brigida (2014b) which modeled the relationship allowing
for Markov regime switching. Lastly, recent research on spark spread options are Hurd and Zhou (2010) which used a Fourier transform to value spark and other spread options, and Deng et al (2001) which derived a closed-form spark spread option pricing model which incorporated the unique aspects of natural gas and electricity price processes.

Through this analysis of spread options, we’ll show conditions under which the value of the option can be decreasing in the volatility of an underlying asset. Moreover, we’ll show the effect of increasing correlation in the underlying price processes on spread option prices. Because spread options methods have supplanted discounted cash flow as a method to value power plants, our results have capital budgeting implications for power generation companies. Moreover, since it is a matter of public policy to ensure there are sufficient market based incentives for private companies to build power plants, our analysis is informative regarding conditions under which power plants are more likely to be built.

The remainder of the article is organized as follows. Section I introduces the classical spread option, and highlights assumptions necessary so it may be applied to spread options on electricity and natural gas. Section II summarizes the most salient uses of electricity and natural gas spread options. Section III defines a spark spread option valuation model and then proceeds to show that the spread option value may be decreasing in underlying electricity or natural gas return volatility. The latter point is the focus of the paper. Section IV discusses implications of the main result, and section V concludes.

I. SPREAD OPTIONS ON ELECTRICITY AND NATURAL GAS

Classical Spread Options

Building on the Black and Scholes (1973) option valuation method, Margrabe (1978) successfully valued the option to exchange some amount of one asset for another asset. Because Margrabe followed the Black-Scholes pricing paradigm, the two assets in this spread option each follow geometric Brownian motion processes. Thus, this pricing model is useful for valuing spread options on equity and foreign exchange, but as we will discuss it is inappropriate when valuing spread options on many commodities.

Price spreads can generally be categorized as locational, calendar (time), processing, and quality. A locational spread may be the difference in the electricity price between two nodes on the network. A calendar spread may be the difference between the price of natural gas for March 2015 and April 2015 delivery. Examples of processing spreads are the crack spread (buying crude oil and selling gasoline and heating oil), or spark spread (buying natural gas and selling electricity) which are the focus of this paper. The processing spreads represent a refining margin. Quality spreads may be the difference between the prices of different grades of crude oil. For a further overview of the many types of spread options see Carmona and Durrleman (2003).

Electricity and Natural Gas Prices

Electricity prices are unique in that they exhibit both strong mean-reversion and price proportional volatility. The latter quality is shared with stock prices and the former is shared with interest rates. However, stock prices do not exhibit mean-reversion and interest rates do not tend to exhibit volatility proportional to the interest rate level (and is
thus not included in the customary models of the evolution of interest rates: Vasicek; Cox-Ingersoll-Ross; Black-Derman-Toy; and Hull-White).

Moreover, unlike interest rates, large deviations from the mean are common and prices tend to revert quite rapidly. These large temporary deviations from the mean price are commonly referred to as 'price spikes'. Price spikes are due to the non-storability of electricity, compounded with near-perfectly inelastic short-term demand for electricity (load). Moreover, for high loads the electricity supply curve is very steep. This is because increasingly inefficient plants must be dispatched to increase supply. So supplying additional units of electricity at high loads incurs substantial marginal costs.

This means for high loads, the supply and demand curves are both very steep, so any change in the load will translate into large price swings. For an overview of electricity prices dynamics and the causes see Bodily and Del Buono (2002). Since natural gas is primarily used for electricity generation, the dynamics of electricity prices dominate natural gas prices.

Because electricity and natural gas price dynamics are so markedly different from stock or interest rate processes, one may not apply stock or interest rate option pricing models to options on electricity and natural gas. Option pricing models are required which are specific to electricity and natural gas price processes (Carmona and Durrleman (2003)).

**Spark and Locational Spread Options**

A spark spread option is defined as a contingent claim with a payoff at expiration of $\max(S_H^T - K_H S_G^T, 0)$, where the expiration of the claim is at time $T$, $S_G^T$ is the spot price of electricity at time $T$ in megawatt hours (MMWh), $S_H^T$ is the spot price of natural gas at time $T$ in millions of British thermal units (MMBtu), and $K_H$ is the heat rate which describes proportionally how much natural gas is required to generate electricity. The heat rate $(K_H)$ is defined as the number of British thermal units (Btu) of input fuel required to generate 1 megawatt hour (MWh) of electricity. Thus, for a particular plant, the heat rate can be written as $K_H = \frac{s_g}{s_e} \text{MMBtu/MWh}$.

Due to the inability to store electricity, the market is completed with electricity and natural gas futures contracts. This means risk-neutral pricing may be used, implying the time $t$ value of the spark spread option is by definition:

$$ C_t = e^{-r(T-t)} E_t^Q \left[ \max\left(F_H^{t,T} - K_H F_G^{t,T}, 0\right) \right] $$

(1)

where $C_t$ is the time $t$ value of the spread option, $E_t^Q$ denotes expectation with respect to the risk-neutral measure, and $r$ is the risk-free rate.

A locational spread option has a payoff $\max(S_H^{t,1} - S_H^{t,2}, 0)$ where the numbers 1 and 2 are used to denote different locations. The time $t$ value of the locational spread option is found analogously.

**II. APPLICATIONS OF SPARK AND LOCAIONAL SPREAD OPTIONS**

*Power Plant Valuation*
A power plant is nothing other than a real option to exchange a certain amount of fuel for another amount of electricity (the heat rate dictates the amount of each). The market price of fuel (natural gas) and electricity at any moment dictates whether exercising the option to convert fuel to electricity is profitable. If it is unprofitable you shut down your plant and earn $0. If it is profitable you run your plant and earn \( S_E - HR \times S_G \) where \( S_E \) is the spot price of electricity per MWh, \( S_G \) is the spot price of natural gas per MMbtu, and \( HR \) is the power plant’s heat rate. Thus your payoff is \( \max(0, S_E - HR \times S_G) \) which is the payoff on a spread option. In sum, a power plant’s value is equivalent to the sum of the values of a series of spark spread call options over the life of the plant.

**Transmission Line Valuation**

The US next generation national electrical grid requires transmission capability between the many present regional grids\(^7\). Further, clean electricity generation is often located far from urban centers and so also requires new transmission lines to be a viable electricity source. Accordingly, firms are presently considering large capital investments in transmission lines highlighting the need for accurate transmission valuation models.

For example, NextEra Energy Resources owns approximately 2,335 MW of wind generating capacity in west Texas. West Texas however usually has electricity prices far lower than the rest of Texas. This prompted NextEra Energy to build a transmission line from its wind farms in west Texas to where the electricity could be distributed to south, east, and north Texas. This transmission line can be valued as a real option with payoff \( \max(0, S_{E,w} - A \times S_{E,w}) \) where \( S_{E,w} \) is the spot price of electricity per MWh in west Texas, \( S_{E,w} \) is a composite spot price of electricity in the other regions of Texas, and \( A \) is an adjustment for transmission line loss.

**Hedging Generation and Transmission Assets**

Spark and locational spread options are widely used by power generators and transmission companies for hedging. Hedging power generation entails both selling electricity, and buying natural gas, forward. The sum of these two transactions, and the plant’s heat rate being given, uniquely defines a spark spread option. It is often more convenient to enter into one spark spread option transaction rather than individual forward transactions. Spark spread call options are generally traded directly between energy companies because the volume of trades is yet insufficient to foster a dealer market. For example, to hedge utilities generally sell spark spread call options to natural gas producers. Hedging transmission assets is analogous.

**III. SPARK SPREAD OPTION VALUATION**

Over the last decade there have been significant improvements in the valuation of power generating facilities through the use of the real options approach. A seminal paper in this regard is Deng et al. (2001) wherein an analytical model was derived for a real option on a gas-fired power plant where the underlying processes exhibit mean-reversion and price proportional volatility. The model however did not include operational constraints (cost and time for plant ramp up/down, minimum up/down-time, commitment constraints, etc.) so it may be considered an idealized model of pure plant generation value.
Note, the spread option model in Deng et al (2001) like that in Margrabe (1978), used the correlation between the two underlying assets. One may alternatively model the spread itself if the underlying price processes are cointegrated (Dempster, Medova, Tang (2008)). However, there is substantial evidence against cointegration between natural gas and electricity prices (particularly with respect to data sampled at the hourly or less frequency).

Later studies have focused on incorporating plant constraints into the real options model (Barz and Tseng (2002), Deng and Oren (2003), Hlouskova, et al., 2005). While most studies value fossil fuel generation facilities, Kiriyama and Suzuki (2004) use real options to value nuclear powered generation with CO2 emissions credits. Prior to the application of real options valuation, power plants were valued using discounted cash flows (net present value).

The following closed-form value of a spark spread call option was derived by Deng et. al (2001):

\[
C = e^{-r(T-t)} \left[ F_e^{T-t} N(d_1) - K_t F_o^{T-t} N(d_2) \right] \tag{2}
\]

where

\[
d_1 = \frac{\ln \left( \frac{F_e^{T-t}}{K_t F_o^{T-t}} \right) + \nu \sqrt{T-t}}{\nu \sqrt{T-t}}
\]

\[
d_2 = d_1 - \nu \sqrt{T-t}
\]

\[
\nu^2 = \frac{\int_{T-t}^{T} \left[ \sigma_e^2(s) - 2 \sigma_e \rho \sigma_o(s) \sigma_o(s) + \sigma_o^2(s) \right] ds}{T-t}
\]

\[
\sigma_e(t) = \sigma_e e^{-k_e t}, \quad \sigma_o(t) = \sigma_o e^{-k_o t}
\]

where \( F_e^{T-t} \) and \( F_o^{T-t} \) denote the time t forward price, of electricity and natural gas respectively, of a contract maturing at time T; \( \nu^2 \) is the time-dependent spread variance; \( \sigma_e \) and \( \sigma_o \) and denote annualized volatility of returns for electricity and natural gas, which may be estimated as the annualized standard deviation of log returns (historical volatility) or by calibrating to present market prices for similar options (implied volatilities); \( k_e \) and \( k_o \) are the rates of mean reversion of natural gas and electricity prices; \( \rho \) is the instantaneous correlation coefficient between electricity and natural gas returns and is a constant in the model; \( K_t \) denotes the plant's heat rate; \( r \) is the risk-free rate; \( N(\cdot) \) denotes the cumulative normal distribution function. Note the use of the risk-free rate means we are pricing the spark spread call option in a complete market (we may replicate the option by trading in forward contracts on electricity and natural gas).

This valuation model was derived assuming the changes in electricity and natural gas forward prices followed the following stochastic differential equations:

\[
dF_e = \kappa_e (\mu_e(t) - \ln F_e)dt + \sigma_e(t)F_e dB^1_t \tag{3a}
\]

\[
dF_o = \kappa_o (\mu_o(t) - \ln F_o)dt + \sigma_o(t)F_o dB^2_t \tag{3b}
\]
where \( \mu \) denotes the long-run mean change, and \( B' \) and \( B'' \) are two Weiner processes with instantaneous correlations \( \rho \). Note these price processes exhibit mean reversion and price proportional volatility.

In this paper we are interested in the behavior of the volatility of the underlying price processes (natural gas and electricity), and how this affects the spread volatility \( \nu \) and ultimately the value of the option. None of the preceding relationships were investigated in Deng et. al (2001) or subsequent studies, and are therefore new to the literature. Note, in the case of a European call or put option valued in a complete market (i.e. the Black-Scholes-Merton model), option values are increasing in volatility. This relationship was proven for options with fixed strike prices by Bergman, Grundy, and Wiener (1996). The relationship is also commonly known as the "Greek" vega, and is analogous to the relationship between the spread option price and spread volatility. We show the conditions under which the underlying price processes reduce the spread volatility and thus the spread option price.

First, let us examine the relationship between the spread option price \( C \) and \( \nu \). Ideally, we would take the partial derivative of the option price \( C \) with respect to \( \nu \). However, \( \nu \) is a function of the volatility of natural gas and electricity, and the forward prices in 3a and 3b are also functions of electricity volatility and natural gas volatility, respectively. This means any change in \( \nu \), which would necessarily be due to changes in either electricity or natural gas volatility, would imply changes in the forward prices. This relationship will confound analytical attempts at taking the partial derivative.

We thus investigate the relationship graphically. Below in Fig. 1 is the relationship between \( \nu \) and the call option price. Note, we varied \( \nu \) by changing the underlying electricity and natural gas volatilities. From Fig. 1 we see that the option premium (C) is monotonically increasing in the spread volatility (\( \nu \)).

We are now ready to investigate how changes in the volatility of electricity (\( \sigma_E \)) and natural gas (\( \sigma_G \)) affect the option price. Since the option price \( C \) is strictly increasing in \( \nu \), any change in the underlying volatilities of electricity and natural gas that increases \( \nu \) will increase the option price and vice versa. Consider:

\[
\nu^2 = \left( \frac{1}{T-t} \right) \int_t^T (\sigma_E^2 - 2\rho \sigma_E \sigma_G + \sigma_G^2) \, ds
\]

and as case (1) set \( \rho = 0 \), then

\[
\nu^2 = \left( \frac{1}{T-t} \right) \int_t^T (\sigma_E^2 + \sigma_G^2) \, ds
\]

and as \( \sigma_E^2 \) or \( \sigma_G^2 \) increases so does \( \nu^2 \). Alternatively, for case two set \( \rho = 1 \). And without loss of generality set \( \sigma_E^2 > \sigma_G^2 \), then

\[
\nu^2 = \left( \frac{1}{T-t} \right) \int_t^T (\sigma_E - \sigma_G)^2 \, ds
\]

and as \( \sigma_G \) increases \( \nu^2 \) decreases which implies the option price decreases. So in case (2) the volatility of natural gas is negatively related to the call option price. This is contrary to what many assume. Note, if we had assumed \( \sigma_E^2 < \sigma_G^2 \) then electricity volatility is
negatively related to the option price (and natural gas volatility is positively related to the option price). In short, for high correlations (only), it is the greater difference, not the greater sum, of volatilities which increases the option price.

