Zackary T. Nelson
James J. Jozefowicz

“Keystone Kops” and Crime Theories: A Panel Data Analysis of Pennsylvania County Crime Rates

Ralph E. Ancil

A Non-Marginalist Approach to The Cubic Value Product Function With Single Variable Input

Johnnie B. Linn III

Unleashing Leviathan: Public Goods Under Involuntary Taxation

Sabri Yilmaz

A Trade Network With a Single Trader Under Asymmetric Information

Johnnie B. Linn III

Erratum
Co-Editors

Orhan Kara & Thomas Tolin, West Chester University

Editorial Board

Michael Hannan, Edinboro University of Pennsylvania
Rocky Huang, Pennsylvania State University – Berks
James Jozefowicz, Indiana University of Pennsylvania
Tracy Miller, Grove City College
Yaya Sissoko, Indiana University of Pennsylvania
Kenneth Smith, Millersville University of Pennsylvania
Roger White, Franklin and Marshall College

The Pennsylvania Economic Review is a biannual, refereed journal sponsored by the Pennsylvania Economic Association. The journal accepts articles and book reviews on any economic topic. Article submission fee is $25 for nonmembers and members of the PEA. The subscription price for institutions and nonmembers is $25 per year. (Members of the Association receive the journal without additional cost.) Send all subscription orders to the editor. Note: All fees and subscriptions prices are subject to revision at the discretion of the Editorial Board.

Copyright ©2000 by the Pennsylvania Economic Association. Permission is granted to reproduce articles published in this journal, so long as reproduced copies are used for non-profit educational or research purposes. For other purposes, permission must be obtained from the Editor. (ISSN 1531-0949).
**Table of Contents**

*Zackary T. Nelson & James J. Jozefowicz*

“Keystone Kops’ and Crime Theories: A Panel Data Analysis of Pennsylvania County Crime Rates ................................................................. 1

*Ralph E. Ancil*

A Non-Marginalist Approach to The Cubic Value Product Function With Single Variable Input 24

*Johnny B. Linn III*

Unleashing Leviathan: Public Goods Under Involuntary Taxation ........................................... 57

*Sabri Yilmaz*

A Trade Network With a Single Trader Under Asymmetric Information ............................. 73

*Johnny B. Linn III*

Erratum ............................................................................................................. 85
PENNSYLVANIA ECONOMIC REVIEW
Style and Submission Instructions for Authors

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

Submission through the following website is encouraged:
http://www.econpea.org/pub/review.html

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.
"KEYSTONE KOPS" AND CRIME THEORIES: 
A PANEL DATA ANALYSIS OF PENNSYLVANIA COUNTY CRIME RATES

Zackary T. Nelson 
Indiana University of Pennsylvania

James J. Jozefowicz* 
Indiana University of Pennsylvania

ABSTRACT

This paper empirically analyzes various factors affecting the crime rate in Pennsylvania counties. Using a balanced panel of data from the 67 counties of Pennsylvania for the period from 1990 to 2009, regression equations are estimated using generalized least squares to control for potential heteroskedasticity and serial correlation while accounting for unobserved county-specific heterogeneity. The dependent variable is the natural logarithm of the Part I county crime rate. The independent variables are categorized according to popular theories on crime, including the economic, social disorganization, and strain theories. Law enforcement variables also are considered. The results suggest that a structural break occurred following the economic expansion of the 1990s. In addition, all three crime theories tested in this study are found to be jointly significant in explaining the Part I county crime rate in Pennsylvania.

INTRODUCTION

Crime in the United States has become a topic of great intrigue and interest, with television shows like *CSI: Crime Scene Investigation*, *NCIS*, and *Law and Order* to captivate and fuel interest among their audiences that tune in every week to watch as heinous crimes are solved neat-and-tidy in forty minutes. While these various shows are entertaining, they reveal a striking trend: crime is becoming a topic of interest to an ever-expanding audience. A wave of academic studies (e.g., DiChello (2011), Kelly (2000), Beccie (1999), and Levitt (2004)) trying to explain crime has emerged. If there is any general consensus across all of these studies, it is that there is no single clear-cut explanation of crime; it is multifaceted. As time passes, different factors are emerging that are changing crime analysis, and explanations of crime remain numerous.

As can be seen in Figure 1, although the Pennsylvania violent crime rate through the 1970s and into the 1980s was much smaller compared to the national average, it
continued to follow the national trend. Consistent with the literature, at the beginning of
the 1980s and into the 1990s, the United States experienced a steep increase in violent
crimes, while Pennsylvania had a much smaller increase in violent crime throughout the
decade. Subsequently, the nation’s violent crime rate began a dramatic decline, which
continues today. At its peak in 1991, the U.S. violent crime rate was 758.2 reported violent
crimes per 100,000 people, but it had declined by nearly half to 386.9 in 2012. In contrast,
Pennsylvania has witnessed a flatter decline in the violent crime rate from its peak of 449.9
per 100,000 people in 1991 to a more modest 348.6 in 2012. Although Pennsylvania’s
violent crime rate remains below the national average, it has been more persistent
throughout the last two decades and experienced a slower rate of decline than the U.S.
Clearly, crime in Pennsylvania remains an important phenomenon to understand.

Studies of crime in a single state are rare (e.g., Beaton (1974), Kovanadzic and
Sloan (2002), and Kovandzic and Vieraitis (2006)). Furthermore, studies focusing on
crime rates in Pennsylvania at the county level are even fewer (e.g., DiChello (2011)).
Arvanities and DeFina (2006), and Phillips and Land (2012) identify several advantages of
panel data analysis of disaggregated units of observation, such as counties. In particular,
these researchers assert that counties exhibit greater homogeneity than states, because they
are relatively small geographic areas, and have less variation in structural factors over time,
which decreases aggregation bias. Additionally, Phillips and Land (2012) observe that
counties’ boundaries do not change over time in contrast to Metropolitan Statistical Areas
(MSAs), and Andreassen (2013) points out that the larger sample sizes afforded by county-
level panel data allow for the inclusion of more control variables to reduce omitted variable
bias, and improve statistical power to yield more precise estimates. The Commonwealth of
Pennsylvania offers interesting data to analyze due to the wide socioeconomic and
demographic breadth across the state. Pennsylvania geography varies from metropolitan
hubs, like Philadelphia and Pittsburgh, to midlevel or growing cities, such as Allentown
and Easton, to rural areas, like Clearfield and Crawford Counties.

This paper extends the existing literature by increasing what Arvanities and
DeFina (2006) refer to as the small number of panel data studies on crime rates. This study
is the first to our knowledge to test the validity of different crime theories in Pennsylvania
counties by conducting a panel data analysis. Potential policy and social implications from
this study include new approaches to targeting at-risk groups to prevent outbreaks of
criminal behavior, as well as, a better understanding of Pennsylvania county-specific traits
that influence crime rates.

The order of this paper is as follows: The second section provides an overview of
existing theories on crime through which regression results will be analyzed. The third
section gives a brief review of the existing literature on crime. The fourth section discusses
the variables used in this study and their expected signs. The fifth section presents the
model employed. Following are the results of the regressions in the sixth section. Last, the
conclusions of this study and potential areas for further research on this topic are
considered.
THEORIES OF CRIME

Different theories to explain crime have been advanced over the last several decades. Chief among these crime theories are the economic theory proposed by Becker (1968), the social disorganization theory formulated by Shaw and McKay (1942), and the strain theory introduced by Merton (1938). Saridakis and Spengler (2012) also studied these theories in an analysis of crime rates in Greece.

Economic Theory of Crime

Perhaps the most well-known crime theory advanced involving economics, Becker (1968) began the trend of analyzing crime through an economics lens. Using a basic supply-and-demand model, the study discusses the optimal policies to combat crime by measuring the social loss from criminal offenses. Furthermore, Becker (1968) wanted to know exactly how many resources and how much punishment should be used to enforce different kinds of legislation. Becker (1968) concludes that criminals focus on the opportunity costs and utility of committing a crime, and whether or not the benefits of the crime outweigh the alternative legal avenues available. Finally, this study finds that there is a socially optimal (or acceptable) level of crime that could be committed and that the punishment of individuals should equal the crime they committed (i.e., an eye-for-an-eye).

Social Disorganization Theory

Focusing on sociological aspects and inequality, social disorganization theory analyzes the mechanisms of social control in a given society. Shaw and McKay (1942) identify many distinct social factors that, when enhanced, cause the various networks of social control to weaken. These factors include poverty, ethnic heterogeneity, and residential mobility as triggers that undermine a community's ability and willingness to informally control its members. Later studies (see Sampson (1987)) have also added family stability to the mechanisms of social control list.

Strain Theory

Again focusing on sociological aspects, Merton (1938) proposed the idea of strain theory. Focusing on inequality, this study argues that the higher the level of relative inequality in an area, the higher the proposed "strain" on the unsuccessful individual, which creates a sense of frustration by their failure to attain the materials associated with success. These high levels of strain, in theory, cause greater inducement for lower-status individuals to commit crime, due to the sense of alienation created. This is only enhanced by greater numbers of successful individuals in the presence of unsuccessful individuals (i.e., higher levels of different types of inequality, such as income or educational attainment).

LITERATURE REVIEW

Kelly (2000) studies the relationship between inequality and crime. This study uses data from urban counties in 1990 and separates the dependent variable into property
crimes and violent crimes to more closely examine the economic, social disorganization, and strain theories of crime, and how they relate to specific types of crime. Of this study’s independent variables, particular attention is paid to income inequality and education inequality. Other independent variables include population density, percentage of female-headed households, percentage of the population that is nonwhite, and poverty rate, among others. Kelly (2000) concludes that the behavior of property and violent crime are different. Inequality has no effect on property crime, but a strong and robust impact on violent crime. Poverty and police activity have a significant impact on property crime, but little effect on violent crime. Finally, this study concludes that property crime is well explained by the economic theory of crime, while violent crime is better explained by strain and social disorganization theories.

Beccs (1999) empirically analyzes the economic theory of crime using U.S. state-level panel data from 1971-1994. The dependent variables are various measures of crime, such as total, property, and violent crimes. Independent variables include population density, income per capita, police employment, unemployment levels, and education attainment levels. Beccs (1999) finds that “property crimes do a better job of conforming to an economic interpretation than do violent crimes and murder.” Furthermore, the study also suggests that auto theft is an urban phenomenon, but that murder is not. It also reaffirms that crime is associated with youth, that the unemployment rate is significant when used as a proxy for “legitimate work,” and that the most consistent determinant for all types of crime is per capita personal income, which was positive and significant in all regressions except crimes of murder.

Ayllon (2009) studies the effects that social and economic factors have on crimes rates, using cross-sectional data from 2000 for the 67 Pennsylvania counties. This study estimates ordinary least squares regressions with property crime and violent crime as dependent variables. Independent variables used in this study include unemployment rates, college educational attainment, poverty rate, percentage of population that is African-American, and per capita income. Ayllon (2009) concludes that criminal activity is strongly related to various socioeconomic determinants, particularly that property crimes are explained well by economic factors, while violent crimes are better explained by demographic and social factors.

DiChello (2011) studies the effects of socioeconomic, demographic, and policing factors for the 67 counties of Pennsylvania using 2000 cross-sectional data. The dependent variable for the study is the Part I crime rate. Independent variables include income, police expenditures, educational levels, population density, gender, and race. The results suggest that population density and race are positively related to crime, while income per capita is negatively related to crime.

This study closely follows Kelly (2000), where the various independent variables chosen are analyzed according to the three different theories of crime, and to a lesser extent, Beccs (1999) who analyzed how well the economic theory explains the crime rate. It also includes variables to account for the law enforcement in a county. The models analyzed are similar to those used by DiChello (2011) and Ayllon (2009).

DATA

In accordance with practices used by both Meredith and Jozefowicz (2008), and Partridge and Rickman (1996), interpolation and extrapolation were used when necessary
to create this balanced panel data set. Studying 1990 to 2009 for all 67 Pennsylvania counties provides a large sample (n=1,340) to analyze, with data gathered from various sources, such as the Bureau of Labor Statistics, the Center for Rural Pennsylvania, and the U.S. Census Bureau.

Like Arvanities and DeFina (2006), DiChello (2011), and Phillips and Land (2012), this study implements the Part I county crime rate (ONECRIME) as the dependent variable. In accordance with Lin (2009), Patalinghug (2011), Phillips and Land (2012), Saridakis and Spengler (2012), and Andrews (2013), the natural logarithm of ONECRIME is taken to adjust for skewed data. Therefore the estimated coefficients should be interpreted as the percent change in the dependent variable due to a one-unit change in the corresponding independent variable as discussed by Buonanno and Montolio (2008). In addition, Ehrlich (1996) asserts that reporting biases in crime data can be mitigated with logarithms, because the natural logarithm of crime rates probably are proportional to actual crime rates since, as Patalinghug (2011) points out, some crimes remain unreported. The Part I county crime rate is defined as the total number of reported Part I crimes per 100,000 people, which includes murder, non-negligent manslaughter, forcible rape, robbery, assault, burglary, larceny-theft, motor vehicle theft, and arson. It was collected from the Pennsylvania State Police Uniform Crime Reporting System website.

Independent Variables and Expected Signs

Economic Theory Independent Variables (ECON)

INCPERCA is the income per capita for each county in Pennsylvania. According to Becsi (1999) and Ayllon (2009), per capita income is positively correlated to crime because higher income means higher potential loot from crime. However, in DiChello (2011) and Patalinghug (2011), per capita income is negatively correlated with crime because of the higher opportunity cost associated with higher per capita incomes in a given county. Because of these conflicting views on income per capita in the literature, the expected sign of INCPERCA is ambiguous.

This study utilizes employment shares as proxies for "legitimate work" in accordance with Becsi (1999). MANUSHR is the share of the county workforce that is employed in the manufacturing industry, while SERVSHR is the share of the county workforce that is employed in the service industry. The reliance on manufacturing industry employment in the state's history, particularly in western Pennsylvania, is an important aspect of local labor market conditions within the timeframe of this study. Also, service industry employment has been steadily increasing in the last few decades nationally, and it is assumed a priori that Pennsylvania will conform to this trend. Due to the lack of empirical research to correlate these particular industry employment shares with crime, the expected signs on both MANUSHR and SERVSHR are ambiguous.

Social Disorganization Theory Independent Variables (SOCIAL)

PERNONW is the percentage of the county population that is not Caucasian. The results of Kelly (2000), Becsi (1999), DiChello (2011), and Ayllon (2009) indicate a positive correlation between the non-Caucasian population and crime. Therefore, the expected sign of PERNONW is positive.
POPDENS is the population density of each county. Becsi (1999), and Buonanno and Montolio (2008), suggest that population density is associated with crime primarily because crime is considered to be an urban phenomenon. Ayllon (2009) also suggests that an increase in population density leads to greater criminal motivation because there are more potential victims for criminals, while the chances of being caught fall. Thus, the expected sign of POPDENS is positive.

FEMWCHU7 is the female share of the county population that is over the age of fifteen and also has children under the age of seven. Due to the relatively unknown relationship with the crime rate, the expected sign of FEMWCHU7 is ambiguous.

HOME5 and BORNINST are implemented in this study to examine the residential mobility within each county, as well as, the county’s “social cohesion.” This is in accordance with social disorganization theory’s assumption that the less control a society has over its members, the higher the potential to commit crime. HOME5 is the percentage of the county population living in the same residence as five years before. BORNINST is the percentage of the county population born in their current state of residence. According to Kelly (2000), “High residential mobility reduces the cohesion of communities and results in lower social control.” Due to this finding and the assumptions of social disorganization theory, the expected signs of both HOME5 and BORNINST are negative.

Strain Theory Independent Variables (STRAIN)

PERMALE is the percentage of the county population that is male. Again, there is historical evidence within crime literature (e.g., Buonanno and Montolio (2008), Andresen (2013)) that points to a positive correlation between the percentage of the population that is male and crime. Thus, the expected sign of PERMALE is positive.

PERYOUNG is the percentage of the county population that is between the ages of 18 and 24. Consistent with the literature (e.g., Buonanno and Montolio (2008), Clement (2009), and DiChello (2011)), the expected sign of PERYOUNG is positive. On the other end of the age spectrum, PEROLD is the percentage of the county population that is 65 years or older. Due to the generally higher levels of income and personal wealth accrued over the lifetime, it is expected that a higher percentage of the senior population is associated with a lower crime rate. Therefore, the expected sign of PEROLD is negative.

UNEMP is the annual county unemployment rate. As explained by Phillips and Land (2012), higher levels of unemployment may be correlated with higher levels of crime due to enhanced levels of motivation to engage in illegal activities. However, they also point out that greater unemployment may reduce the number of potential targets and enhance guardianship of possessions as unemployed workers may remain at home more often and result in fewer criminal opportunities. Due to this debate, the expected sign of UNEMP is ambiguous. POVERTY is the percentage of the county population that falls below the poverty threshold. Higher levels of poverty are often correlated with higher rates of crime as found by Choe (2008). Thus, the expected sign for POVERTY is positive.

LNIINCOME is the natural logarithm of income for each county in Pennsylvania. It is used in this study as a proxy for income inequality in the absence of consistently accurate county-level Gini coefficients. This approach to measuring income inequality is used by Ace and Gallagher (2000). According to strain theory, higher levels of income inequality create heightened stress in lower-status individuals, increasing the chances of these individuals committing crimes. However, it may also be the case that as income
decreases, *ceteris paribus*, crime may increase as people resort to crime. As a result, the expected sign of LNIINCOME is ambiguous.