In fact, if $\sigma_E = \sigma_G$ then as $\rho \to 1$, this implies $\nu \to 0$, and the option premium approaches its intrinsic value: $\max\left[ P_E^{ST} - K_H P_G^{ST}, 0 \right]$. Intuitively, this is because if $\rho = 1$
and \( \sigma_E = \sigma_G \) then the spread will always stay the same and so there is no benefit to owning an option. Alternatively, under these conditions, and fixing time, forward prices, and the heat rate, we have:

\[
\lim_{\nu \to 0} d_1 = \lim_{\nu \to 0} \frac{\ln \left( \frac{F_E^T}{K_H F_G^T} \right)}{\nu \sqrt{T - t}} \to \infty \quad \Rightarrow \quad \lim_{\nu \to 0} N(d_1) \to 1
\]

and,

\[
\lim_{\nu \to 0} d_2 = \lim_{\nu \to 0} \left( d_1 - \nu \sqrt{T - t} \right) \to \infty \quad \Rightarrow \quad \lim_{\nu \to 0} N(d_2) \to 1
\]

so finally,

\[
\lim_{\nu \to 0} C = \lim_{\nu \to 0} e^{-r(T-t)} \left[ F_E^{T,t} N(d_2) - K_H F_G^{T,t} N(d_2) \right] \to e^{-r(T-t)} \left( F_E^{T,t} - K_H F_G^{T,t} \right)
\]

which is the option’s intrinsic value.

This result is exhibited graphically in figures 3-7 in the appendix. Figure 3 shows the option time value for varying levels of natural gas and electricity price volatility assuming the correlation between the two underlying processes is 0.99. Note when natural gas and electricity volatility are equal, the option value decreases to 0.

Figures 4-7 fix the natural gas volatility and show the option time value for varying correlations and electricity volatility. In figure 4 natural gas volatility is fixed very low (at 1%) and so the option value behaves much like a standard call option with a fixed strike price. In the following figures (5 through 7) natural gas volatility is fixed progressively higher (25%, 50%, then 100%) showing, for high correlations, the option value decreases to zero as electricity volatility approaches natural gas volatility.

**IV. IMPLICATIONS**

We have shown spread option values may be decreasing in the volatility of an underlying asset. This means increasing volatility may not benefit the owner of generation assets or transmission capabilities as has previously been assumed. Further this result has also clear implications for understanding and trading spark spread options.

An interesting corollary of the above analysis is that if we fix natural gas and electricity volatility and allow the correlation to vary the option value will also be affected. Specifically, holding all else constant power plant generation value decreases in an increasing correlation between natural gas and electricity returns. Thus, not considering time-varying correlations would imply misvaluation of generation.

In the chart below the dashed line is the 90-day rolling correlation between Henry-Hub natural gas price returns and PJM-west electricity price returns. The solid line is the percent by which the month-ahead generation is under or over valued when assuming a constant correlation of 16% (the correlation estimated at the beginning of the period). Note, during the 2008 financial crisis the correlation between natural gas and electricity price returns dropped to below -10%, causing the generation to be about 25% undervalued. For a moderate size natural gas-fired power plant this amounts to an approximate $126,000 undervaluation of daily generation value.
V. CONCLUSION

This paper has summarized the applications and valuation of spread options on electricity and natural gas. Within a particular valuation framework, we showed that the spread option time value may be decreasing in the volatility of an underlying process. This point is consequential and new to the literature. We then discuss implications of this finding with respect to the real options valuation of natural gas fired power plants. In particular we note when the value of a power plant can be decreasing in the volatility of an underlying asset. Further, we note our analysis implies power plant values will decrease as the correlation between the underlying price series increases. These results are important to electricity producers which may wish to understand the underlying factors which affect their power plants value. The results are also important as a matter of public policy, because high plant values will tend to ensure there is sufficient generating capacity.
Footnotes:

1 You can store electricity by pumping water uphill or compressing gas underground. The proposed Solana solar plant in Arizona intends to store electricity by heating tank of a saline solution which will retain heat that later may be extracted to generate electricity via a steam turbine. At present, however, due to associated costs most electricity is produced when demanded.

2 Natural gas is most common however you may also use coal as the fuel.

3 There are 8 independent system operators in North America.

4 To value a locational spread option, such as the value of an electricity transmission line, simply replace $K_u F_{G}^{S,T}$ with the electricity forward price at another location.

5 The above shows mean-reversion affects the volatility process. Therefore, using GBM and setting the drift to zero will not reproduce mean-reversion – as many trying to price spark spread options with Black-Scholes have tried to do.

6 Say the value of a 1500 MW plant run 24 hours earning a spread of $50MWh - 8*$4.5MMBtu, means approximately $504,000 in generation value which reduced 25% would be $378,00 for a $126,000 undervaluation.
REFERENCES


APPENDIX

Figure 3

Option Time Value (assuming Correlation = 0.89)
Figure 5

Option Time Value (assuming Nat. Gas Volatility = 25%)
Figure 6

Option Time Value (assuming Nat. Gas Volatility = 50%)
Figure 7

Option Time Value (assuming Nat. Gas Volatility = 100%)

[Graph showing a 3D plot of option time value]
QUANTITATIVE EASING: WHERE IS THE GROWTH? WHERE DID THE RESERVES GO?

Hillard Neumann
Northern State University

John Meyer
Northern State University

ABSTRACT

Since the end of the “Great Recession,” almost six years ago, real economic growth has been anemic. Inflation has fallen below the Federal Reserve’s monetary policy target and at times flirted with deflation. All this in spite of the Federal Reserve’s Quantitative Easing program and zero bound short term interest rates. Quantitative Easing injected trillions of dollars into the U.S. financial system, while interest rates across the time spectrum have fallen to historically low levels. The Federal Reserve’s balance sheet has ballooned as holdings of treasury securities and mortgaged backed securities sky-rocketed. The traditional transmission mechanism of monetary policy has broken down. Using a Quantity Theory approach and applying descriptive statistical techniques, this paper explores the fundamental monetary reasons for this dismal economic outcome and the over-all breakdown in monetary policy.

BACKGROUND

Liberals and conservatives alike are frustrated by the recuperative efforts of the U.S. government in the aftermath of the 2008-2009 financial crises. Almost every knowledgeable commentator would concede that the various wrongdoers have gone unpunished. Compromised and entangled firms have been bailed out, and the financial institutions that were “too big to fail” were resired via .025% Federal Reserve discount loans, zero lower bound federal funds, and 3% Treasury rates. So far, the only meaningful relief for the citizenry has been the increase of the FDIC depository insurance from $100,000 to $250,000. This article describes the QE phenomena and illustrates why the results are so unsatisfactory.
In a March 24, 2013 *Wall Street Journal* editorial, Reuven Brenner (McGill University) and Martin Fridson (Standard and Poor’s) picked up on Federal Reserve Chairman Ben Bernanke’s likening of his policies to the U.S. World War II monetary regime. Bernanke had noted the success of the Fed in maintaining “a ceiling of 2.5% on long term treasury bonds for nearly a decade (1940’s)” and suggested that interest rates could be held down even after the U.S. economy eventually revives. But Brenner and Fridson claimed this might not work this time, because political support and patriotic fervor during the war were unique phenomena stoking U.S. Treasury bond purchases in lieu of more competitive bond rates. Moreover, they pointed out that back then government control over the economy was exceptional, including price controls and rationing. Additionally, they noted there was no other safe place for people to put their money during this era. They went on to mention how the currently pending U.S. unfunded entitlement situation, global competitiveness, and pressures for government spending do not auger well for the ability of the U.S. to pay down its national debt cheaply in the future. Soon, we will be coming upon the second anniversary of the Brenner-Fridson editorial and just now the media seem to be taking notice of their arguments.

**PROBLEM STATEMENT**

While it’s true that Bernanke’s pragmatic quantitative easing (QE) policies have not prompted worrisome inflationary pressures so far, the big question is what will happen when increased demand and money velocity do eventually return to the U.S. economy. Currently, the optimists think the U.S. will continue to be the world’s strongest capital magnet, and the pessimists think that uncontrolled inflation is on the horizon. This paper will survey these arguments in the context of Quantitative Easing (QE). We will then present a model based on the quantity theory of money in order to evaluate likely causation for the path taken by key economic variables. The model will then be utilized to explain the historical patterns applying descriptive statistics and graphical analysis to suggest where the monetary policy transmission system broke down, and lastly, discuss the implications for future monetary policy challenges.

*Retracing Monetary History; Keynesian Management or Self-Correcting Mechanism*

A fundamental analytical divide among economists since John Maynard Keynes’ *General Theory of Employment, Interest, & Money* (1936) has never been resolved because the differing views rest upon different assumptions about human behavior; thus, the same evidence can be cited by either side to support their arguments. The unresolved theoretical question; whether persistent high levels of unemployment can be properly characterized as an “equilibrium,” colors the current commentary about QE and the prospects of avoiding future inflation. If Keynes was right, equilibrium can exist at present high unemployment levels and the monetary policy goals of “full” Humphrey-Hawkins employment and stable prices cannot be met without additional massive government intervention including “fiscal stimulus” to revive demand.

On the monetary policy side, Keynes focused on what he termed “liquidity preference” as a fundamental impediment of self-correcting restoration of demand. Classical theory is more confident of investment motivations and tends to rely upon supply side price and interest rate adjustments to restore aggregate demand. Gold standard
advocates would take classical theory even further, and hold that governments can’t be trusted to restrain “beggar thy neighbor” motivations of aggressive policy in trade matters (devaluing their currency to gain trade advantage). They like the automatic self-correcting mechanism of the gold standard, in effect, as a natural law, that can’t be broken without ultimately getting set aright. We think truth lies amidst these theories rather than residing exclusively in any one of them.

It is of passing interest, however, that Keynes penned a popular pamphlet, *How to Pay for the War* (1940), holding that the price system should be allowed to function as normally as possible despite wartime restrictions on private purchasing and compulsory savings plans aimed at war shortages. The idea behind the restrictions was to avoid runaway inflation of critical commodity supplies by controlling consumption. Recently, however, most so-called “Keynesian” stimulus plans try to increase consumption, but Keynes always maintained a conviction that price mechanisms should be allowed to work. Instead of temporarily stimulating consumption, Keynes favored capital expenditures. His activism was mainly focused upon enabling the capitalist investment system to pull itself out of persistent unemployment.

As emphasized by Alan H. Meltzer (2013), “Keynesianism” has morphed into a misguided and broadened debate about Fed activism. (See {Taylor, 2012} for a brief analysis of Congressional opposition proposals to limit this movement.) Some excerpts from Meltzer’s September 24, 2013 testimony to the Senate Budget Committee summarize his point:

John Maynard Keynes is frequently cited as the intellectual father of short-term policies to restore growth by increasing government spending to stimulate private consumption. The 2009 stimulus implemented that policy by offering sizeable temporary tax reduction to middle income taxpayers and temporary payments to state and local governments...

To write my book on Keynes’ work, I read most of his books and papers...No one who has read Keynes’ work carefully can find him favoring policies to boost consumer spending. He opposed them throughout his life. Keynes favored temporary deficits to replace private investment, but he opposed permanent deficits.

Keynes did advocate for activist expansionary fiscal policy during periods of economic malaise and high and persistent unemployment. He believed the government should act as the employer of last resort through deficit spending. We are not there presently. What is needed now is to revive the economy sufficiently to mitigate “involuntary” unemployment.

Christina Romer, former chair of the President’s Council of Economic Advisors, authored seminal research on the U.S. history of monetary policy (Romer, 1999). Her results cast doubts on the ability of the current fiat money standard to moderate the business cycle better than did the gold standard by demonstrating that the frequency and severity of recessions didn’t significantly differ prior to 1913 (gold standard) and after World War II (dollar standard convertible to gold) and post 1971 (fiat standard). Romer’s work would also seemingly support the notion that the Fed significantly contributed to the World War I rates of inflation that set the high water mark of modern history, and which were followed by the record deflation of 1920-21.
As mentioned above, Brenner and Fridson differ with Bernanke’s likening of the current policies to the U.S. WWII monetary regime and cast doubts about the ability of the U.S. to hold down interest rates long enough to pay down its national debt via cheap money policy. In his recent book “The Federal Reserve and the Financial Crisis,” (2013) Bernanke concedes that the Fed’s famous tightening of 1937 prolonged the Great Depression, and that its’ loosening also caused the Great Inflation of the 1970’s. The general trend of monetary policy in recent history has been to avoid over-corrections of both inflation and deflation. We think there is merit in both concerns.

As demonstrated by our graphics, we conclude that although there is no implicit “third objective” (lowering the cost of government debt) to the so-called dual mandate of the Federal Reserve Reform Act of 1977 and the Full Employment Act of 1978, the Fed is going to have its hands full just bailing itself out. Similarly, we are responding to several recent books and articles suggesting that central banks have more to worry about than price stability and full employment. See Adam S. Posen’s testimony before the Joint Economic Committee hearing on “The Fed at 100: Can Monetary Policy Close the Growth Gap and Promote Sound Dollar?” April 18, 2013. See also (Posen, 2010 and 2013), (Bernanke, 2013), (Meltzer, 2013), and (Taylor, 2013).