Other variables include HSDROP, which is the county high school dropout rate. As predicted by strain theory, the relatively few economic opportunities generally available to high school dropouts increase levels of stress as a result of educational inequality. Studying income inequality and educational attainment, Jenkins and Jozełowicz (2006), Chiswick and Chiswick (1987), and Jimenez (1986) propose that rising educational attainment may lead to higher levels of income inequality through different factors, such as fewer numbers of highly-skilled workers vs. uneducated workers, a declining wage rate, and skewed income based on educational levels. This increased strain may cause an increased potential to commit crime. Consequently, the expected sign of HSDROP is positive.

PERBACHE is the percentage of the county population that holds a Bachelor’s degree or higher. In the study conducted by Kelly (2000), college educational attainment reflected both increased economic opportunity and greater socialization. This notion is upheld by Feesi (1999), who found increasing educational expenditures to increase the opportunity costs of committing crime, thus reducing criminal motivation. Due to these findings, the expected sign of PERBACHE is negative.

*Law Enforcement Variables (ENFORCE)*

POLPERCA is the total number of police officers divided by the total population for each county. Patalinghug (2011) employs a similar variable. As the number of police per capita in a county rises, the chances of committing a crime and being able to get away with it fall, suggesting a diminished incentive for the individual to commit a punishable offense. According to Becker (1968), this suggests that the opportunity cost of committing a crime is very high compared to the “legitimate” alternative, where the rewards carry little-to-no chance of punishment. Thus, the expected sign of POLPERCA is negative.

ONEARRES is the number of arrests per county for Part I crimes. Typical consensus (e.g., Vanagunas (1979), Kelly (2000), Aylon (2009), and DiChello (2011)) suggests that as the number of arrests increases in an area, crime rates fall. This thinking implies a negative correlation with crime. Alternatively, arrest rates may go up because proactive law enforcement agencies are targeting the enforcement of specific crimes. This increased focus on particular crimes will inherently cause the number of crimes reported by the police to go up because previously unaccounted-for crimes are now being pursued. Therefore, the expected sign of ONEARRES is ambiguous.

ONECLEAR is the number of crimes per county that are cleared by charges being filed divided by the total number of Part I crimes recorded. Given that this variable accounts for actual charges being brought against accused individuals, it may be a more realistic indicator of “police aggressiveness” than the arrest rate, because not all people who are arrested by police are brought-up on charges. Furthermore, because this variable accounts for the actual charges filed for the crimes allegedly committed, it would stand to reason that those being charged with the crime are “being made an example of” to the rest of society, creating a “fear factor” of an increased opportunity cost and decreased rewards for committing a crime (i.e., deterring them from criminal behavior). The expected sign of ONECLEAR is negative in accordance with Buonanno and Montolfo (2008).
A complete list of variable descriptions and their expected signs can be found in Table 1.

**Descriptive Statistics**

The descriptive statistics can be found in Table 2. The Part I county crime rate (ONECRIME) has a mean of 2,232.32 crimes per 100,000 people. ONECRIME ranged from a minimum of 456.10 (Mifflin County in 2001) to a maximum of 7,377.60 (Fayette County in 1997). Looking at specific observations of the variables, Pike County, which borders both New York and New Jersey, while also being the closest Pennsylvania county to the New York City metropolitan area, has the lowest percentage of the county population that is non-Caucasian (PERNONW). Cameron County displays the lowest value with 0.18%. Alternatively, Philadelphia County has the highest percentage of the county population that is non-Caucasian with 57.5%. Finally, Montgomery County displays the highest levels of income per capita (INCPERCA), with a value of $63,002.00 per person, well above the average level of income per capita of $24,211.87 for all Pennsylvania counties.

**MODEL**

**Model Specification**

This study uses a balanced panel data set for all 67 Pennsylvania counties for the period from 1990 to 2009. To control for potential heteroskedasticity and serial correlation while accounting for unobserved county-specific heterogeneity, the models are estimated by generalized least squares (GLS) as recommended by Jirata et al. (2014). Andreassen (2013) confronted heteroskedasticity in his model. The empirical model is as follows:

\[
\ln \text{ONECRIME}_{it} = \beta_0 + \beta_1 \text{ECON}_{it} + \beta_2 \text{SOCIAL}_{it} + \beta_3 \text{STRAIN}_{it} + \beta_4 \text{ENFORCE}_{it} + \alpha_i + \epsilon_{it}
\]

(1)

**Econometric Issues**

As pointed out by Bushway and Reuter (2002), the longest post-World War II economic expansion in the United States occurred during the 1990s. Coincidentally, crime rates in the nation fell, particularly in the latter half of that decade. Furthermore, Phillips and Land (2012) identify the need to research the timeframe of the Great Recession since their data sample omitted 2008-2009. Since the period under consideration spans 1990 until 2009, it is reasonably conceivable that a structural break exists in the data. Thus, a Chow test was performed, and the results confirm the presence of a structural break between the 1990-2000 and the 2001-2009 periods.

**RESULTS**

The results of the GLS regressions can be found in Table 3. In Model 1, period weights were employed for the GLS analysis, and F-tests were run to test for the joint
significance of the economic theory, social disorganization theory, and strain theory
groups, as well as, the law enforcement group. Each of these groups was found to be
jointly significant at the 1% level. The $R^2$ of Model 1 is 0.598.

The Model 1 results indicate that INCPERCA is statistically significant at the 1%
level and displays a negative sign. SERVSHR is statistically significant at the 1% level
and also has a negative sign. MANUSHR is not statistically significant. The regression
results from Model 1 for the ECON variables suggest that as income per capita in a county
increases, the crime rate declines perhaps due to the higher opportunity cost of committing
a crime, which reduces the motivation to engage in criminal activity. Cantor and Land
(1985), and Phillips and Land (2012) suggest this explanation. This finding is consistent
with DiChello (2011) and Patalinghug (2011) who also obtained negative signs on this
covariate. The share of employment in the service industry, which has a significant
negative relationship with crime, may speak to the sectoral shift of the U.S. labor market to
the service sector over the traditional “blue-collar” manufacturing sector, seen in the “Rust
Belt” of western Pennsylvania. As jobs shift increasingly to the service sector, those
individuals employed in the manufacturing sector may experience declining “legitimate”
labor market opportunities.

PERNONW is significant at the 5% level, FEMWCHU7 and HOME5 are both
statistically significant at the 1% level, and POPDENS and BORNINST are not statistically
significant. The regression results from Model 1 for the SOCIAL group reveal an expected
positive sign on the PERNONW variable, which could be due to Pennsylvania’s racial
homeness outside of urban areas, which account for less than one-third of
Pennsylvania’s counties according to the Center for Rural Pennsylvania. This positive
correlation between PERNONW and the dependent variable is consistent with the findings
of Becsi (1999), Kelly (2000), Choe (2008), Aylton (2009), and DiChello (2011). The
relative absence of urban counties as opposed to rural counties in Pennsylvania may also
explain the insignificant negative sign on population density. The significance and
negative signs on FEMWCHU7 and HOME5 are upheld by social disorganization theory,
and suggest that the presence of children and lower levels of resident turnover cause people
to “plant their roots” in a community and become more invested in its welfare. Such social
cohesion deters crime as noted by Andresen (2013).

UNEMP is significant at the 5% level, while PERYOUNG, PEROLD,
PERBACHE, POVERTY, and LINCOME are statistically significant at the 1% level.
The lack of significance of PERMALE could be due to the fact that the specific age
brackets of males in the county were not taken into account. The positive sign on UNEMP
suggests that greater joblessness may increase the motivation to commit crime as discussed
by Cantor and Land (1985), and Phillips and Land (2012). The negative sign of
PERYOUNG may reflect a lack of motivation on the part of younger individuals to engage
in criminal activities as found by Andresen (2013). PEROLD displays a positive sign,
which indicates that senior citizens may represent “easy” low-risk, high-reward targets and,
therefore, greater opportunities for criminals as explained by Cantor and Land (1985), and

The unexpected positive sign on PERBACHE could be confirmation of
educational inequality, because as stated by the Brookings Institution (2003), Pennsylvania
as a whole has suffered from a serious “brain drain” of human capital in its counties, due to
a “...aging workforce and a net loss of young, educated workers.” Given the rising costs of
attending a university, and the increased earnings gap between those with a high school
diploma and college graduates, again supported by the Brookings Institution (2003), this may suggest increased strain on individuals who are not willing or able to go to college. This sign also is consistent with studies done on the relationship between educational attainment and income inequality, such as Jimenez (1986), and Chiswick and Chiswick (1987). In a similar vein, the significance of LNINCOME suggests that income inequality increases the presence of motivated offenders as found by Andresen (2013). The positive sign on POVERTY, and the corresponding relative economic hardships associated with it, also support the validity of the strain theory of crime. Choe (2008) obtained similar results.

All three ENFORCE group variables are statistically significant at the 1% level. Andresen (2013) argues that the positive sign on POLPERCA is indicative of persistent crime problems over time, which is apparent in Figure 1. The significance of ONEARRES is consistent with this study’s expectations and is upheld by DiChello (2011), and Jacob and Rich (1981), who found a positive correlation between robbery and arrest rates, and suggest that, “...as the police devote more attention to robbery arrests, they also record more robbery offenses.” As noted in Buonanno and Montolio (2008), DiChello (2011), and Saridakis and Spengler (2012), ONECLEAR, as a measure of police aggressiveness, reduces the rewards from committing crime and increases the opportunity costs of crime resulting in fewer motivated offenders, which concurs with its negative sign.

Structural Break

Since the results of a Chow test indicate a structural break in the data, the regression findings for the 1990-2000 and 2001-2009 time periods are separately presented to highlight any differences. Model 2 in Table 3 presents the regression results for 1990-2000, which encompasses the longest post-World War II economic expansion. The $R^2$ of Model 2 is 0.578. The signs for MANUSHR, SERVSRH, PERNOW, POPDENS, FEMWCHUT, HOME5, BORNINST, UNEMP, PERYOUNG, PEROLD, HSDROP, POVERTY, LNINCOME, POLPERCA, ONEARRES, and ONECLEAR are robust from Model 1 to Model 2. However, INCPERCA, PERMALE, and PERBACHE experience sign changes. Specifically, INCPERCA and PERMALE carry positive signs, but only INCPERCA is significant. PERMALE remains insignificant, and PERBACHE has a negative sign, but loses its statistical significance. The positive correlation between INCPERCA and the dependent variable indicates enhanced criminal opportunity as higher levels of income and wealth during the expansion translated into heightened consumption of expensive items, which criminals find attractive to steal. Becsi (1999), Aylion (2009), Patalinghug (2011), and Phillips and Land (2012) support this notion.

Model 3 in Table 3 includes the regression results for 2001-2009, which encompasses the 2001 recession and the entirety of the Great Recession. The $R^2$ of Model 3 is 0.483. The signs for MANUSHR, SERVSRH, PERNOW, POPDENS, FEMWCHUT, BORNINST, UNEMP, PEROLD, HSDROP, POVERTY, LNINCOME, POLPERCA, ONEARRES, and ONECLEAR are robust across Models 1-3. Notably, INCPERCA remains positive as in Model 2, but loses its statistical significance. The sign on HOME5 turns to positive, but loses its significance. The sign on PERMALE becomes negative, which is consistent with Model 1, but now it is significant at the 1% level. This may reflect the disproportionately negative impact of economic downturns on men, which reduces criminal opportunities and increases guardianship as they spend more time at home while out of work. Anecdotally, the Great Recession especially hurt employment in
traditionally male-dominated industries, such as manufacturing and construction. PERYOUNG is positive and significant at the 1% level in Model 3 in contrast to Models 1-2 where it was negative and significant at the same level. This finding is consistent with the notion that a period of economic downturn may produce motivated offenders as mentioned by Andresen (2013). PERBACHE resumes its positive sign from Model 1, but it is not statistically significant.

Robustness Checks

To test the robustness of the results from Model 1, Model 4 in Table 3 presents the results of a random effects (REM) specification. The findings are largely robust in signs and significance with those of Model 1. One exception is a change in sign from positive to negative on HSDROP, but the estimated coefficient is not statistically different from zero. Additional variables, which become insignificant in the REM, include INCPERCA, SERVSHR, PERBACHE, and POVERTY. However, MANUSHR achieves significance at the 5% level in the REM while remaining positive in sign.

To further assess whether or not the findings are sensitive to the specification of the model, cross-section weights are utilized instead of time period weights in the GLS analysis. The results of this alternative approach appear in Model 5 of Table 3. Once again, the findings are largely robust in signs and significance with those of Model 1. However, PERNONW and UNEMP become insignificant while retaining their positive signs, and HSDROP remains positive, but achieves significance at the 1% level.

CONCLUSION

This study implements panel data analysis for the 67 counties of Pennsylvania between the years 1990 and 2009 to study how the major crime theories in the literature explain the Part I county crime rate. The theories include Becker’s (1968) economic theory of crime, Shaw and McKay’s (1942) social disorganization theory, and Merton’s (1938) strain theory. The findings indicate that all three crime theories play a role in explaining crime rates in Pennsylvania counties. For example, a one-unit increase in the share of employment in the service sector will reduce the crime rate by 1.15% on average, ceteris paribus. Furthermore, a one-unit increase in the share of the county female population with young children leads to an average 5.56% reduction in the crime rate, ceteris paribus, and a similar increase in the county population living in the same residence as 5 years earlier causes the crime rate to fall by 2.89% on average, all else equal. More of the strain theory variables are highly significant relative to the variables in the other two theory groups. A one-percentage-point increase in the county unemployment rate will raise crime by 1.16% on average, ceteris paribus. Levitt (2001) obtained a similar magnitude using state-level panel data. Similarly, a one-unit increase in the poverty rate results in a 1.09% increase in the crime rate on average, all else equal. Given the proportion of Pennsylvania’s rural counties, and the typical emphasis on urban centers in the existing crime literature, any robust unexpected signs obtained in this analysis may actually be an accurate depiction of factors influencing county crime rates in Pennsylvania.

The analysis also reveals noteworthy differences between periods of economic growth and periods of economic contraction. Specifically, greater per capita income encourages crime during an economic expansion as criminal opportunity increases, and a larger male population in a county raises crime rates during a period of prosperity perhaps
due to enhanced motivation. However, the male presence reduces criminal activity afterward perhaps because of less opportunity to commit crimes as more men may remain out of work. There also appear to be more motivated youth offenders during times of slower economic growth.

Policy Implications

With respect to policymaking in Pennsylvania, the results of this study contribute to the knowledge base for formulating policies and legislation focused on the problem of crime. The significant negative relationship between the clearance rate (i.e., cases that actually lead to charges being filed) and the crime rate suggests that a policy of “setting an example” may be a valid deterrent to crime.

Additionally, this research offers policy alternatives consistent with the theories of crime, which may be effective for reducing occurrences of criminal activity, and extend beyond the typical emphasis on law enforcement remedies. The findings of this study indicate that lower population density is associated with higher crime rates. Given the extensive rural nature of Pennsylvania relative to many other states, crime-fighting policies should focus on urban counties and rural counties separately, as the same tactics in urban areas may not be as effective in rural areas. Efforts to build strong community ties and to provide support for single mothers with young children should be considered to strengthen social cohesion and monitor citizens informally.

The results of this analysis also offer insights into the behavior of crime rates during periods of economic upturn, as well as, times of recession. As discussed by Saridakis and Spengler (2012), a downturn in economic activity results in weaker aggregate demand with associated increases in unemployment rates, and a higher likelihood of budget reductions, which have the potential to decrease clearance rates. Both of these phenomena have a tendency to exacerbate crime rates, and Saridakis and Spengler (2012) argue that good economic policy is simultaneously good crime-fighting policy, because crime acts as an automatic destabilizer of the economy. Pennsylvania policymakers thus should consider employment growth policies to promote service industry employment, as generally discussed by Shields and Novak (2003), and provide job training to reduce crime by offering legitimate opportunities in the labor market to would-be offenders.

Extensions of the Research

Due to the broad categories of crimes that the Part I crime rate encompasses, there may be some omitted factors that were not included in this study. Further research on this topic should investigate specific crime categories from the Part I rate, such as violent or property crimes, because some literature has shown that these two categories behave differently. Another interesting point for future research would be accounting for Part II crime categories, which include drug offenses, prostitution, and other lower-level crimes.

Future research on this subject should also try to include more independent variables. Using police data, such as police expenditures, conviction rates, and measures of the effectiveness of policing policies, may be important additions. As mentioned earlier, the creation of a Gini coefficient for each county would help to better account for income
inequality. It would also behoove future researchers to carefully look at the crime theories mentioned in this study, and the variables that were chosen to represent them.

ENDNOTES

*Corresponding author: Department of Economics, 213 McElhaney Hall, Indiana University of Pennsylvania, Indiana, PA 15705. Email: James.Jozefowicz@iup.edu.

1A variable to account for the number of police in a given county each year (POLPERCO) was also included, but was ultimately discarded as it did not control for population size and was highly correlated with ONEARRES.

2This idea conforms to Kessler and Levitt (1999), who found sentencing enhancements to promote deterrence rather than immediately confirming incapacitation.

3Of the 67 counties that comprise the Commonwealth of Pennsylvania, 48 counties are designated as rural counties, according to the Center for Rural Pennsylvania. This means that only about 28% of Pennsylvania’s counties are urban counties.

(We appreciate helpful suggestions and assistance from Brittany A. DiChello, Christopher R. Jeffords, Jingze Jiang, Alida V. Merlo, Neil R. Meredith, and participants at the 2014 Pennsylvania Economic Association Conference. We especially are indebted to Yaya Sisoko whose recommendations significantly improved this paper. Any and all mistakes are our own. We gratefully acknowledge Debbie Bacco for her superb editorial assistance.)