**The Dual Mandate, Disciplined Discretion, and Quantitative Easing**

Although it is fashionable to refer to QE as "unconventional" policy, not everyone agrees it is really unconventional. St. Louis Fed President James Bullard characterized it as a "substitute for ordinary policy easing" (meaning the lowering of short term interest rates). Bullard concluded that QE was "an effective tool when the policy rate is near zero"; in other words, something that can be done when interest rates are "zero lower bounded." See (Romer, 2013) and (Bullard, 2011). Posen, in his April 18, 2013 testimony to the Joint Economic Committee, acknowledged the right of Congress to set monetary policy goals, but insisted upon "instrument independence" for central banks. He went on to state that broad central bank mandates like "price stability" are not supported by significant correlation, in cross-sectional studies, with the outcomes that are actually delivered (Posen, 1999). Instead, he defines instrument independence as "disciplined discretion" whereby there is a hierarchy of medium term (2-3 years) policy goals with sufficient flexibility to spare the economy "sharp swings over too short of a time frame", including temporary neglect of inflation targeting goals. Thus, "disciplined discretion" would allow the Fed to ignore inflation targets for 2-3 years, provided it meets its other policy goals (that it has chosen not to ignore) e.g., as a matter of near term FOMC discretion.

To support this version of the unconventional policy argument, Posen dismisses the existence of any current threat to soundness of the dollar and argues that economic activity and inflation pressures usually move together. He maintains that "leaning against the wind" (e.g. attempting to pre-emptively curb inflation), and QE, have limited effect on the real economy; (albeit that some asset prices and interest rate spreads have been historically affected). Moreover, he says that even if QE fails to be sufficient to restore full employment (a goal he disputes) nothing will happen to the dollar’s soundness (albeit that high unemployment and lost income would continue to mount). Additionally, he points out that the measures taken by the Fed since 2008 are similar to the quantitative means of purchase and sale of a wide range of private and public sector assets conducted by central banks for the preceding 200 or so years.
Posen (2013) concedes, in a similar vein to Romer (1999), that the modern history of central banks does include examples of high inflation caused by monetizing of public debt, but he insists that this has not been occurring today under QE; and that neither the enabling of public debt levels nor lowering government borrowing costs are current goals of Fed policy (2013):

The golden age of extreme price stability and sound dollar that some like to evoke as an ideal for today occurred while central banks were engaged in precisely this kind of quantitative easing and tightening behavior in the United States, before 1913, when ad hoc groups of private financiers responded to panics and fluctuations in liquidity through their purchases and sale of private assets, in the absence of the Federal Reserve. I defy anyone to come up with a relevant historical example, let alone a consistent statistical pattern in the data, of the mere fact of central banks engaging in large scale asset purchases leading to significant inflation or the erosion of financial markets.

So too, Bernanke recently likened the Fed’s response to the 2008-09 economic crisis to the traditional role of a central bank. In his book, “The Federal Reserve and the Financial Crisis” (2013), he simply identifies traditional central bank actions as lender of last resort (providing liquidity for important institutions), regulation of the financial system, and controlling the money supply as the Fed’s primary functions. He seemingly sidesteps the monetization of debt issue with a terse explanation and a statement that the amount of money in circulation is not increased by QE operations:

And in particular, when the Fed buys assets as part of an LSAP or QE program, this is not a form of government spending. It does not show up as a government spending because we are not actually spending money. What we are doing is buying assets, which at some point will be sold back to the market, and so the value of those purchases will be earned back. In fact, because the Fed gets interest on the securities we hold, we actually make a very nice profit on these LSAPs. What we have done over the past three years is transfer about two hundred billion dollars in profits to the Treasury. That money goes directly to reducing the deficit. So these actions are not deficit-increasing; they are in fact significantly deficit-reducing.

The authors herein think there is an important distinction to be made between interim and ultimate money creation, and thus we see a fallacy in this reasoning. The fact is the Fed purchases these assets by crediting the (seller) bank a reserve deposit that is created out of thin air. When these assets (Treasury Securities and MBS) are sold back to the market, real value is regained, but it has not been “earned.” Instead, it is just being restored from a fictional value credit on the Fed’s books to a marketplace reality. In other words, unless the selling bank (of the QE purchases) was paid in existing cash, (which never happens) new credit (money) has been created; it has been monetized. We think the unwinding process is another matter (see Taylor below) and the two steps may be years apart in execution. The interim monetization “has persisted, and will persist, until actual unwinding ultimately occurs”.

22
Perhaps part of the confusion and controversy arises from some misunderstanding of what QE actually is:

1) Quantitative Easing causes the lowering of short-term and, most recently, long-term interest rates via large quantities of asset purchases (large scale asset purchases, LSAP). The necessary dollar amounts need to be large enough to force interest rates to go down on debt instruments (as the price is forced up). But national debt is not immediately and directly affected because bank reserve credits are swapped for bonds in the typical case. While the liabilities (reserves) of the Federal Reserve are increased, so too, are the assets. Due to the swap, the Fed merely issues a credit out of thin air and therefore doesn’t have to print money to acquire the securities and the amount of currency or debt instruments in circulation is unchanged (at least not concurrently).

2) The purpose of QE is to reduce short term rates and thus stimulate borrowing. During QE1 and QE2, bond yields and mortgage rates actually went up, probably due to the purchase of mortgage backed securities (MBS). However, during these early phases, the purchase of discounted value MBS became a Fed objective due to the collapse of that market and the perceived need to stabilize the industry.

3) QE has obviously stimulated stock prices. But any associated and temporary lowering of long-term mortgage interest rates is claimed to benefit borrowers more than investors. (Try making this argument to someone with less than a 700 credit rating!) The current problems with access to credit are probably due to Dodd-Frank lending restrictions and requirements more than to QE, but common sense dictates that lenders are also likely to be more cautious when historically low interest rates (and spreads) prevail.

4) Theoretically, QE should encourage businesses and savers alike to take on more risk, so long as they believe short-term rates will remain “zero lower bound” for their foreseeable investment horizon. Even the Fed, however, has made it clear that short-term rates will have to rise eventually. It is not clear yet whether business borrowing for real investment will pick up in response to QE, but at least improved solvency has increased the ability of banks to lend.

The U.S. has never formally mandated a stated inflation target, but the overwhelming bulk of current opinion holds that a core rate of about 2% inflation over the long term is a good thing (Bernanke, 2013). According to Benn Steil and Manuel Hinds, inflation targeting theory hopes to achieve long term price stability characteristics of the gold standard without any of the problems associated with it (Steil & Hinds, 2009):

The nineteenth-century rule gold standard was convertibility of paper money and token coinage into gold on demand, but its effect has been successfully mimicked
since the 1980’s, at least in some countries for some periods, by what is now known as inflation targeting. The major difference between the two is that under the gold standard absolute price levels tended to be stable over the long run (that is, inflation was followed by deflation), whereas under inflation targeting it is the rate of price increases that tends to stabilize (that is, above-target inflation is followed by a return to target inflation, not deflation).

It is interesting to note that most of the inflation Hawks and rules-based policy advocates on the Federal Reserve Board have been bankers instead of academic economists. Among academics, the Doves probably outnumber the Hawks. But the foremost currently publishing critic of Fed policy is John B. Taylor, a Stanford academic and author of the “Taylor rule.” Taylor would seemingly agree with the authors herein that reserve credits used to purchase QE assets are created out of thin air. He says this may be justified in the case of emergency liquidity swaps during a financial panic, but he is quite uneasy about the Fed’s ability to unwind the QE asset “purchases” without creating serious inflation pressures (Taylor, 2013):

When the Fed engages in its current policy of quantitative easing, it finances its purchase of mortgage-backed securities or federal debt by crediting the banks with these deposits. The deposits — called bank reserves — normally are increased during times of financial stress, as on 9/11/2001..., or during the panic in the fall of 2008...

The policy {QE} is a drag on the economy in part because people do not know how the bank reserves will be unwound, as they must be eventually.

Recent reports are that these risks are becoming a worry to a number of policy makers at the Fed. People recognize that the Fed will eventually have to undo the interventions and reverse the large-scale asset purchases. If the asset sales are too slow, inflation will rise as bank reserves used to finance the asset purchases flow out of the banks and money growth increases. If the asset sales are too fast or abrupt, there will be a recession as interest rates spike. Those who say not to worry about the interventions because inflation has not increased ignore the fact that the interventions can be a drag on the economy without increasing inflation in the short run.

Frequent op-ed commentator Alan S. Blinder is usually thought to be an inflation Dove, but at least with respect to unwinding QE asset purchases, Taylor is not alone in his view that QE is monetization (Blinder, 2013):

...once you approach the zero lower bound on nominal short rates, there are two main options for pushing longer-term interest rates down further. The central bank can use stronger “central bank talk” as a commitment device (which is not QE). Or it can reduce “the spreads of longer-term interest rates over expected policy rates through asset purchases financed by money creation” (which is QE).
Will the Fed be able to unwind QE skillfully enough to avoid serious inflation? Of course nobody knows the answer for sure. But perhaps a graphic review of the current situation would be instructive.

MODEL AND EVIDENCE

Because we want to demonstrate relative relationships instead of absolute amounts, our exhibits will show the relationships to GDP so that some idea of the relative effects upon the economy can be appreciated. The idea is not necessarily to demonstrate debt and unwinding burdens upon the economy per se. Instead, we want to convey some idea of the scale relationships as hopefully indicating what these relationships have been in the past, and associated current trend lines.

Since the end of the Bretton Woods era the current U.S. economic performance, in comparison, has been far below that of previous expansion phases. The first row of Table 1 summarizes the average of the growth of real gross domestic product and inflation, as measured by the CPI and PCE, over the five expansion periods beginning March 1975. The second row of Table 1 presents the average growth rate of real gross domestic product and inflation in the current expansion, since the end of the “Great Recession” second quarter of 2009. Clearly, economic growth, as measured by the percentage change in real gross domestic product has not kept pace with the average of the previous five expansion periods. Average inflation, referring to both price indices, is below that of the previous expansions and below the 2% inflation target of the monetary policy arm of the Federal Reserve.

It is evident from Table 1 that the current economic expansion is anemic compared to the average of previous expansions. This raises the question of why QE has produced such poor economic performance. Monetary policy under quantitative easing has been extremely aggressive. However, there appears to be a breakdown in the transmission mechanism of monetary policy. The transmission mechanism ordinarily runs from open market purchases to lower interest rates, and then to greater real investment and economic growth. This mechanism has apparently failed. Commercial bank reserves and the monetary base have expanded by trillions of dollars, interest rates are “zero bound,” yet investment and real economic growth are still stymied after six years of QE. Where has the economic system and transmission mechanism of monetary policy broken down? To shed light on this outcome we turn to the well-known and modern version of the quantity theory of money model to descriptively explain this event.

The quantity theory of money is simply stated in the following widely known and used equation.

\[ M \times V = P \times y \]  

(1)

Where:

- \( M \) = the Money Supply
- \( V \) = the Velocity of Money
- \( P \) = the Price Level
- \( y \) = real income (Real Gross Domestic Product)

Transforming the traditional quantity theory into its percentage change, we have:
\[ \% \Delta M + \% \Delta V = \% \Delta P + \% \Delta y \]  
(2)

We can break equation (2) down further by using the well-known money supply process equation:

\[ M = m^* B \]

Where:
- \( m \) = the money multiplier
- \( B \) = the monetary base

Transforming the money supply equation into its percentage change, we have:

\[ \% \Delta M = \% \Delta m + \% \Delta B \]  
(3)

Substituting (3) into (2)

\[ \% \Delta m + \% \Delta B + \% \Delta V = \% \Delta P + \% \Delta y \]  
(4)

Rearranging equation (4) we have the \( \% \Delta y \) or real economic growth equation:

\[ \% \Delta y = \% \Delta m + \% \Delta B + \% \Delta V - \% \Delta P \]  
(5)

Again, rearranging equation (4) we have the \( \% \Delta P \) or the inflation equation:

\[ \% \Delta P = \% \Delta m + \% \Delta B + \% \Delta V - \% \Delta y. \]  
(6)

Using readily available data, equations (5) and (6) can be applied to explain the sub-par performance of real gross domestic product growth and explain why inflation is below both the rates in previous expansions and below Federal Reserve’s target rate of 2%. (According to the Federal Reserve Bank of St. Louis, the five-year break-even inflation expectation is currently only 1.5 %.)

Both real economic growth and inflation are positively related to the money multiplier, the monetary base (bank reserves) and the velocity of money. It is instructive to observe the path of these economic data. Figures 1 and 2 show the path of M1 and M2 since January 1st 1975. M1 and M2 have, respectively, grown at an annual rate of 13.83% and 6.82% for each of the 5 years since June 2009, the start of the post-great recession expansion.

The path the money supply takes has been closely related to that of the monetary base, Figure 3 plots the course of the monetary base from 1970 to 2014. Prior to 2008 with the onset of QE the path of base money remained relatively flat, growing at a relatively constant rate. In August 2008 base money took a vertical path increasing by a rate of 59.5% annually from August 2008 to October 2014. The annual increase in the monetary base by nearly 60% did not correspond to a proportionate increase in either monetary aggregate M1 or M2.
This graphically displayed result raises the question of why the obvious empirical breakdown between the monetary base and the money supply has occurred. The quantity theory of money and the money supply equation provide two avenues of explanation. First, the role of the velocity of money and, second, the role of the money multiplier must each separately be patterned to acquire an understanding of this disconnect.

Figures 4 and 5 illustrate the course of the velocity of the monetary aggregates M1 and M2. M1 velocity has increased over time and peaked in March 2008. After that M1 velocity fell from 10.624 to 6.178 July 2014, a decrease in velocity of 6.7% per year. M2’s velocity fell similarly by 3.3% annually.