REFERENCES


Figure 1. U.S. and Pennsylvania Violent Crime Rates
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Description</th>
<th>Expected Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variable</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ONECRIME</td>
<td>The Part 1 county crime rate per 100,000 people for all 67 counties in the state of Pennsylvania for the years 1990 to 2009. <em>Pennsylvania State Police Uniform Crime Report</em>.</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Independent Variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Economic Theory of Crime Group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INCPERCA</td>
<td>The income per capita for each of the 67 counties in the state of Pennsylvania for the years 1990 to 2009. <em>Bureau of Economic Analysis and U.S. Department of Agriculture's Economic Research Service</em>.</td>
<td>(?)</td>
</tr>
<tr>
<td>MANUSHR</td>
<td>The share of the county workforce that is employed in the manufacturing industry, linearly interpolated and extrapolated based on 1990, 2000, and 2010 data points. <em>U.S. Census Bureau</em>.</td>
<td>(?)</td>
</tr>
<tr>
<td>SERVSHR</td>
<td>The share of the county workforce that is employed in the service industry, linearly interpolated and extrapolated based on 1990, 2000, and 2010 data points. <em>U.S. Census Bureau</em>.</td>
<td>(?)</td>
</tr>
<tr>
<td><strong>Social Disorganization Theory Group</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERNONW</td>
<td>The percentage of the county population that is not Caucasian for all 67 counties in the state of Pennsylvania for the years 1990 to 2009. <em>U.S. Census Bureau</em>.</td>
<td>(+)</td>
</tr>
<tr>
<td>POPDENS</td>
<td>The number of people within a square mile for each of the 67 counties in the state of Pennsylvania for the years 1990 to 2009. <em>U.S. Census Bureau</em>.</td>
<td>(+)</td>
</tr>
<tr>
<td>FEMWCHU7</td>
<td>The female share of the county population that is over the age of 15 and has children under the age of 7, linearly interpolated and extrapolated using 1990, 2000, and 2010 data points as needed. <em>U.S. Census Bureau</em>.</td>
<td>(?)</td>
</tr>
<tr>
<td>HOME5</td>
<td>The percent of the population living in the same residence as five years before, linearly interpolated and extrapolated using 1990, 2000, and 2010 data points as needed for each of the 67 counties in the state of Pennsylvania. <em>U.S. Census Bureau</em>.</td>
<td>(-)</td>
</tr>
<tr>
<td>Variable</td>
<td>Description</td>
<td>Location</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>BORNINST</td>
<td>The percent of the population born in their state of residence, linearly interpolated and extrapolated using 1990, 2000, and 2010 data points as necessary for each of the 67 counties in the state of Pennsylvania. U.S. Census Bureau.</td>
<td>(+)</td>
</tr>
<tr>
<td>PERMALE</td>
<td>The percentage of the county population that is male for each of the 67 counties in the state of Pennsylvania for the years 1990 to 2009. U.S. Census Bureau.</td>
<td>(+)</td>
</tr>
<tr>
<td>UNEMF</td>
<td>The county unemployment rate for each of the 67 counties in the state of Pennsylvania for the years 1990 to 2009. U.S. Census Bureau.</td>
<td>(?)</td>
</tr>
<tr>
<td>PERYOUNG</td>
<td>The percentage of the county population that is between the ages of 18 and 24 for each of the 67 counties in the state of Pennsylvania for the years 1990 to 2009. U.S. Census Bureau.</td>
<td>(+)</td>
</tr>
<tr>
<td>PEROLD</td>
<td>The percentage of the county population that is 65 years or older for each of the 67 counties in the state of Pennsylvania for the years 1990 to 2009. U.S. Census Bureau.</td>
<td>(-)</td>
</tr>
<tr>
<td>HSDROP</td>
<td>The county high school dropout rate for each of the 67 counties in the state of Pennsylvania for the years 1990 to 2009. U.S. Census Bureau.</td>
<td>(+)</td>
</tr>
<tr>
<td>PERBACHE</td>
<td>The percentage of the county population that holds a Bachelor’s degree or higher for each of the 67 counties in the state of Pennsylvania for the years 1990 to 2009. U.S. Department of Agriculture’s Economic Research Service.</td>
<td>(-)</td>
</tr>
<tr>
<td>POVERTY</td>
<td>The percent of the total population for each of the 67 counties in the state of Pennsylvania that falls below the poverty threshold for the years 1990 to 2009. Small Area Income Poverty Estimate.</td>
<td>(+)</td>
</tr>
<tr>
<td>LNINCOME</td>
<td>The natural log of income for all 67 counties in the state of Pennsylvania for the years 1990 to 2009. U.S. Census Bureau</td>
<td>(?)</td>
</tr>
<tr>
<td>POLPERCA</td>
<td>The total number of police officers divided by the total population for each of the 67 counties in the state of Pennsylvania for the years 1990 to 2009. Pennsylvania Uniform Crime Reports.</td>
<td>(-)</td>
</tr>
<tr>
<td>Variable Name</td>
<td>Description</td>
<td>Expected Sign</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------</td>
<td>---------------</td>
</tr>
<tr>
<td>ONEARRES</td>
<td>The total number of arrests per county for Part I crimes for all 67 counties in the state of Pennsylvania for the years 1990 to 2009. <em>Pennsylvania Uniform Crime Reports</em></td>
<td>(?)</td>
</tr>
<tr>
<td>ONECLEAR</td>
<td>The number of crimes per county that are cleared by charges being filed divided by the total number of crimes recorded for Part I crimes for all 67 counties in the state of Pennsylvania for the years 1990 to 2009. <em>Pennsylvania Uniform Crime Reports</em></td>
<td>(-)</td>
</tr>
</tbody>
</table>
Table 2. Descriptive Statistics