The money supply is also directly related to the movement of the money multiplier. Figure 6 illustrates the movement of the M1 money multiplier. The exhibit illustrates the money multiplier took practically a vertical drop in value before the end of the great recession. The precipitous fall began August 2008 when the value was 1.069 and thereafter drifted to .709 in May of 2014. Additionally, the M2 money multiplier has stabilized at its historical low level below 1.0. A similar movement in the M2 money multiplier occurred falling by 12.15% per year from 9.161 in August 2008 to 2.856 in May 2014. Both the movement in the velocity of money and the money multiplier have worked against the intended stimulus of quantitative easing. The short answer to why this has been the case is that money sitting in bank reserves doesn’t affect the velocity of money or the money multiplier. To do that it needs to be loaned out.

The movement of key monetary policy tools and their corresponding economic variables begs the question of where have the monetary base and bank reserves gone? To what use have these newly created reserves been put by financial institutions? To explore this issue we will turn our sights on interest rate movements and federal government debt financing. The short answer is that banks stuck the vast bulk of these resources into non-productive government security investments instead of lending them out. Initially this helped with the serious solvency problems facing financial institutions, and recently it has been alleged to be largely due to Dodd-Frank requirements.

Figure 7 illustrates the movement of interest rates since 2000. It is evident quantitative easing was successful in bringing interest rates down along the term structure from short to long term. These rates today remain at historically low levels. Although both long term interest rates have fallen and short term rates remain at zero-bound, the US economic system is stalled and economic growth remains tepid.

Figure 8 plots the ratio of Treasury Debt Outstanding to GDP along with Federal Reserve Bank holdings of Treasury Securities to GDP. Figure 8 shows the dramatic upswing in the two quantities after the onset of QE. What is significant, however, is the rapid acceleration of the total public debt to GDP and the increasing proportion of Federal Reserve Treasury debt to GDP since the onset of QE. At least in these relative terms, the unwinding task would appear to be formidable in scale as well as potential impact.

Figure’s 9 and 10 demonstrate the sad story of the impact of the 2008-09 economic crisis upon the slow progress that had previously been made upon the debt to GDP ratio (despite the narrowing gap with monetary base to GDP). Not long after the onset of QE both trend lines had more or less doubled. This would seem to belie any claims that QE has not dramatically increased the monetary base via monetization of QE which was financed by the creation of new Federal Reserve credits (bank reserves) credited to banks in payment for purchased treasury securities and mortgage backed securities (MBS).
Figure 11 delineates the relationship of total federal debt and the movement of interest rates. The chart will disappoint those who doubt the possible future negative effects of mounting federal debt to GDP ratios. Quite obviously, as the proportion of federal debt to GDP since the onset of the 2008-09 economic crisis has risen, interest rates on the types of securities QE have all gone down. One would expect as Federal debt increases market interest rates would need to rise, but the intended influence of QE has actually reduced interest rates. The differences between the rates of decline for each type reflect the nature and volumes of respective Fed purchases.

As mentioned previously and demonstrated in the figures, it can’t be said that QE will not significantly affect long term interest rates as shown by the ten year and mortgage rates. At best, these exhibits show that the affects upon short rates are more profound. The anticipated rebound of the ten year and mortgage rates, after initial declines, has not been captured by the data because these events did not begin to occur until 2013.

CONCLUSION

As it appears from our descriptive statistics and graphics, the breakdown of the transmission mechanism of monetary policy is directly related to diminished velocity of both M1 and M2. The extreme fall-off in the money multiplier provides additional evidence why QE failed in stimulating economic activity. It is quite evident that injection of bank reserves and the monetary base did reduce interest rates as expected, however, it appears these reserves ended up financing government spending activity as Federal Reserve Banks purchased Treasury securities and mortgaged backed securities at historically high levels.

Brenner and Fridson’s prognostication of more costly if not troublesome experiences in dealing with U.S. debt levels doesn’t bode well for the Fed in unwinding QE. Holding critical Treasury interest rates down may eventually prove to be difficult insofar as the Fed’s unwinding task itself appears from our graphs to be quite formidable. Ben Bernanke and Janet Yellen may get the last laugh, however, because even long term rates have recently settled back in response to slackening aggregate demand. The effectiveness of QE policy as economic stimulus is currently suspect, however it has not yet been disproven by the post-2008 events.

Your authors are naturally inclined to share Brenner and Fridson’s concerns. But, of course, we would rather be wrong than witness the unvarnished results of persistent and increasingly burdensome debt to GDP levels and corresponding “spending like a drunken sailor.” Senator John McCain, an old Navy man, recently made the quip that the levels of U.S. spending would actually seem insulting to a drunken sailor. Moreover, it appears patently obvious that QE created bank reserves will not fully stimulate the economy if they remain in the banks due to restrictive lending standards. It appears to us that the government has been working at cross purposes to its well-intentioned objectives. We sincerely hope that smooth unwinding of QE will not be checked by excessive regulation and unjustified fiscal stimulus.
REFERENCES


Table 1. Real GDP Growth and Inflation Over the Five Expansions from 1975 – 2007 Compared to June 2009 to May 2014 Expansion

<table>
<thead>
<tr>
<th></th>
<th>Real Domestic Growth Rate</th>
<th>Gross Product Price Index Inflation</th>
<th>Personal Consumption Expenditure Inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average of Five Expansions</td>
<td>3.114%</td>
<td>5.758%</td>
<td>5.054%</td>
</tr>
<tr>
<td>Expansion June 2009 to May 2014</td>
<td>1.819%</td>
<td>1.778%</td>
<td>1.513%</td>
</tr>
</tbody>
</table>
Figure 4

Velocity of M1
Figure 6

M1 Multiplier

Figure 7

Long Term Interest Rates: Five Year Treasury, Ten Year Treasury and Mortgage Rates

Short Term Interest Rates: Federal Funds Rate and Three Month Treasury Rate

Figure 9

Total Treasury Debt to GDP (Right Scale)

Federal Reserve Monetary Base to GDP (Left Scale)

00 01 02 03 04 05 06 07 08 09 10 11 12 13 14
Figure 10

Federal Reserve Bank Holdings of Treasury Debt & Mortgaged Backed Securities
(Billions of Dollars)

FRB Treasury Security Holdings
FRB Mortgaged Backed Security Holdings
Figure 11

Federal Debt to GDP
(Right Scale %)

Short and Long Term Interest Rates: Federal Funds,
3 Month Treasury, 5 Year Treasury,
10 Year Treasury & Mortgage Rates
(Left Scale %)

THE EFFECT OF RACE ON WORK COMMUTING TIME IN THE
PITTSBURGH, PA METROPOLITAN AREA

Robert R. Dunn
Washington & Jefferson College

Johnathan T. Williams
Washington & Jefferson College

ABSTRACT

This paper examines the impact of race on commuting time to work in the Pittsburgh, Pa metropolitan statistical area (MSA) from 2000 to 2010. Increased commuting time is a potential consequence of a spatial mismatch of employment growth, which has occurred predominantly in suburban areas, and minority location, which is predominantly centralized in metropolitan areas. The unit of analysis for this study is the census tract and other relevant factors that may explain commuting time for workers are considered, including earnings. The estimation results for both 2000 and 2010 show that residents in census tracts with a higher proportion of black residents face longer commutes to work. When the data are disaggregated into the central county in the MSA, Allegheny, and the surrounding six counties the results hold for Allegheny and the surrounding counties in 2000 and for Allegheny in 2010. The surrounding counties do not show evidence of increased commuting time in census tracts with a greater proportion of black residents in 2010. Finally calculations of the increased time spent commuting show that the impacts on a neighborhood can be substantial.
INTRODUCTION

Since John Kain’s seminal work (1964, 1968) on a potential spatial mismatch between urban minority residents and suburban job opportunities, many papers have examined numerous factors related to this hypothesis. The core of the spatial mismatch hypothesis is that there is a spatial distance between minorities that live predominantly in inner cities and employment opportunities that have increasingly moved to suburban locations. The belief is that minorities are unable to gain access to these employment opportunities for several reasons, including but not limited to discrimination in the housing market, increased search costs for potential jobs, and increased travel costs for jobs that are available. These issues and others have been examined in a vast empirical literature. This paper will focus on one specific aspect of this potential relationship between minority location and employment opportunity, the travel time to work for employed individuals. The hypothesis is that minorities who are employed face a longer commute to work than their white counterparts who are employed in the Pittsburgh metropolitan statistical area (MSA). Ihlaniyfeld (2006) identifies increased commuting costs as a main consequence of spatial mismatch.

The framework of the analysis is taken from Gabriel and Rosenthal (1996). The authors utilized a unique data set to examine the impact of race on commuting times at the household level. This paper performs a similar analysis at the neighborhood level (based on Census tracts) for the Pittsburgh MSA between 2000 and 2010. Analysis at the neighborhood level allows for new insight into the impacts of race on commuting times outside the individual household unit and the estimation results from Gabriel and Rosenthal (1996) provide a useful and reliable benchmark for the current estimates. The findings presented here indicate that after controlling for relevant variables residents in neighborhoods with larger shares of black residents in the Pittsburgh MSA face a longer commute to work than residents of neighborhoods with more white residents. The estimation results are both statistically and economically significant and are similar to the results obtained by Gabriel and Rosenthal (1996) at the household level. A percentage point increase in the proportion of the population that is black results in a 0.3 percent increase in travel time to work for residents of the neighborhood. Considering the average travel time for Pittsburgh workers was just over 25 minutes in 2000 and 2010 this would add approximately 4.5 seconds of commuting time each way or 9 seconds per day. However this analysis is performed at the neighborhood level and the average census tract in the analysis contains approximately 1,500 commuting workers. Based on the estimation results the average impact on a census tract is an additional two hours of commuting time each way or 4 hours of additional commute time each day. Neighborhoods with greater proportions of black residents are likely to face significantly longer commuting times to employment compared with neighborhoods with more white residents.

LITERATURE REVIEW

Gabriel and Rosenthal (1996) provide an overview of the many empirical studies focused on the spatial mismatch hypothesis in urban areas and their work is the basis for the current analysis. They argue that a common shortcoming of these studies is a failure to account for compensating differentials with respect to wages and house prices. While
minority workers could face longer work commute times, mobility of these workers should allow for a long-run adjustment of wages and house prices. The authors employ a unique set of data from the American Housing Survey in 1985 and 1989. The data is comprised of information on individual urban housing units and corresponding information on neighboring units. This allows the authors to use a fixed effects estimator to control for neighborhood characteristics. Their results show that black workers do endure longer work commutes after controlling for amenities, housing prices, and earnings.

A further extension looks at the mobility of households and finds that undercompensated black households, those that are not completely compensated for a longer commute through higher wages or lower house prices, are less likely to move. If long-run adjustment were to take place undercompensated black households should exhibit a greater propensity to move in an effort to capture high wages or lower house prices. By controlling each offsetting variable Gabriel and Rosenthal (1996) found that families, not individuals, are more likely to move if their compensations do not cover the extra time spent in commuting. However, a deeper look at the results showed that black families are less likely to move if undercompensated. This poses a serious implication; if black families are less likely to move if undercompensated, then some form of discrimination must be present that acts as a barrier to entry into the areas where compensation would improve.

The current analysis is closely related to the work of Gabriel and Rosenthal (1996) in that it attempts to determine if similar effects are found for larger neighborhoods (Census tracts) in metropolitan Pittsburgh. The authors note that neighborhood specific data is desirable for study such as this. Following their study earnings and house prices are controlled for along with other neighborhood demographic characteristics that could explain commute times.

Along with the consideration of higher wages or lower housing prices that may compensate workers for longer commutes access to automobiles and method of transportation are important factors when explaining longer commutes. Taylor and Ong (1995) utilize the American Housing Survey for 1977-78 and 1985 and conclude that the distance minorities must travel to work is not leading to longer commute times but rather it is lack of access to an automobile that leads to longer commute times. They also note that commute times for minorities were stable or declining at the time of their study. Hess (2005) reports similar results for the Buffalo – Niagara Falls area in New York. The author does not find evidence of spatial mismatch as low wage jobs and lower income black households are both centralized within the metropolitan area. The important point is that lower income black households lack automobile access. Based on these findings and the importance of transportation method the percent of workers that commute via automobile is controlled for in the empirical model.

Gottlieb and Lentnek (2001) examine the potential for spatial mismatch in Cleveland by investigating four types of neighborhoods: low-income predominantly black, low-income predominantly white, lower-middle-income predominantly black, and lower-middle-income predominantly white. They find that black households do have access to employment opportunities but do tend to experience longer commute times. They also report the possibility of job discrimination as suburban black households showed a greater tendency to work in the central city compared with white households. Related to this, Boustan and Margo (2009) focus on employment decentralization and a
reduction in work opportunities by utilizing the centrality of U.S. Postal Service processing facilities within metropolitan areas. The authors are able to show that blacks do show a preference for postal work and that the affect is stronger in more segregated areas; however the strength of the relationship is decreasing over time.

Martin (2001) constructs a spatial mismatch index that shows that the geographic distance between employment and black residents in metropolitan areas increased from 1970 to 1990. The impact of this increase was mitigated to some extent by population shifts but did not have a large enough impact to reduce the mismatch. In a separate paper Martin (2001b) notes that commuting subsidies could increase welfare for black residents who are not near employment centers, though not enough to lead to a net welfare gain. The largest impact on the issue of separation from employment is restricted residential mobility for black households. The current paper will provide evidence of considerable differences in commuting times to work in the Pittsburgh MSA. This may provide support for the idea of commuting subsidies proposed by Martin (2001b).

Stoll (2005) considers the relationship between metropolitan job sprawl and spatial mismatch for African Americans. In his study job sprawl is defined as the proportion of employment outside a 5-mile radius of the central city. The author then constructs a spatial mismatch index that provides a measure of how disconnected African Americans are from employment in approximately 300 U.S. metropolitan areas in 2000. This is based on the index of dissimilarity that is commonly used to measure neighborhood segregation. Theoretically it is possible that job sprawl could work to reduce spatial mismatch if the supply of housing increases quickly enough to reduce house prices and increase affordability for minorities. However, Stoll (2005) finds a strong positive relationship between metropolitan job sprawl and increased spatial mismatch for African Americans and Latinos.

The average job sprawl index measure for all metropolitan areas is 64.7 percent and the average African American spatial mismatch index measure is 53.5 percent for 2000. For metropolitan areas with more than 500,000 people the job sprawl index is 71.1 percent and the African American spatial mismatch index is 56.5 percent. The Pittsburgh metropolitan statistical area (MSA) has a job sprawl index of 64.2 percent and an African American spatial mismatch index of 55.6 percent for 2000; both are in line with the all metropolitan areas numbers. These statistics indicate that minorities in Pittsburgh are likely to face longer commuting time to places of employment and our results provide further evidence of this.

Although the empirical literature dealing with spatial mismatch and related issues is vast theoretical contributions have been relatively limited. Brueckner and Martin (1997) are an early exception to this as they construct an urban equilibrium framework that incorporates a spatial mismatch of employment and minority locations. The theoretical results obtained are consistent with those of Gabriel and Rosenthal (1996) described above and note that increased commute time is an impact of spatial mismatch faced by employed blacks. This is the exact issue examined in the current paper.

Gobillon, Selod, and Zenou (2007) focus on specific mechanisms that may lead to spatial mismatch. The authors argue that the emergence of theoretical models of spatial mismatch well after the presence of many empirical studies has led to a lack of specific mechanisms that may lead to spatial mismatch. They review numerous theoretical and empirical contributions and identify seven mechanisms. These include: (1) workers refusal of jobs with long commutes; (2) lower job search efficiency with greater distance
from employment; (3) less intensive job searches; (4) high search costs restrict area of job search; (5) employer discrimination against segregated communities; (6) refusal to hire employees with long commute times; (7) belief that white customers dislike contact with minority workers. The first and sixth mechanisms deal directly with the effects of increased commute times, the focus of this paper.

DATA AND METHODOLOGY

The data utilized in this study are measured at the census tract level and come from the 2000 Census of Population and from the 2010 American Community Survey five-year estimates. Due to changes in statistical methodology with respect to the variables included in this analysis it is not possible to combine the observations into a single dataset. Due to this the estimations for 2000 will be run separately from the estimations for 2010. While it is necessary to use caution when examining these two sets of results as the datasets cannot be directly compared due to the methodology change, the estimation should provide some insight to any changing patterns within the Pittsburgh metropolitan area. Census tracts are small subdivisions of counties that generally contain between 2,500 and 8,000 people with a mean population of approximately 4,000 residents according to the definition from the Census Bureau. By design they are intended to be homogenous based on population, economic, and housing characteristics. In addition to having common attributes these communities should have access to a similar bundle of amenities. This provides the basis for the focus on the impacts of race on commuting time at the neighborhood level. For the year 2000 the Pittsburgh metropolitan area had 719 census tracts with complete data for the empirical analysis. In addition to looking at race and commuting time at the metropolitan level, the observations are disaggregated into tracts within Allegheny County, the county containing the City of Pittsburgh, and tracts in the counties surrounding Allegheny that are part of the metropolitan area (Armstrong, Beaver, Butler, Fayette, Washington, Westmoreland). There are 415 tracts in Allegheny County and the remaining 304 are in the other six counties in the MSA. For 2010 there are 696 observations for the MSA with 388 in Allegheny County and 308 in the other six counties.

The census tract data in this analysis follow closely the variables examined by Gabriel and Rosenthal (1996) at the individual household level. The dependent variable in this study is the natural log of average travel time to work as measured in minutes taken commuting one way to work daily. This statistic was calculated by finding the aggregate minutes of travel time to work for all workers in the particular census tract and dividing that by the number of workers that commute to work in each tract. A longer commute time could be the result of greater distance from the place of employment or it could be the result of different methods of transportation. Based on this and evidence from previous research discussed in the literature review the method of commute will be controlled for in the estimation. This is measured as the percent of workers in the census tract that commute to work by car, truck, or van. The expectation is that increased access to an automobile within neighborhoods would likely reduce travel times due to increased travel speeds. However, it is also possible that neighborhoods with a higher percentage of automobile commuting are located farther from employment which could lead to longer commute times, particularly if commuting occurs during high traffic times.

Of the factors included to explain variation in the commute time to work, the variable of greatest interest is the percent of the population within a census tract that
identifies as black or African American. The expectation is that tracts with greater proportions of black residents will experience longer average commute times to work. It is important to note that the current analysis does not allow for identification of the specific reason that could cause longer commute times, although several relevant factors are included as control variables. However, the analysis can provide insight into one specific aspect of the spatial mismatch hypothesis. It also allows for consideration of the relationship between race and commuting time at the neighborhood level rather than the individual household level. In addition to analyzing the black population, the percent of the population that is Hispanic or Latino in each tract is also included to determine if these minorities encounter similar experiences. It should be noted that the Pittsburgh area has a considerably lower proportion of Hispanic or Latino residents, just over 1.25 percent of the population, than the nation as whole, which has a proportion of over 16 percent.

Gabriel and Rosenthal (1996) point out the importance of controlling for a compensating differential with respect to commute time. Essentially, workers may be willing to endure longer commutes if they are receiving compensation in the form of higher earnings. Similarly, workers facing longer commutes could be rewarded with lower housing prices. To account for this median household income for the previous year is included in each regression estimation. Given the high correlation between income and housing prices within census tracts these two variables cannot be included in the same estimation. While the results reported control for household income, estimations using housing prices produce similar results and are available from the author upon request.

Related to this, several neighborhood demographic characteristics are controlled for in the estimation. Educational attainment may be a determinant of commuting time if increasing levels of education allow for a larger set of employment opportunities and greater choice with regard to residential location. This is accounted for by including a measure of the percent of population within the census tract that has a college degree or higher. The expected influence of this variable is ambiguous as neighborhoods with greater education attainment could be located in more distant suburbs that require longer commutes or they could be closer to places of employment that result in shorter commuting times.

Two other neighborhood characteristics that may have relationships with commute time are the median age of the census tract and the percent of households that can be described as married couple families in each tract. The relative importance of employment and location varies over the lifecycle and it is therefore worthwhile to control for different age profiles within the census tracts. Families may show a preference for suburban living, particularly those with children. These families may be willing to endure longer commutes to obtain more land and larger homes and are not basing location decisions primarily on proximity to employment.

Finally, the mobility characteristics of census tracts could be important determinants of commute times. Given that employment is a major motivation for migration, tracts with higher percentages of immigrants, including households that move within a county and households that move from a different county, could have reduced travel times if new residents initially locate near their place of work. Two variables are included, one to measure the percent of the population that moved into the census tract and previously lived in the same county and one to measure the percent of the population that moved into the census tract and previously lived in a different county. For the data from the 2000 Census of Population, the same county variable is calculated as the percent of the
population that lived in a different residence within the same county in 1995 while the
different county variable measures the percent of the population living in the tract in 2000
that lived in a different residence in a different county in 1995. For the data from the 2010
American Community Survey five-year estimates the variables provide information on the
percent of the population that lived in a different residence the previous year. This measure
is then averaged for the observation for the five-years prior to 2010 to provide a five-year
estimate. The difference in methodology is noticeable in the summary statistics as the
mobility statistic is considerably higher for the 2000 dataset compared with the 2010
dataset.

The resulting log-level empirical model takes the following form and will be
estimated separately for the data from 2000 and from 2010:

\[ \ln(\text{time}) = \beta_0 + \beta_1 \text{Black} + \beta_2 \text{Hispanic} + \beta_3 \text{College} + \beta_4 \text{Age} + \]
\[ + \beta_5 \text{Married} + \beta_6 \text{Automobile} + \beta_7 \text{Samecounty} + \beta_8 \text{Differentcounty} + \]
\[ + \beta_9 \text{Income} + \epsilon_i \]

White’s test for heteroskedasticity shows that the null hypothesis of
homoskedasticity is rejected; therefore robust standard errors are used. Examination of
the variance inflation factors (VIFs) shows that multicollinearity is not a problem with the
data outside of the issue with median household income and median house value
discussed above (Wooldridge 2013). The mean VIF for each estimation ranged between
three and four with no single measure registering above eight. Summary statistics for
2000 and 2010 can be found in Tables 1 and 2 respectively.

ESTIMATION RESULTS

Table 3 displays the cross-sectional estimation results for the Pittsburgh
metropolitan area for 2000 and Table 4 presents disaggregated results for Allegheny
County and for the other six counties outside Allegheny but within the MSA. The results
indicate that census tracts with a greater proportion of black residents experience longer
commute times to work, approximately 0.4 percent longer for each percentage point
increase in the black population.

While the magnitude may seem small initially this would mean that a worker
within a census tract facing the average commute time for the sample (25 minutes) would
have their commute increased four percent each way (equivalent to 1 minute each way)
with a ten percentage point increase in the black population. Conversely, a tract with zero
black residents could be expected to experience a commute to work that is four percent
quicker as the mean black population for all tracts in the 2000 sample is just over ten
percent with a range of 0 to 98 percent. In the 2000 sample, 13 percent of the census
tracts have a black population of 20 percent or more and more than seven percent of the
tracts have a black population of 50 percent for greater. The implication is that a worker
in a census tract with a black population of 20 percent (a ten percentage point increase
above the mean) could expect to commute an additional two minutes per day, 10 minutes
per week, and eight hours per year, assuming a five day work week and 48 weeks of work
per year. In a tract with a black population of 50 percent (40 percentage points higher
than the mean) a worker is expected to face a 16 percent longer commute or an additional
eight minutes per day, 40 minutes per week, and 32 hours per year, the equivalent of four full days of work. This provides evidence that neighborhoods that contain a higher proportion of black residents are at a disadvantage with regard to commuting time to work and devote considerably more of their limited time to getting to and from employment.

The disaggregated results for 2000 are similar in that neighborhoods with more black residents face longer commutes in Allegheny County and in the surrounding counties, with the largest impact found in the surrounding counties.

The estimation results for the metropolitan area for 2010 are reported in Table 5 and the disaggregated results are reported in Table 6. The results show a similar pattern although the estimated impact of a percentage point increase in the population that is black drops to 0.27 which implies that a ten percentage point increase in black population increases commute times in a census tract by 2.7 percent.

In the 2010 sample the mean black population for all tracts increased to just more than 11 percent so a ten percentage point increase would again correspond to a black population of approximately 20 percent. As in the previous sample the range of black population within tracts is 0 to 98 percent. Slightly more than 16 percent of the tracts had a black population of 20 percent or greater and nearly eight percent had a black population of 50 percent or more. Workers within a census tract where the proportion of black residents is 20 percent could expect to commute an additional 1.35 minutes per day, 6.75 minutes per week, and 5.4 hours per year to work. In a tract where 50 percent of the population is black, workers would experience a 10.8 percent longer commute to work which translates to an additional 5.4 minutes per day, 27 minutes per week, and 21.6 hours per year. While the impacts of race on commute time appear to have decreased relative to the 2000 sample they are still clearly economically meaningful.

The disaggregated results also show some potential improvement on this issue. While neighborhoods with more black residents in Allegheny County still face longer commutes this relationship is not statistically significant in the surrounding counties for 2010. These results are in line with those reported by Gabriel and Rosenthal (1996) and provide evidence that the relationship they describe between race and commuting time at the household level is also present at the census tract, or neighborhood, level.

The other two factors that prove to have a statistically significant relationship with travel time to work in 2000 and 2010 are the prevalence of married couple families and mobility of residents. As expected census tracts with a greater proportion of married couple families are enduring longer commuting times to work, a finding that is consistent across the two time periods and the county disaggregation. This likely reflects a location decision that is primarily based on neighborhood attributes and not on proximity to employment. Neighborhoods that are attractive to families may well be in more residential locations and therefore result in a longer commute to employment centers. Also as expected the mobility of residents proves to be a determinant of commute time. For the year 2000 sample residents in the Pittsburgh metropolitan area and in Allegheny County that moved into a census tract from a different county experience reduced commute times while in the surrounding counties this is true only of residents who moved into a census tract from within the same county. The pattern is the same for the 2010 sample with the exception of the metropolitan area where all movers, those within the same county and those from a different county, have shorter commutes. The findings on mobility seem reasonable as immigrants to a metropolitan area may locate near employment initially and therefore have a shorter commute. In the surrounding counties it is movement within a
county that is associated with reduced commuting time and this may reflect secondary moves of households to more desirable locations once employment is established.

The additional demographic factors for the census tracts measuring education, age, and income are not statistically significant in the estimations with the exception of a higher proportion of college graduates being associated with shorter commutes in the 2010 sample in the counties outside of Allegheny. However, even in this instance it is only significant at the ten percent level. In addition, access to an automobile did not prove to have a statistically significant relationship with commute time. This could be due to the offsetting factors of automobiles allowing for potentially faster commutes compared with mass transit while also being more prevalent in more distant neighborhoods and suburbs that would tend to increase commute time.