<table>
<thead>
<tr>
<th>Variables</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variable</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ONECRIME</td>
<td>7377.60</td>
<td>456.10</td>
<td>2252.52</td>
<td>1021.19</td>
</tr>
<tr>
<td><strong>Independent Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INCPerCA</td>
<td>63002.00</td>
<td>12158.00</td>
<td>24211.87</td>
<td>7546.11</td>
</tr>
<tr>
<td>MANUSHR</td>
<td>47.54</td>
<td>2.52</td>
<td>19.91</td>
<td>7.19</td>
</tr>
<tr>
<td>SERVSHR</td>
<td>67.10</td>
<td>21.66</td>
<td>38.10</td>
<td>7.67</td>
</tr>
<tr>
<td>PERNOW</td>
<td>57.50</td>
<td>0.18</td>
<td>4.99</td>
<td>6.94</td>
</tr>
<tr>
<td>POPDENS</td>
<td>11745.00</td>
<td>11.20</td>
<td>454.44</td>
<td>1423.19</td>
</tr>
<tr>
<td>FEMWDCHU7</td>
<td>9.52</td>
<td>0.42</td>
<td>6.19</td>
<td>1.17</td>
</tr>
<tr>
<td>HOMES</td>
<td>77.34</td>
<td>41.57</td>
<td>65.20</td>
<td>5.22</td>
</tr>
<tr>
<td>BORNINST</td>
<td>94.30</td>
<td>16.66</td>
<td>79.87</td>
<td>11.81</td>
</tr>
<tr>
<td>PERMALE</td>
<td>262.43</td>
<td>30.96</td>
<td>49.20</td>
<td>6.11</td>
</tr>
<tr>
<td>UNEMP</td>
<td>17.30</td>
<td>2.60</td>
<td>6.23</td>
<td>1.99</td>
</tr>
<tr>
<td>PERYOUNG</td>
<td>34.17</td>
<td>6.50</td>
<td>13.36</td>
<td>3.16</td>
</tr>
<tr>
<td>PEROLD</td>
<td>24.90</td>
<td>8.97</td>
<td>16.08</td>
<td>2.27</td>
</tr>
<tr>
<td>HS dropout</td>
<td>40.84</td>
<td>4.89</td>
<td>13.46</td>
<td>3.79</td>
</tr>
<tr>
<td>PERBACHE</td>
<td>43.26</td>
<td>7.30</td>
<td>17.07</td>
<td>6.75</td>
</tr>
<tr>
<td>POVERTY</td>
<td>30.69</td>
<td>1.50</td>
<td>11.53</td>
<td>3.54</td>
</tr>
<tr>
<td>LNINCOME</td>
<td>24.76</td>
<td>17.96</td>
<td>21.49</td>
<td>1.32</td>
</tr>
<tr>
<td>POLPerCA</td>
<td>0.0019557</td>
<td>0.001806</td>
<td>0.0011474</td>
<td></td>
</tr>
<tr>
<td>ONEARRES</td>
<td>29904</td>
<td>13</td>
<td>1294.883</td>
<td>3072.50043</td>
</tr>
<tr>
<td>ONECLEAR</td>
<td>315.4</td>
<td>9.7</td>
<td>30.99552</td>
<td>11.1080508</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------</td>
<td>--------------------------</td>
<td>--------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>INCPECA</td>
<td>-5.52E-06** ((-2.795199))</td>
<td>1.23E-05** ((2.373108))</td>
<td>4.92E-06 ((1.556232))</td>
<td>-1.71E-06 ((-0.482465))</td>
</tr>
<tr>
<td>MANUSHR</td>
<td>0.0001862 ((1.168634))</td>
<td>0.002739 ((1.384688))</td>
<td>0.003794 ((1.714511))</td>
<td>0.003572** ((2.226566))</td>
</tr>
<tr>
<td>SERVSHR</td>
<td>-0.011505*** ((-4.484303))</td>
<td>-0.004405 ((-1.213247))</td>
<td>-0.016910*** ((-4.457151))</td>
<td>-0.004715 ((-1.665783))</td>
</tr>
<tr>
<td>PERNONW</td>
<td>0.007147** ((2.008316))</td>
<td>0.012110** ((2.641466))</td>
<td>0.027595*** ((4.702604))</td>
<td>0.007176** ((2.004744))</td>
</tr>
<tr>
<td>POPDENS</td>
<td>-2.56E-05 ((-1.414961))</td>
<td>-8.17E-05** ((-2.342814))</td>
<td>-7.99E-05*** ((-3.295461))</td>
<td>-2.98E-05 ((-1.674813))</td>
</tr>
<tr>
<td>FEMWCHU7</td>
<td>-0.055580*** ((-4.053603))</td>
<td>-0.005821 ((-0.246385))</td>
<td>-0.092222*** ((-5.919743))</td>
<td>-0.073429*** ((-5.051524))</td>
</tr>
<tr>
<td>HOME5</td>
<td>-0.028922*** ((-10.88703))</td>
<td>-0.041975*** ((-11.16597))</td>
<td>0.006428* ((1.675396))</td>
<td>-0.026388*** ((-9.807195))</td>
</tr>
<tr>
<td>BORNINST</td>
<td>0.000910 ((0.999846))</td>
<td>0.004244*** ((3.024081))</td>
<td>0.001503 ((1.346844))</td>
<td>0.000591 ((0.648989))</td>
</tr>
<tr>
<td>PERMALE</td>
<td>-0.000667 ((-0.561409))</td>
<td>8.72E-05 ((0.088076))</td>
<td>-0.048350*** ((-6.377584))</td>
<td>-0.001036 ((-0.876833))</td>
</tr>
<tr>
<td>UNEMP</td>
<td>0.011572** ((2.250347))</td>
<td>0.026746*** ((3.816986))</td>
<td>0.002652 ((0.402280))</td>
<td>0.013929** ((2.042959))</td>
</tr>
<tr>
<td>PERYOUNG</td>
<td>-0.015268*** ((-3.294263))</td>
<td>-0.024581*** ((-3.933746))</td>
<td>0.019441*** ((2.974518))</td>
<td>-0.015920*** ((-3.264684))</td>
</tr>
<tr>
<td>PEROLD</td>
<td>0.021180*** ((3.318557))</td>
<td>0.016823 ((1.131702))</td>
<td>0.006055 ((0.697669))</td>
<td>0.012970** ((1.971308))</td>
</tr>
<tr>
<td></td>
<td>HSDROP</td>
<td>PERBACHE</td>
<td>POVERTY</td>
<td>LNINCOME</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>----------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>0.000245 (0.083645)</td>
<td>0.005628 (0.952171)</td>
<td>0.006219** (2.085628)</td>
<td>-0.000733 (-0.254848)</td>
</tr>
<tr>
<td>PERBACHE</td>
<td>0.009916*** (4.659124)</td>
<td>-0.001455 (-0.420818)</td>
<td>0.004011 (1.015344)</td>
<td>0.003679 (1.257035)</td>
</tr>
<tr>
<td>POVERTY</td>
<td>0.010864*** (3.399256)</td>
<td>0.012393*** (3.209353)</td>
<td>0.017139*** (3.444669)</td>
<td>0.005703 (1.610157)</td>
</tr>
<tr>
<td>LNINCOME</td>
<td>0.082654*** (6.966608)</td>
<td>0.079054*** (5.299837)</td>
<td>0.125097*** (6.432432)</td>
<td>0.087307*** (7.439754)</td>
</tr>
<tr>
<td>POLPERCA</td>
<td>79.54095*** (7.809184)</td>
<td>214.1930*** (11.16547)</td>
<td>31.38247*** (2.581372)</td>
<td>86.47513*** (8.480420)</td>
</tr>
<tr>
<td>ONEARRES</td>
<td>2.75E-05*** (3.882084)</td>
<td>1.92E-05 (1.210410)</td>
<td>8.55E-06 (1.179457)</td>
<td>2.58E-05*** (3.742579)</td>
</tr>
<tr>
<td>ONECLEAR</td>
<td>-0.002092*** (-3.042353)</td>
<td>-0.001403* (-2.002078)</td>
<td>-0.003945*** (-3.122134)</td>
<td>-0.001897*** (-2.739261)</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>7.913602*** (20.98030)</td>
<td>7.569751*** (14.86428)</td>
<td>6.901993*** (9.050032)</td>
<td>7.732693*** (20.17693)</td>
</tr>
<tr>
<td>ADJUSTED R²</td>
<td>0.453196</td>
<td>0.577670</td>
<td>0.482594</td>
<td>0.437604</td>
</tr>
</tbody>
</table>

Parentheses contain t-statistics; Significance Levels: *** at 1%, ** at 5%, * at 10%.

* F-tests were run on the three crime theory groups and the law enforcement group to determine their joint significance.
The results indicate that each group is jointly significant at the 1% level.
A NON-MARGINALIST APPROACH TO THE CUBIC VALUE PRODUCT FUNCTION WITH SINGLE VARIABLE INPUT

Ralph E. Ancil
Geneva College

ABSTRACT

In this paper the cubic total value product function is examined with a view to providing an alternative to the method of marginal analysis. Instead of finding the maximum net return with a single variable input using the differential calculus, an algebraic method is used to derive a model where the production process is conceived as a departure from the case of uniform coefficients in the average value product function and a linear average value product function is used instead. This function is paired with the linear portion of the (quadratic) average value product function to derive an equation whose solution yields the point of maximum net returns. Furthermore, the equations of marginal analysis are themselves derived from the algebra and an adjusted linear average value product function is used to replace the marginal value product function. This algebraic method is further applied to derive additional results including the use of break-even points to arrive at the solution. Diminishing returns are left embedded in the analysis and average and total output are expressed as gross returns. The analysis is followed with a brief discussion of the implications this approach has for firm behavior and orthodox theory with emphasis on the importance of a constant average value in understanding production decisions.

INTRODUCTION

This paper is a continuation of work showing that for a subset of total revenue and cost equation pairs simple algebra and break-even analysis are sufficient to find the optimal level of output (or input). Past work (Ancil, 2010; Ancil, 2011) examined a number of common cases, and models, used to find the profit maximizing level of output using an extended version of break-even analysis, averages, and algebra. In the present paper,
another model is explored, the cubic production function, not primarily for its economic properties, but for its mathematical ones, though it is treated as a total value product function and paired with a total factor cost equation. An alternative method is introduced whose key feature is the linear average value product (LAVP) function based on the special case of uniform coefficients. This special case is then used as a standard to adjust to changes in the coefficients and to derive other important results.

Criticism of the marginal approach to economic analysis has a long history. Institutionalist economists, among others, have often expressed their dissatisfaction with it. As an historical example we can refer to Michigan State University's Robert Solo (1968) who writes of marginal analysis: "Emphasis is perpetually on the incremental rather than on the average or the total effects of action." (pp. 48-49, original emphasis). But this emphasis, he argues, is unrealistic; it does not reflect how businesses actually make their decisions and refers to the well-known surveys of businesses beginning with the classic work of Charles Hitch of Oxford University which conclude that businesses do not base their costing and pricing decisions on marginal calculations. (p. 49)

When it comes to producer's choice of resource inputs, Solo points out that these are sometimes fixed in regard to a given product output. When this is the case "...the marginalist conception of business as hinging upon the calculations of the cost effects of incremental or decremental change in resource combinations is irrelevant as an explanation of producer's choice." (p. 51) For Solo this is significant because Leontief's well-known and most important input-output analysis assumes "...such fixity to be the general and prevailing condition of the economy." (p. 51) He concludes: "It is...very doubtful that any business entity in fact possesses or attempts to gather the information that would be required in order to determine the range of cost effects of marginal variations in resource inputs." (p. 52)

Solo dismisses the counter-argument that, although the models of marginalism cannot be taken to describe literally economic reality, businesses should be thought of as if they behaved this way. In his words:

It could be argued that even if marginal choice isn't practicably feasible, it is the way that businesses would choose if they could, and since it is really what they want, that therefore what they do must be an approximation to marginalism. This would be like saying that since the quickest most direct route between two points is by flying, that anyone wanting to go from A to B must, by inference, want to fly. Since they cannot fly but instead are obliged to walk, therefore walking must be understood as an approximation of flight. Obviously nonsense. (pp. 52-53)

An extreme statement to be sure, but not surprising from someone in the institutionalist school of thought.