CONCLUSION

This analysis has examined one issue related to the spatial mismatch hypothesis at the census tract level for the Pittsburgh metropolitan area for the years 2000 and 2010. It closely follows the work of Gabriel and Rosenthal (1996) that examined the relationship between race and commute to work time at the individual household level. While the unit of analysis and the geographic setting of the study differ from that of Gabriel and Rosenthal (1996) the findings are similar and take into account the importance of controlling for a compensating differential associated with longer commute times. Ultimately the results show that census tracts with a larger proportion of black residents experience longer commute times to work on a scale that is economically substantial. This result is statistically significant at the one percent level for the Pittsburgh metropolitan area for the 2000 and 2010 samples. When the samples are disaggregated by central and surrounding counties the results holds at the one percent level for Allegheny County and the surrounding counties in 2000 and for Allegheny County in 2010. Other important factors proved to be the percent of households characterized as married couple families and the mobility or residents in a census tract. The findings contribute to the large empirical literature on the impacts of race with regard to employment and employment opportunities and suggest that minorities in the Pittsburgh metropolitan area for the time periods studied endure longer commutes to places of employment.
ENDNOTES

1. The index of dissimilarity is a measure of how evenly distributed the white and black populations are among geographical units (e.g. census tracts) that constitute a metropolitan statistical area. The minimum value is zero, a situation where whites and blacks are evenly distributed across an MSA, while the maximum value is 100, a situation where whites and blacks and completely segregated within an MSA.

2. The job sprawl index is a measure of the proportion of jobs in an MSA that are located outside of a five mile radius centered on the central business district of the MSA. A higher measurement implies that jobs in that MSA are more decentralized. The African American spatial mismatch index is a combination of the index of dissimilarity and the job sprawl index in that it measures segregation between the African American population in an MSA and jobs in that MSA. A higher measure indicates that the African American population is located further from job opportunities. The measure can also be interpreted as the proportion of the African American population or the proportion of jobs that would have to relocate to create an even distribution of the population and jobs (Stoll 2005).

3. Taking the natural log of the average travel time to work and utilizing the log-level functional form allows for a nonlinear relationship between the proportion of a census tracts' population that is black and the average travel time to work as the estimates allow for a constant percentage effect as opposed to a constant unit effect (Wooldridge 2013). The unit effect of the proportion of the population that is black becomes larger as this proportion increases while the percentage effect remains constant. In addition, possible quadratic effects of the black population in a census tract (the variable of interest) were examined by estimating the percentage of the population that is black and its square. The estimates were statistically insignificant which makes it unlikely that substantial increasing or decreasing marginal effects are present.
REFERENCES


<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Travel Time</td>
<td>719</td>
<td>25.1</td>
<td>4.5</td>
<td>10.3</td>
<td>60.6</td>
</tr>
<tr>
<td>College</td>
<td>719</td>
<td>29.3</td>
<td>16.0</td>
<td>0</td>
<td>86.5</td>
</tr>
<tr>
<td>Median Age</td>
<td>719</td>
<td>39.8</td>
<td>4.6</td>
<td>14.1</td>
<td>54.6</td>
</tr>
<tr>
<td>Married Couple Families (%)</td>
<td>719</td>
<td>48.9</td>
<td>16.9</td>
<td>0</td>
<td>84.0</td>
</tr>
<tr>
<td>Automobile</td>
<td>719</td>
<td>84.7</td>
<td>14.3</td>
<td>18.8</td>
<td>100</td>
</tr>
<tr>
<td>Black Pop. (%)</td>
<td>719</td>
<td>10.4</td>
<td>21.1</td>
<td>0</td>
<td>98.2</td>
</tr>
<tr>
<td>Hispanic Pop. (%)</td>
<td>719</td>
<td>0.7</td>
<td>0.9</td>
<td>0</td>
<td>7.4</td>
</tr>
<tr>
<td>Median HH Income (000s)</td>
<td>719</td>
<td>38,157</td>
<td>16,075</td>
<td>7,042</td>
<td>147,298</td>
</tr>
<tr>
<td>Moved – Same County</td>
<td>719</td>
<td>23.1</td>
<td>7.9</td>
<td>0</td>
<td>57.2</td>
</tr>
<tr>
<td>Moved – Different County</td>
<td>719</td>
<td>9.9</td>
<td>7.6</td>
<td>0</td>
<td>64.4</td>
</tr>
</tbody>
</table>
### Table 2. Summary Statistics – Pittsburgh MSA, 2010

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Travel Time</td>
<td>696</td>
<td>25.2</td>
<td>4.2</td>
<td>10.3</td>
<td>44.6</td>
</tr>
<tr>
<td>College</td>
<td>696</td>
<td>26.9</td>
<td>17.1</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Median Age</td>
<td>696</td>
<td>42.0</td>
<td>5.9</td>
<td>15.0</td>
<td>64.2</td>
</tr>
<tr>
<td>Married Couple Families (%)</td>
<td>696</td>
<td>45.6</td>
<td>17.1</td>
<td>0</td>
<td>85.3</td>
</tr>
<tr>
<td>Automobile</td>
<td>696</td>
<td>84.0</td>
<td>14.7</td>
<td>9.2</td>
<td>99.2</td>
</tr>
<tr>
<td>Black Pop. (%)</td>
<td>696</td>
<td>11.2</td>
<td>20.6</td>
<td>0</td>
<td>98.4</td>
</tr>
<tr>
<td>Hispanic Pop. (%)</td>
<td>696</td>
<td>1.3</td>
<td>1.8</td>
<td>0</td>
<td>15.3</td>
</tr>
<tr>
<td>Median HH Income (000s)</td>
<td>696</td>
<td>48,568</td>
<td>20,965</td>
<td>10,902</td>
<td>178,542</td>
</tr>
<tr>
<td>Moved – Same County</td>
<td>696</td>
<td>8.4</td>
<td>6.1</td>
<td>0</td>
<td>43.0</td>
</tr>
<tr>
<td>Moved – Different County</td>
<td>696</td>
<td>3.8</td>
<td>4.4</td>
<td>0</td>
<td>50.0</td>
</tr>
</tbody>
</table>
Table 3. Estimation Results – Pittsburgh MSA, 2000

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>T-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>College</td>
<td>-0.00028</td>
<td>0.32</td>
</tr>
<tr>
<td>Median Age</td>
<td>0.00016</td>
<td>0.07</td>
</tr>
<tr>
<td>Married Couple Families (%)</td>
<td>0.00437</td>
<td>2.50**</td>
</tr>
<tr>
<td>Automobile</td>
<td>0.00072</td>
<td>0.60</td>
</tr>
<tr>
<td>Black Pop. (%)</td>
<td>0.00399</td>
<td>8.24***</td>
</tr>
<tr>
<td>Hispanic Pop. (%)</td>
<td>0.00434</td>
<td>0.55</td>
</tr>
<tr>
<td>Median HH Income (000s)</td>
<td>-0.00056</td>
<td>0.43</td>
</tr>
<tr>
<td>Moved – Same County</td>
<td>-0.00258</td>
<td>1.20</td>
</tr>
<tr>
<td>Moved – Different County</td>
<td>-0.00439</td>
<td>2.66***</td>
</tr>
<tr>
<td>Constant</td>
<td>3.014</td>
<td>16.35***</td>
</tr>
</tbody>
</table>

N = 719
F-Stat = 27.82
R-squared = 0.25
Table 4. Estimation Results – County Disaggregation, 2000

<table>
<thead>
<tr>
<th>Variable</th>
<th>Allegheny County</th>
<th>Other Counties</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>T-Statistic</td>
</tr>
<tr>
<td>College</td>
<td>0.00071</td>
<td>0.73</td>
</tr>
<tr>
<td>Median Age</td>
<td>-0.00204</td>
<td>0.72</td>
</tr>
<tr>
<td>Married Couples (%)</td>
<td>0.00423</td>
<td>1.96**</td>
</tr>
<tr>
<td>Automobile</td>
<td>0.00151</td>
<td>1.33</td>
</tr>
<tr>
<td>Black Pop. (%)</td>
<td>0.00324</td>
<td>6.48***</td>
</tr>
<tr>
<td>Hispanic Pop. (%)</td>
<td>-0.00104</td>
<td>0.11</td>
</tr>
<tr>
<td>Median HH Income (000s)</td>
<td>-0.00124</td>
<td>0.82</td>
</tr>
<tr>
<td>Moved – Same County</td>
<td>-0.00156</td>
<td>0.58</td>
</tr>
<tr>
<td>Moved – Different County</td>
<td>-0.00664</td>
<td>2.68***</td>
</tr>
<tr>
<td>Constant</td>
<td>2.978</td>
<td>13.04***</td>
</tr>
</tbody>
</table>

N = 415; 304
F-Stat = 18.53; 14.74
R-squared = 0.32; 0.28
Table 5. Estimation Results – Pittsburgh MSA, 2010

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>T-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>College</td>
<td>-0.00048</td>
<td>0.61</td>
</tr>
<tr>
<td>Median Age</td>
<td>-0.00041</td>
<td>0.31</td>
</tr>
<tr>
<td>Married Couple Families (%)</td>
<td>0.00398</td>
<td>4.47***</td>
</tr>
<tr>
<td>Automobile</td>
<td>0.00015</td>
<td>0.15</td>
</tr>
<tr>
<td>Black Pop. (%)</td>
<td>0.00265</td>
<td>5.11***</td>
</tr>
<tr>
<td>Hispanic Pop. (%)</td>
<td>-0.00013</td>
<td>0.04</td>
</tr>
<tr>
<td>Median HH Income (000s)</td>
<td>0.00020</td>
<td>0.24</td>
</tr>
<tr>
<td>Moved – Same County</td>
<td>-0.00324</td>
<td>2.12**</td>
</tr>
<tr>
<td>Moved – Different County</td>
<td>-0.00712</td>
<td>3.13***</td>
</tr>
<tr>
<td>Constant</td>
<td>3.060</td>
<td>30.91***</td>
</tr>
</tbody>
</table>

N = 696  
F-Stat = 24.20  
R-squared = 0.23
Table 6. Estimation Results – County Disaggregation, 2010

<table>
<thead>
<tr>
<th>Variable</th>
<th>Allegheny County</th>
<th>Coefficient</th>
<th>T-Statistic</th>
<th>Other Counties</th>
<th>Coefficient</th>
<th>T-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>College</td>
<td>-0.00021</td>
<td>0.25</td>
<td>-0.00296</td>
<td>1.90*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median Age</td>
<td>-0.00019</td>
<td>0.12</td>
<td>-0.00187</td>
<td>0.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married Couple Families (%)</td>
<td>0.00372</td>
<td>3.59***</td>
<td>0.00321</td>
<td>1.90*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Automobile</td>
<td>0.00102</td>
<td>0.92</td>
<td>0.00300</td>
<td>1.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black Pop. (%)</td>
<td>0.00224</td>
<td>4.32***</td>
<td>0.00225</td>
<td>1.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic Pop. (%)</td>
<td>0.00462</td>
<td>1.14</td>
<td>-0.00681</td>
<td>1.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median HH Income (000s)</td>
<td>-0.06058</td>
<td>0.66</td>
<td>0.00280</td>
<td>1.92*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moved – Same County</td>
<td>0.00001</td>
<td>0.02</td>
<td>-0.00871</td>
<td>3.06***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moved – Different County</td>
<td>-0.00909</td>
<td>3.27***</td>
<td>-0.00181</td>
<td>0.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>3.007</td>
<td>27.29***</td>
<td>2.843</td>
<td>9.78***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N = 388; 308
F-Stat = 15.70; 16.74
R-squared = 0.24; 0.32
FACTORS CONTRIBUTING TO HIGHER STUDENT ACHIEVEMENT IN PENNSYLVANIA CHARTER SCHOOLS

Kerry A. Adzima
Penn State Erie, The Behrend College

ABSTRACT

Charter schools in Pennsylvania serve over 120,000 K-12 students or approximately 6.7 percent of all students in the state. Recent exam scores in Pennsylvania reveal that some charter schools are outperforming traditional public schools making it increasingly important that researchers focus on the factors contributing to these charter school success stories. In this study, I find that higher attendance rates, more certified teachers, lower student to teacher ratios, and smaller enrollments overall have a positive influence on student achievement. In addition, charter schools stressing academic achievement and parental involvement within their mission statements also observe better outcomes.

INTRODUCTION

Charter schools are publicly funded elementary or secondary schools operated independently of the local school board. These schools have the freedom to implement their own curriculum and educational philosophy and are unbound from some of the rules, regulations, and statutes that apply to other public schools. However, charter schools are still held accountable for producing certain results, and must abide by the standards and procedures set forth in each school’s charter. Unlike most traditional public schools, poor performance can result in the closing of a charter school by the school’s authorizers.

By law, charter schools cannot charge tuition and are financed by local, state and federal tax dollars based on enrollment. Charter schools cannot discriminate on any basis but may give preference to those students who live in the district where the charter school is located. In cases where there is not enough space to accommodate every child that applies, students are selected through a random lottery.

In the late 1980’s, Philadelphia, Pennsylvania was the first city in the United States to experiment with developing charter schools by creating them inside some of its
public schools. When the experiment was deemed successful, school officials in other cities refined this approach and tried to implement the idea in their communities. In Minnesota, where the first charter school law was written in 1991 and the first 'official' charter school was opened in 1992, educators developed three basic values for these schools: opportunity, choice and responsibility for results. Over the past two decades, charter schools have subsequently become one of the most debated forms of school choice in both the United States and in other countries across the globe.

One of the biggest debates among charter school supporters and critics deals with student test scores. Supporters claim that charter school students are performing at least as well and in some cases better than their traditional public school counterparts, while critics claim just the opposite. Research tends to reveal mixed results depending on the sample used and methodology employed, which helps to fuel these already heated discussions.