However, one finds similar, though more selected and qualified criticism, from the recent writings of orthodox (marginalist) economists. One example will suffice. After reviewing a number of cost studies, mainstream economists Keat and Young (2009) summarize: "A large majority of these studies have concluded that marginal cost in the short run is relatively constant...The upward-sloping - U-shaped - average and marginal cost curves postulated by economic theory tend to be the exception in empirical findings." (p. 308) They also say that "such results should make economists pause and reexamine some of their theoretical conclusions..." Though they themselves believe there is enough
ambiguity in the way the research was conducted that the possibility remains that
marginalist results, i.e., eventually rising marginal costs and diseconomies of scale, will
still be validated. (p. 308)

But if resource inputs are generally administered in fixed relations and if marginal
cost is often constant, and thus identical to average cost over the relevant range, is it not
reasonable to "pause and reexamine" our model and consider what an alternative one of,
say, a production function would be like? If we took averages, totals, and linear functions
and their slopes as the central concepts in the model and left marginal methods as, well, as
marginal, would it yield the same results as the present mainstream model? What would be
the implications for the firm? That is precisely what this paper does.

In the following analysis the cubic production function is selected because of its
central place in the development of the basic microeconomic model as evidenced in its
continued presence in economic textbooks even though it is seldom used if at all by
researchers. (See below.) The problem is to find an alternative, non-marginal, method to
determine the same result as the orthodox method achieves, that is, the quantity at which
production is maximized or, in the case of the value product, the quantity of variable
input at which net returns are maximized. At the same time, this search will involve the
use of algebra rather than differential calculus.

All this means that this paper is predominantly a presentation of mathematical
results. Since this research challenges mainstream thinking, it is important to display the
results clearly. Once this is done, some discussion about the implications for the firm can
be entertained. In what follows it will be shown that:

a. a linear function with algebraic treatment gives the same solution as marginal
   analysis and is called the linear average value product (LAVP);

b. the marginal product can be derived from the average product and is called
   the average value product adjusted (AVPA);

c. a variation of the method can be derived which emphasizes the shifting of the
   function to obtain the solution and is called the average value product
   inverted (AVPI);

d. the functions leading to the solution can be expressed in gross returns rather
   than net returns;

e. the firm can be understood to make decisions about the hiring of variable
   input amounts based on the goal of achieving or maintaining a constant
   average value for the fixed input which coincides with the maximization of
   net returns.

Extended treatments of some detail are given in two appendices.

COMMENT ON THE LITERATURE

The cubic production function is well known in the economic literature as an
example of an economic model relating a variable input to the quantity of output. The
short run cubic production function emphasizes the role of a limiting factor which causes
diminishing returns. The object in production theory is to determine the amount of the
variable input, either labor or capital in a one variable case, which will produce the
maximum output.
The model does, however, present difficulties which limit its usefulness in actual economic research. Yunker (2008) writes: "Although the cubic production function is ubiquitous in principles and intermediate textbooks, it finds little or no application in professional economic research, possibly because polynomials are notoriously difficult to work with in mathematical manipulations." To substantiate this he cites the number of "hits" on article abstracts in the EconLit database showing that while other functions have relative frequent hits, led by Cobb-Douglas with 141, the cubic production function had zero. He concludes that because of its pedagogical usefulness, though, it will likely continue to be retained in the textbooks.

A sampling of textbooks from managerial economics to intermediate micro theory largely supports Yunker. To name only a few one could cite Thomas and Maurice in their book Managerial Economics Foundations of Business Analysis and Strategy (2011) who give one of the more detailed examinations of the cubic production function. Keat and Young (2009) and Pindyck and Rubinfeld (2009) likewise examine the function. The same is true for Allen, Weigelt, Doherty and Mansfield (2009), and also Perloff (2012) and Salvatore (2012). On the other hand, Varian (2014) does not use this function in the treatment of production theory.

In the typical exposition of this function, the model is introduced and explained in terms of its total output, marginal and average product curves. The discussion proceeds to the law of diminishing returns and the stages of production. Stage I is defined to lie from the origin to the maximum point on the average product curve. Stage II from the maximum point on AP to the point where MP is zero. Stage III is the range where MP is negative.

The emphasis throughout is on the marginal changes and the achievement of maximum output which is identified when the slope of a line tangent to the total product curve is zero. The technique for this is the use of differential calculus although some texts use a non-calculus approach and reserve the use of calculus for footnotes or appendices. Even here the method remains marginal analysis.

In the following discussion the analysis proceeds algebraically and without the use of marginal concepts. The result will still be useful for pedagogical purposes and the equations just as simple to solve as the final ones in orthodox treatments, though there will no need to use differential calculus.

**THE VALUE PRODUCT FUNCTION WITH UNIFORM COEFFICIENTS**

We begin with the familiar short-run cubic production function with uniform coefficients but converted to a total value product equation (TVP) by multiplying the former by dollars and also using the corresponding average value product function (AVP):

\[
\text{TVP} = ax + bx^2 - cx^3 \\
\text{AVP} = a + bx - cx^2,
\]

and where \( a = b = c \).

We note that \( \text{AVP} = \text{TVP} \) at the net revenue maximizing level of input, NRMI, which is one:

\[
ax + bx^2 - cx^3 = a + bx - cx^2.
\]
Since the coefficients are equal, division by one of them gives,

\[ x + x^2 - x^3 = 1 + x - x^2. \]  \hspace{1cm} (4)

This can be re-written as,

\[ 2x^2 - x^3 - 1. \]  \hspace{1cm} (5)

The relevant solution clearly is at \( x = 1. \)

Thus, for all cases where the coefficients are equal, the NRMI is one, and AVP = TVP at that point. (See Figure 1.)

To verify the results using marginal analysis, we set marginal value product (MVP) equal to marginal factor cost (MFC) in the usual manner, and we have from equation (1),

\[ MVP = 1 + 2x - 3x^2. \]  \hspace{1cm} (6)

Solving algebraically yields,

\[ x = (\pm \sqrt{4 + 1})/3 \]
\[ = 1 \text{ (or } -1/3). \]  \hspace{1cm} (7)

It is also true that the AVP (equation 2) and TVP (equation 1) must equal each another at the break-even point on the x-axis when \( y = 0, \) we have:

\[ 1 + x - x^2 = 0. \]  \hspace{1cm} (9)

Solving in the usual manner gives,

\[ x = (\pm \sqrt{5 + 1})/2 \]
\[ = 1.6180, \text{ or } -0.6180. \]  \hspace{1cm} (11)

Using these results for TVP verifies the equality with AVP for these x-values. This is also true for any TVP/AVP equations when the coefficients are equal. See Figure 1.

Putting this hypothetical function in the context of the firm’s decision-making processes, the standard narrative would describe the firm as hiring additional amounts of the input variable until \( x = 1. \)

**AN ALTERNATIVE METHOD**

The usual interpretation of the production function or, as in this case, the total value product function, is that firms compare small changes in marginal value product and marginal factor costs until the two are equal (MVP=MFC) and net return is maximized, or at least, they act as if they proceeded in this manner. The critical concept here is that the third term \( (x^2 \text{ in TVP}) \) is the term expressing diminishing returns which constrains the production/value product function, making the determination of a maximum point necessary.
However, this same result can be obtained in a different way. Assume two simple linear functions that coincide with each other and are given by,

\[ y_1 = x, \text{ and } y_2 = \sqrt{1} \, x, \text{ and } \]
\[ x = \sqrt{1} \, x. \]  
(12)
(13)

Add 1 to the left side and k to the right side under the square root sign,

\[ 1 + x = \sqrt{(1 + k) \, x}. \]  
(14)

Squaring both sides and recalling that for this special case \( x=1 \), we have,

\[ 1 + 1 = \sqrt{(1 + k)^2} \]
\[ (2)^2 = (\sqrt{(1 + k)})^2 \]
\[ 4 = 1 + k, \text{ and} \]
\[ k = 3. \]  
(15)
(16)
(17)
(18)

(It should be clear that this applies for the set of real numbers and that the focus is on the first quadrant.)

Rewriting equation (14), we have,

\[ 1 + x = \sqrt{4} \, x. \]  
(19)

This result is the same as with marginal analysis where the left side equation intersects the right side equation at the NRMI. Adding "k" to the right side changes the slope in such a way as to account for the effect of the diminishing returns term (DR) but it spreads these (gross) returns, or their effect, evenly over the entire function. It is, in other words, equivalent to solving the AVP function (equation 9) and replacing the term \( "x^3" \) with \( \sqrt{4} \, x \) to arrive at equation (19).