In Pennsylvania, the 2011-12 results for the Pennsylvania System of School Assessments (PSSA) revealed noticeable differences between traditional district schools and charter schools. For example, 49.85 percent of traditional school districts statewide made Adequate Yearly Progress (AYP) compared to 58.97 percent of Pennsylvania's charter schools. The cities of Pittsburgh and Philadelphia saw the most prominent differences between traditional and charter schools. In Pittsburgh, 12 percent of the traditional district schools made AYP for the 2011-2012 school year compared to 57 percent of charter schools. Thirteen percent of the traditional schools in the Philadelphia School District achieved AYP status compared to 61 percent of charter schools.1

In contrast to these findings, a 2011 study by the Center for Research on Education Outcomes (CREDO) at Stanford University found that students in Pennsylvania charter schools on average lost 29 days of learning in reading compared with their peers in traditional Pennsylvania public schools, and in math they fell behind by 50 days. The study compared charter school students with a computer-generated student representing traditional public school students. Each charter school student was matched with a "virtual twin" taking into account test scores as well as characteristics such as race, gender, and eligibility for free or reduced lunches. Critics of this study claim that the generalizations made about Pennsylvania charter schools in CREDO'S report are flawed because the research methodology did not make use of randomized control trials.2

Even though the debate over test scores is a long way from being settled, it is hard to argue against the fact that there are some charter schools making a positive impact on the lives of families in the K-12 education system. The evidence of this can be found most predominantly by looking at the number of students who are still waiting to be selected and admitted to their local charter school.

Although there is a vast literature comparing traditional public school and charter school test scores, there are relatively few studies that just focus on the factors that influence charter school performance. In this paper, I contribute to the former strand of literature by examining the various factors that play a role in the academic success of charter schools in Pennsylvania. Using annual report card data, I find that higher attendance rates, more certified teachers, lower student to teacher ratios, and smaller enrollments overall have a positive influence on student achievement. In addition, I find that charter school missions which stress academic performance and an active parental role in the education process also result in higher student outcomes.
The following sections of the paper present a review of the literature, a description of the data, the empirical model and regression results, and a conclusion with my thoughts on future studies relating to charter school performance.

LITERATURE REVIEW

The education production function has long been a standard method for evaluating the inputs that affect student achievement going back at least as far as the landmark Coleman Report released in 1966. For example, Hanushek (1996), Hedges and Greenwald (1996) and Krueger (1999) all use production functions to analyze how financial resources impact student achievement. The results of these studies show conflicting results with Hanushek (1996) observing a negligible impact and the latter two seeing positive effects.

Over time as data has become more accessible, the variety of inputs included as part of the education production function has increased allowing researchers to focus on characteristics such as class size (Angrist and Lavy 1999 and Hoxby 2000), peer effects (Hanushek et al. 2003), and teacher characteristics (Rivkin, Hanushek, and Kain 2005). Family background has also been cited as an important part of the education production function. More specifically, parental education and income have been included as inputs in studies by Murnane, Maynard, and Ohls (1981), Hanushek (1992), Ehrenberg and Brewer (1994), Ferguson and Ladd (1996), and Goldhaber and Brewer (1997). Houtenville and Conway (2008) find the impact of parental effort in the production of student achievement to be consistently associated with higher student outcomes.

When specifically examining just the charter school literature, a meta-analysis performed by Betts and Tang (2011) revealed mixed results when looking at production functions and other methods that compare traditional public schools and charter schools using math and reading scores. In some grades and locations, charter schools outperform or perform basically the same as traditional public schools and in others they underperform. However, a few exceptions include reading scores at the elementary level and both math and reading in middle schools where there is no evidence of negative effects of charter schools and, in some cases, confirmation of positive effects. The authors also noted that studies focusing on urban areas tended to find larger gains in academic achievement for charter students compared to studies that looked at a wider range of locations. More specifically, charter schools in New York City and Boston produced larger achievement gains compared to charter schools in most other areas.

The nonrandom assignment of charter school policies and characteristics to each school makes it hard to determine why some charter schools succeed and others do not. Betts and Hill (2006) found that many of the factors that might affect achievement are difficult and sometimes impractical to measure. There are a few studies however that have revealed some preliminary conclusions regarding what factors may contribute to increasing test scores in charter schools. For example, Hoxby, Kang and Murarka (2009) analyzed New York City charters and found some support for higher achievement when the schools used a longer school day/year, time devoted to reading instruction, a school mission statement emphasizing academic achievement, “small rewards/small punishment” discipline policies, and teacher pay systems not based exclusively on experience. Dobbie and Fryer (2011) also looked at New York City charters and estimated that 50 percent of the variation in school effectiveness was explained by extended time, high-dosage tutoring, the use of data to inform instruction, frequent formal or informal teacher feedback, and a
self-reported focus on academic achievement. Angrist, Pathak, and Walters (2011) found longer school hours and a self-reported adherence to a “no excuses” philosophy toward student achievement and behavior to be important characteristics influencing charter school success. Similarly, Furgesson, et al. (2011) revealed that charter management organizations using longer school days, comprehensive behavior policies, and rigorous teacher coaching tended to get improved results. Finally, Gleason et al. (2010) found some support of higher achievement in charter middle schools with smaller enrollments and those using ability grouping.

DATA

Across the nation, the number of charter schools and students served has risen to 5,997 and 2,278,388 respectively. In addition, there are approximately 920,000 students on waitlists hoping for a chance to enroll in a charter school. Table 1 displays the growth in charter schools and students served in the United States as a whole.3 The table shows that the percent of charter schools to all public schools has risen from 1.7 in 1999-00 to 6.3 in 2012-13 and the percent of charter school students to all public school students has increased from 0.7 in 1999-00 to 4.6 in 2012-13.

In Pennsylvania, where charter school law was first written in 1997, the number of charter schools has risen from 47 in 1999-00 to 175 in 2012-13, and the number of students served has risen from 11,413 to 120,102. The number of students on Pennsylvania charter school waiting lists is currently estimated to be 44,000. Table 2 displays the growth in charter schools and students served in Pennsylvania from 1999-00 through 2012-13. This table shows that Pennsylvania has been following the same trend in charter school growth observed throughout the nation. The percent of charter schools to all public schools in Pennsylvania has risen from 1.5 in 1999-00 to 5.7 in 2012-13 and the percent of charter school students to all public school students in Pennsylvania has increased from 0.6 in 1999-00 to 6.7 in 2012-13.4

Data for this study were collected from the Pennsylvania Department of Education (PDE). The two dependent variables were constructed using the Pennsylvania System of School Assessment (PSSA) standards-based, criterion-referenced test used to measure a student’s academic performance. All Pennsylvania students in grades 3 through 8 and grade 11 are assessed in reading and mathematics annually. Students are classified as achieving at one of four levels based on their performance on this test. The categories are “advanced,” “proficient,” “basic,” or “below basic.” Because Pennsylvania is required to meet certain targets on the PSSA according to the federal No Child Left Behind Act (78% proficient or advanced in math and 81% proficient or advanced in reading in 2011-12), I used these categories in my analysis. More specifically, the dependent variables include the percent of students in grades 3 through 8 and grade 11 who are classified as advanced and proficient in reading (ACHIEVE_READING) and mathematics (ACHIEVE_MATH) for the 2010-11 and 2011-12 school years.

The independent variables were collected from each charter school’s annual report card for each year. The relevant variables that schools are required to report include: the percent of certified teachers, the percent of students eligible for a free lunch, the percent of female students, the percent of minority students, the number of students enrolled in special education, the date the school opened, the student to teacher ratio, enrollment numbers, the attendance rate, and subsidies. Each charter school also provides its mission statement.
which I used to create a set of dummy variables to examine how a school’s philosophy influences academic performance. Both a time dummy and a dummy variable that captures whether the school’s grade span includes high school students (11th grade) were included as control variables as well. Finally, two control variables were used to capture the effects of the surrounding schools: the average district score for which the charter school is located (DISTRICT_ACHIEVE_READING) or (DISTRICT_ACHIEVE_MATH) and the percent of minority student in the district for which the charter school is located (DISTRICT_MINORITY).

After deleting the schools with missing observations the remaining sample included 80 charter schools and 787 observations. The descriptive statistics are reported in Table 3. On average, 61.5 percent of charter students in this sample are scoring at the advanced and proficient level in reading and 66.2 percent at the advanced and proficient level in mathematics. These numbers are slightly higher relative to the home districts which see 57.1 and 61.9 percent of their students scoring at the advanced and proficient level in reading and mathematics respectively. Charter schools in this sample have been open an average of 9.6 years; have an average total enrollment of 629.6 students, a student to teacher ratio of 17.4 to 1, a 90.6 percent teacher certification rate, and receive approximately $9,236.59 in subsidies per pupil. The students in this sample have a 94.3 percent attendance rate, with 68.4 percent of them eligible for a free lunch, 70.6 percent being of minority status, 52.0 percent being female, and 13.4 percent classified as having special needs, and spend approximately 1,107 hours in the classroom each school year.

EMPIRICAL METHODS AND RESULTS

The empirical model follows from a standard education production function which assumes that student achievement is a function of school and home district characteristics and resources, and student demographics. The theoretical model can be described as

\[
\text{ACHIEVEMENT}_{c,t} = f (\text{SCHOOL CHARACTERISTICS}_{c,t}, \text{SCHOOL RESOURCES}_{c,t}, \text{STUDENT DEMOGRAPHICS}_{c,t}, \text{DISTRICT CHARACTERISTICS}_{c,t})
\]

where all variables are for charter school \( c \) at time \( t \) and defined as follows:

\( \text{ACHIEVE_READING}_{c,t} \) and \( \text{ACHIEVE_MATH}_{c,t} \) = percent of students who are classified as advanced and proficient in reading and mathematics respectively on the PSSA

\( \text{SCHOOL CHARACTERISTICS}_{c,t} \) = number of years the charter school has been in operation (YEARS_OPEN), total charter school enrollment (ENROLLMENT), student to teacher ratio (STUD_TEACH_RATIO), student attendance rate (ATTEND_RATE), a time dummy (YEAR), a dummy to capture the school’s that include 11th grade (HIGH_SCHOOL), and seven dummy variables to capture various aspects of the school’s mission.

The dummies include (1) stressing academic achievement (ACADEMIC), (2) emphasizing parental involvement as an important element of student success (PARENTS), (3) a belief that students should take responsibility for their own learning and actions (RESPONSIBLE), (4) a focus on high quality teachers (TEACHERS), (5) promising an individualized education (INDIVIDUALIZED), (6) a special emphasis on technology (TECHNOLOGY), and (7) using project-based learning techniques (PROJECT_BASED).
SCHOOL RESOURCES$\varepsilon_{it} =$ subsidies per pupil (SUBSIDY)$^3$, percent of certified teachers (TEACH_CERTIFIED)

STUDENT DEMOGRAPHICS$\varepsilon_{it} =$ percent of students eligible for a free lunch (FREE_LUNCH), percent of special education students (SPECIAL_EDU), percent of minority students (MINORITY), and percent of female students (FEMALE).

DISTRICT CHARACTERISTICS$\varepsilon_{it} =$ percent of students who are classified as advanced and proficient in reading and mathematics respectively on the PSSA in the charter school’s home district (DISTRICT_ACHIEVE_READING) and (DISTRICT_ACHIEVE_MATH) and the percent of minority student in the district school for which the charter school is located (DISTRICT_MINORITY).

The following equation represents the models used in the analysis:

$$\text{ACHIEVE\_READING}_{it} = \beta_0 + \beta_1\text{YEARS\_OPEN}_{it} + \beta_2\text{STUD\_TEACH\_RATIO}_{it} + \beta_3\text{TEACH\_CERTIFIED}_{it} + \beta_4\text{ATTEND\_RATE}_{it} + \beta_5\text{TIME}_{it} + \beta_6\text{HIGH\_SCHOOL}_{it} + \beta_7\text{ENROLLMENT}_{it} + \beta_8\text{FREE\_LUNCH}_{it} + \beta_9\text{INSTRUCT\_HOURS}_{it} + \beta_{10}\text{SPECIAL\_EDU}_{it} + \beta_{11}\ln\text{SUBSIDY}_{it} + \beta_{12}\text{MINORITY}_{it} + \beta_{13}\text{FEMALE}_{it} + \beta_{14}\text{ACADEMIC}_{it} + \beta_{15}\text{PARENTS}_{it} + \beta_{16}\text{TEACHERS}_{it} + \beta_{17}\text{RESPONSIBLE}_{it} + \beta_{18}\text{INDIVIDUALIZED}_{it} + \beta_{19}\text{TECHNOLOGY}_{it} + \beta_{20}\text{PROJECT\_BASED}_{it} + \beta_{21}\text{DISTRICT\_ACHIEVE\_READING}_{it} + \beta_{22}\text{DISTRICT\_MINORITY}_{it} + \varepsilon_{it},$$

where $\varepsilon_{it}$ represents the error term.

A second equation is also used with the dependent variable ACHIEVE_MATH$\varepsilon_{it}$ and the same independent variables (other than DISTRICT_ACHIEVE_MATH) used in equation (1).

A random effects specification is used over an OLS pooled regression because of the high values found when performing the Breusch-Pagan Lagrange multiplier test. The results of the random effects model using cluster-robust standard errors are reported in Table 4. In columns 2 and 3 the coefficients and standard errors are reported using the percent of students who are advanced and proficient in reading as the dependent variable and in columns 4 and 5 the results come from using the percent of students who are advanced and proficient in mathematics as the dependent variable.

In each model, the percent of certified teachers and the percent of female students are found to be positive and statistically significant. Holding all other variables constant, a ten percentage point increase in the number of certified teachers leads to between a 2.1 percentage point and 2.7 percentage point increase in the number of students who are advanced and proficient, and a ten percentage point increase in the number of female students leads to between a 4.8 percentage point and 5.2 percentage point increase in the number of students who are advanced and proficient. The expenditure per pupil (subsidies) is also positive and statistically significant in both models. A one percentage point increase in expenditures leads to an approximate 2.3 percentage point increase in the number of students who are advanced and proficient. On the reading exam only, for every additional year a school is open, the number of students who are advanced and proficient increases by 0.74 percentage points and a ten percentage point increase in the attendance rate leads to an 8.30 percentage point increase in the number of students who are advanced and proficient in this subject, all else equal.
Both regressions also revealed that a higher number of students eligible for a free lunch and a higher number of special education students leads to a decrease in the number of students who are advanced and proficient in each subject. More specifically, a ten percentage point increase in the number of students eligible for a free lunch leads to an approximate 2.9 percentage point decrease in the number of students who are advanced and proficient, and a ten percentage point increase in the number of special education students reveals between a 3.1 and 5.8 percentage point decrease in the number of students who are advanced and proficient. A ten percentage point increase in the number of minority students indicates about a 1.0 percentage point decrease in the number of students who are advanced and proficient in the reading specification only.

According to the math specification, schools that stress academic achievement start out with more students (5.7 percentage points higher) at the advanced and proficient level while schools that stress an individualized education and student responsibility start out with less (4.4 and 7.0 percentage points lower, respectively). On the reading assessment, schools that stress student responsibility start out with fewer students at the advanced and proficient level (7.4 percentage points lower), while all other dummy variables associated with the mission statement were found to be statistically insignificant. In both models, a dummy to account for charter schools that have a high school score (11th grade) revealed that these schools start out with fewer students at the advanced and proficient level (between 8.6 and 11.0 percentage points lower.)

CONCLUSION

Although not all charter schools have been successful, the growth in the overall number of schools, students served, and the sheer number of students just waiting for a chance to enroll in a charter school provides a fairly strong argument in favor of analyzing these schools more closely. While the literature on charter schools is abundant when comparing traditional public schools and charter schools, there is still a void when it comes to studying the success and failures of charter schools themselves. In this paper, I contribute to the literature by providing a study that examines the factors that have an impact on the academic achievement of charter school students in Pennsylvania. Not only did I control for the factors that are normally studied when analyzing tests scores such as school resources and student demographics, I also examined the mission statements of each individual school. Consistent with other studies, I found that schools stressing academic achievement have initially higher student outcomes in mathematics, while schools that stress student responsibility tend to start out at lower levels of academic achievement.

Both policy makers and educators should find the results of this study useful when examining their own local charter schools and when developing new schools in the future. Additional studies of this kind should be undertaken continuing to analyze what inputs are contributing to the charter school success stories in other states. By developing a set of characteristics that contribute to higher achievement, current and future charter school authorizers can implement those factors they have control over and be better able to create schools that foster improved student outcomes.
Endnotes
1The statistics regarding Adequate Yearly Progress were taken from http://pacharters.org/2012/10/charter-schools-outperforming-district-schools-in-cities-statewide/.  
2The CREDO report is available at http://credo.stanford.edu/research-reports.html.  
3Charter school statistics were obtained from the National Alliance for Public Charter Schools, The Public Charter Schools Dashboard, available at http://dashboard.publiccharters.org/dashboard/home  
4Ibid  
5In the regression model the log of subsidy was used for easier interpretation, however in both cases, the variable was found to be statistically insignificant. The regression was also run without logging subsidy but the variable was still found to be statistically insignificant and had only a negligible impact on the overall model.
REFERENCES


<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Schools</th>
<th>% of Charter Schools to All Public Schools</th>
<th>Number of Students</th>
<th>% of Charter School Students to All Public School Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999-00</td>
<td>1,542</td>
<td>1.7</td>
<td>349,714</td>
<td>0.7</td>
</tr>
<tr>
<td>2000-01</td>
<td>1,941</td>
<td>2.1</td>
<td>458,664</td>
<td>1.0</td>
</tr>
<tr>
<td>2001-02</td>
<td>2,313</td>
<td>2.5</td>
<td>580,029</td>
<td>1.2</td>
</tr>
<tr>
<td>2002-03</td>
<td>2,559</td>
<td>2.7</td>
<td>666,038</td>
<td>1.4</td>
</tr>
<tr>
<td>2003-04</td>
<td>2,959</td>
<td>3.1</td>
<td>789,479</td>
<td>1.6</td>
</tr>
<tr>
<td>2004-05</td>
<td>3,383</td>
<td>3.6</td>
<td>897,643</td>
<td>1.8</td>
</tr>
<tr>
<td>2005-06</td>
<td>3,689</td>
<td>3.9</td>
<td>1,019,620</td>
<td>2.1</td>
</tr>
<tr>
<td>2006-07</td>
<td>3,999</td>
<td>4.2</td>
<td>1,165,200</td>
<td>2.4</td>
</tr>
<tr>
<td>2007-08</td>
<td>4,299</td>
<td>4.4</td>
<td>1,293,560</td>
<td>2.6</td>
</tr>
<tr>
<td>2008-09</td>
<td>4,640</td>
<td>4.8</td>
<td>1,445,954</td>
<td>2.9</td>
</tr>
<tr>
<td>2009-10</td>
<td>4,913</td>
<td>5.0</td>
<td>1,627,403</td>
<td>3.3</td>
</tr>
<tr>
<td>2010-11</td>
<td>5,259</td>
<td>5.4</td>
<td>1,812,889</td>
<td>3.7</td>
</tr>
<tr>
<td>2011-12</td>
<td>5,618</td>
<td>5.8</td>
<td>2,051,809</td>
<td>4.2</td>
</tr>
<tr>
<td>2012-13</td>
<td>5,997</td>
<td>6.3</td>
<td>2,278,388</td>
<td>4.6</td>
</tr>
<tr>
<td>Year</td>
<td>Number of Schools</td>
<td>% of Charter Schools to All Public Schools</td>
<td>Number of Students</td>
<td>% of Charter School Students to All Public School Students</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------</td>
<td>--------------------------------------------</td>
<td>-------------------</td>
<td>----------------------------------------------------------</td>
</tr>
<tr>
<td>1999-00</td>
<td>47</td>
<td>1.5</td>
<td>11,413</td>
<td>0.6</td>
</tr>
<tr>
<td>2000-01</td>
<td>65</td>
<td>2.0</td>
<td>18,981</td>
<td>1.0</td>
</tr>
<tr>
<td>2001-02</td>
<td>77</td>
<td>2.4</td>
<td>28,453</td>
<td>1.6</td>
</tr>
<tr>
<td>2002-03</td>
<td>91</td>
<td>2.9</td>
<td>32,862</td>
<td>1.8</td>
</tr>
<tr>
<td>2003-04</td>
<td>102</td>
<td>3.2</td>
<td>41,114</td>
<td>2.3</td>
</tr>
<tr>
<td>2004-05</td>
<td>109</td>
<td>3.3</td>
<td>48,212</td>
<td>2.6</td>
</tr>
<tr>
<td>2005-06</td>
<td>116</td>
<td>3.6</td>
<td>55,630</td>
<td>3.0</td>
</tr>
<tr>
<td>2006-07</td>
<td>119</td>
<td>3.6</td>
<td>59,976</td>
<td>3.2</td>
</tr>
<tr>
<td>2007-08</td>
<td>125</td>
<td>3.9</td>
<td>67,275</td>
<td>3.8</td>
</tr>
<tr>
<td>2008-09</td>
<td>127</td>
<td>3.9</td>
<td>73,390</td>
<td>4.1</td>
</tr>
<tr>
<td>2009-10</td>
<td>134</td>
<td>4.1</td>
<td>79,535</td>
<td>4.5</td>
</tr>
<tr>
<td>2010-11</td>
<td>145</td>
<td>4.5</td>
<td>90,954</td>
<td>5.2</td>
</tr>
<tr>
<td>2011-12</td>
<td>162</td>
<td>5.1</td>
<td>105,591</td>
<td>6.0</td>
</tr>
<tr>
<td>2012-13</td>
<td>175</td>
<td>5.7</td>
<td>120,102</td>
<td>6.7</td>
</tr>
<tr>
<td>Variables</td>
<td>Mean</td>
<td>Standard Deviation</td>
<td>Minimum</td>
<td>Maximum</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------</td>
<td>--------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>DISTRICT ACHIEVE</td>
<td>57.13</td>
<td>16.30</td>
<td>22.00</td>
<td>95.8</td>
</tr>
<tr>
<td>READING</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DISTRICT ACHIEVE</td>
<td>61.88</td>
<td>14.87</td>
<td>22.90</td>
<td>97.00</td>
</tr>
<tr>
<td>MATH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACHIEVE READING</td>
<td>58.72</td>
<td>19.45</td>
<td>5.90</td>
<td>100.00</td>
</tr>
<tr>
<td>ACHIEVE MATH</td>
<td>66.16</td>
<td>21.42</td>
<td>0</td>
<td>100.00</td>
</tr>
<tr>
<td>YEARS OPEN</td>
<td>9.62</td>
<td>3.78</td>
<td>1.00</td>
<td>15.00</td>
</tr>
<tr>
<td>STUD TEACH RATIO</td>
<td>17.40</td>
<td>5.81</td>
<td>7.30</td>
<td>32.00</td>
</tr>
<tr>
<td>TEACH CERTIFIED</td>
<td>90.62</td>
<td>10.11</td>
<td>49.00</td>
<td>100.00</td>
</tr>
<tr>
<td>ATTEND RATE</td>
<td>94.29</td>
<td>2.11</td>
<td>87.90</td>
<td>98.00</td>
</tr>
<tr>
<td>ENROLLMENT</td>
<td>629.56</td>
<td>456.85</td>
<td>90.00</td>
<td>2664.00</td>
</tr>
<tr>
<td>FREE LUNCH</td>
<td>68.39</td>
<td>26.75</td>
<td>0.00</td>
<td>100.00</td>
</tr>
<tr>
<td>INSTRUCT HOURS</td>
<td>1107.59</td>
<td>104.55</td>
<td>913.00</td>
<td>1517.00</td>
</tr>
<tr>
<td>SPECIAL EDU</td>
<td>13.44</td>
<td>6.83</td>
<td>0.00</td>
<td>40.95</td>
</tr>
<tr>
<td>SUBSIDY</td>
<td>9236.59</td>
<td>1409.03</td>
<td>7,400.00</td>
<td>13,643.00</td>
</tr>
<tr>
<td>MINORITY</td>
<td>70.58</td>
<td>33.40</td>
<td>1.41</td>
<td>100.00</td>
</tr>
<tr>
<td>FEMALE</td>
<td>52.01</td>
<td>5.18</td>
<td>33.33</td>
<td>70.18</td>
</tr>
<tr>
<td>DISTRICT MINORITY</td>
<td>64.97</td>
<td>29.82</td>
<td>2.45</td>
<td>97.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Random Effects Results Using Percent of Students Classified as Advanced and Proficient in Reading and Mathematics for 2010-11 and 2011-12

<table>
<thead>
<tr>
<th>Variables</th>
<th>Reading</th>
<th>Mathematics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Standard Error</td>
</tr>
<tr>
<td>YEAR OPEN</td>
<td>0.737**</td>
<td>0.323</td>
</tr>
<tr>
<td>STUD TEACH RATIO</td>
<td>-0.228</td>
<td>0.250</td>
</tr>
<tr>
<td>TEACH CERTIFIED</td>
<td>0.218*</td>
<td>0.129</td>
</tr>
<tr>
<td>ATTEND RATE</td>
<td>0.830*</td>
<td>0.483</td>
</tr>
<tr>
<td>ENROLLMENT</td>
<td>0.002</td>
<td>0.003</td>
</tr>
<tr>
<td>FREE LUNCH</td>
<td>-0.292***</td>
<td>0.062</td>
</tr>
<tr>
<td>INSTRUCT HOURS</td>
<td>-0.015</td>
<td>0.010</td>
</tr>
<tr>
<td>SPEC EDU</td>
<td>-0.378**</td>
<td>0.170</td>
</tr>
<tr>
<td>SUBSIDY</td>
<td>23.463***</td>
<td>9.475</td>
</tr>
<tr>
<td>MINORITY</td>
<td>-0.105**</td>
<td>0.051</td>
</tr>
<tr>
<td>FEMALE</td>
<td>0.481***</td>
<td>0.168</td>
</tr>
<tr>
<td>ACADEMIC</td>
<td>2.142</td>
<td>2.177</td>
</tr>
<tr>
<td>PARENTS</td>
<td>0.816</td>
<td>3.380</td>
</tr>
<tr>
<td>TEACHERS</td>
<td>0.103</td>
<td>3.172</td>
</tr>
<tr>
<td>RESPONSIBLE</td>
<td>-7.431**</td>
<td>3.018</td>
</tr>
<tr>
<td>INDIVIDUALIZED</td>
<td>0.955</td>
<td>3.270</td>
</tr>
<tr>
<td>TECHNOLOGY</td>
<td>-3.515</td>
<td>3.562</td>
</tr>
<tr>
<td>PROJECT BASED</td>
<td>-0.367</td>
<td>3.663</td>
</tr>
<tr>
<td>DISTRICT ACHIEVE READING</td>
<td>0.487***</td>
<td>0.139</td>
</tr>
<tr>
<td>DISTRICT ACHIEVE MATH</td>
<td></td>
<td>0.460***</td>
</tr>
<tr>
<td>DISTRICT MINORITY</td>
<td>0.331***</td>
<td>0.083</td>
</tr>
<tr>
<td>YEAR DUMMY</td>
<td>-0.006</td>
<td>1.143</td>
</tr>
<tr>
<td>HIGH SCHOOL DUMMY</td>
<td>-8.595**</td>
<td>3.593</td>
</tr>
<tr>
<td>OBSERVATIONS</td>
<td>787</td>
<td>787</td>
</tr>
<tr>
<td>R-SQUARED</td>
<td>0.45</td>
<td>0.33</td>
</tr>
<tr>
<td>BREUSCH-PAGAN LAGRANGE MULTIPLIER</td>
<td>304.64</td>
<td>389.08</td>
</tr>
</tbody>
</table>

Note. For the coefficient estimates, the asterisks indicate significance as follows: *** = 1%, ** = 5%, * = 10%.