Continuing in this way we can multiply equation (14) by 2, recalling that \( k = 3 \), to get:

\[ 2 + 2x = 2 \sqrt{(1 + 3) \, x} \]
\[ = \sqrt{4(1 + 3)} \, x \]
\[ = \sqrt{4(4 + 12)} \, x \]
\[ = \sqrt{16} \, x. \]  
(20)
(21)
(22)
(23)

We note that at this point, the expression under the square root sign can be re-written as,

\[ 2 + 2x = \sqrt{(2^2 + (3)(2)(2)) \, x}. \]  
(24)

One may hypothesize that the new 2s on the right reflect the other coefficients on the original AVP function \( (2 + 2x - 2x^2) \), so that in general we have,

\[ a + bx = \sqrt{(b^2 + 3ac) \, x}. \]  
(25)

This is easily tested. When the coefficients are all 3s \( (AVP = 3 + 3x - 3x^2) \), we have,
\[ 3 + 3x = \sqrt{(9 + 3(3))} \times \] 
\[ = \sqrt{36} \times x. \] 
\[ (26) \] 
\[ (27) \]

When they are all 4s (AVP = 4 + 4x - 4x^2), we have,

\[ 4 + 4x - \sqrt{(16 + 3(4)(4))} \]
\[ = \sqrt{64} \times x. \] 
\[ (28) \] 
\[ (29) \]

This is for the uniform case where all the coefficients are equal and x = 1.

**Figure 2** shows what these equations (19 and 23) look like graphically. The left side is clearly the linear portion of the original AVP function while the right side equation is the linear average value product function (LAVP) which replaces the term cx^2.

**CHANGING THE COEFFICIENTS**

So far we have assumed the coefficients are the same. What happens if they are not? The DR term or its effect is reflected in the LAVP and may be called an “internal cost” because it is internal to the TVP (or production) function. External costs can be included which change the coefficients in the following manner, where a positive total factor cost (TFC) and average factor cost (AFC) are added:

\[ \text{TFC} = x \]
\[ \text{AFC} = 1, \]
\[ (30) \]
\[ (31) \]

and recalling the original example where all the coefficients are 2's, we have,

\[ \text{TVP} = 2x - 2x^2 - 2x^3, \]
\[ \text{AVP} = 2 + 2x - 2x^2. \]
\[ (32) \]
\[ (33) \]

Subtracting AFC = 1 from the left side of AVP gives the net AVP function (NAV):\n
\[ \text{NAV} = 1 + 2x - 2x^2. \]
\[ (34) \]

The coefficients are no longer all the same. Using the right hand part of equation (25) with the first two terms of equation (34), we now have:

\[ 1 + 2x = \sqrt{(4 + 3(2)(1))} \times x \]
\[ = 0.8604 = x. \]
\[ (35) \]
\[ (36) \]

From marginal analysis, we have,

\[ \text{MVP} = 1 + 4x - 6x^2 \]
\[ (37) \]

and

\[ x = (\sqrt{5} + \sqrt{2})/3\sqrt{2} \]
\[ = 0.8604. \]
\[ (38) \]
\[ (39) \]
As shown in Figure 3, the original AVP line is shifted down by 1 while the slope is held constant but in the other function, the LAVP, the slope is lessened from $\sqrt{16}$ to $\sqrt{10}$. The two functions intersect at the NRMI. In both approaches the firm hires the variable input until $x = .8604$ though in the method advocated here the firm thinks of this amount as a linear average calculated from the productive process understood as a whole and implemented from the beginning, not as in marginal analysis as a series of very small comparisons between input and output.

THE LONG FORMULA

We can envision an AVP function with uniform coefficients ($a = b = c$) which serves as a standard to evaluate changes in the coefficients. As in equations (34) and (35), the coefficients were all 2 until a cost was introduced altering $a = 2$ to $a = 1$. As in equations (20-29), continued experimentation shows a relationship between the coefficients which allows us to calculate the changes in the slope of the right hand expression and thus correctly solve for the NRMI. That formula is:

$$\text{Slope} = \sqrt{4b^2 + (a - b)3c + (c - b)3b}.$$  \hfill (40)

This can be understood as reflecting the differences from the slope, $b$, on the expression “$bx$.” The “$b$” can be understood as the coefficient that is common to both the original uniform set of coefficients and the new set which departs from the uniform case. The formula measures the difference between these two cases weighted by the terms “$3c$” and “$3b$.” Or, this can be thought of as the differences of coefficients in the final equation from the slope, $b$, on the term, “$bx$.”

For example, consider,

$$\text{AVP} = 1 + 2x - x^2.$$  \hfill (41)

The coefficients $a = 1$ and $c = 1$ are taken as departures from the uniform case where $a = b = c = 2$. The new slope is:

$$\text{Slope} = \sqrt{16 + (1 - 2)3 + (1 - 2)6}$$
$$= \sqrt{7}.$$  \hfill (42)

Thus, the whole equation becomes,

$$1 + 2x - \sqrt{7} x$$
$$1/(\sqrt{7} - 2) - x$$
$$1.5486 = x.$$  \hfill (44)

The long formula (equation 40) can be reduced to the short form (equation 23):

$$\text{Slope} = \sqrt{4b^2 + 3ac - 3bc + 3bc - 3b^2}$$
$$= \sqrt{4b^2 - 3b^2 + 3ac}.$$  \hfill (47)
\[ = \sqrt{b^2 + 3ac}. \] \hspace{1cm} (49)

Using this result in the full equation, we have the general form,
\[ a + bx = \sqrt{3ac + b^2} \cdot x, \] \hspace{1cm} (50)
and finally,
\[ x = \frac{a}{\sqrt{3ac + b^2} - b}. \] \hspace{1cm} (51)

We can compare this result with the general expression for marginal analysis:
\[ TVP = ax + bx^2 - cx^3, \] \hspace{1cm} and \hspace{1cm} (52)
\[ MVP = a + 2bx - 3cx^2. \] \hspace{1cm} (53)

After working through the algebra for MVP, we have,
\[ x = (\sqrt{3ac + b^2} + b)/3c. \] \hspace{1cm} (54)

Equating equation 54 with equation 51, we have,
\[ a/(\sqrt{3ac + b^2} - b) = (\sqrt{3ac + b^2} + b)/3c. \] \hspace{1cm} (55)

We observe the sign on “b” is reversed and the two equations are essentially reciprocals of one another, though the numerator of the one and the denominator of the other differ. The calculus approach uses a larger numerator \((\sqrt{3ac + b^2} + b)\) divided by \(3c\), while the algebraic method divides “a” by a smaller denominator \((\sqrt{3ac + b^2} - b)\). In other words, the right-hand slope from equation 50 \((\sqrt{3ac + b^2})\) is adjusted by adding or subtracting the left-hand slope \((b)\). The differences in the denominators is counter-balanced by the differences in the numerators in such a way that the solution is the same.\(^1\)

There is an optimum only when the LAVP slope is larger than “b.” In the extreme case of an infinite slope the optimal input is zero. This occurs when “c” becomes infinitely large, the overall fraction (equation 51) approaches zero. In the other extreme where “c” becomes infinitesimally small (approaches zero), the “optimum” becomes infinitely large with constant (average) returns to scale. In other words, we are left effectively with the left side of the original equation (50), i.e., the linear portion of the original AVP function. An alternative way of putting it is that for the special case where the slopes are equal (and “c” literally is zero) the lines are parallel and there is no optimum.

What assurance do we have that this equation (51) yields not merely an optimal solution but the optimal solution or that it is not merely a special case? Since we obtain a maximum (or minimum) by taking the derivative of the main function, setting the resulting equation equal to zero, and solving for “x”, it is sufficient to show that from equation (50, or 51) one can derive this first derivative, the same as that from the calculus. This is easily done in three steps by (1) squaring both sides of equation (50), (2) subtracting out the term \((bx)^2\) common to both sides of the equation, and (3) dividing through by “a”, a factor common to all terms. The result is the first derivative. We have,
\[ (a + bx)^2 = (\sqrt{b^2 + 3ac} \cdot x)^2 \] \hspace{1cm} (56)
\[ a^2 + 2abx = 3acx^2 \] \hspace{1cm} (57)
\[ a + 2bx - 3cx^2 = 0. \] \hspace{1cm} (58)

32
This outcome is the same as equation (53) from marginal analysis, and solving for “x” results in equation (54). The algebraic and calculus-based equations are thus fully equivalent. (The only exception is the special case where a = 0 which is treated below in note 2 of Appendix A.)

It should be clear that from here (equation 58) it is a simple matter of determining algebraically the point at which diminishing returns set in, that is, the familiar relation, x = b/3c which is the axis of symmetry. This is the same as the second order condition derived from marginal analysis. (The axis of symmetry on AVP, x = b/2c, is, of course, well-known.)

THE AVERAGE VALUE PRODUCT, ADJUSTED

Given the importance of the marginal concept in production theory, it is instructive to see that we can derive an adjusted average value product curve (AVPA) to replace the marginal value product curve (MVP). In budget-line fashion, we can use the y-intercept on the AVP function (the “a” value) and subtract the relevant function – the right hand side of equation (50) – from it to intersect the average factor cost curve (AFC).

Using the coefficients from equation (37), where the average factor cost is AFC = 1, we have,

\[ \text{AVPA} = 2 - \left( \sqrt{(4 + 6) - 2} \right)x \]
\[ = 2 - \left( \sqrt{10} - 2 \right)x \]
\[ = 2 - 1.1623x. \]

(59) \hspace{1cm} (60) \hspace{1cm} (61)

This intersects AFC at y = 1 with a NRMI of .8604.

For AFC = 1.5, a = .5, and x = .7743,

\[ \text{AVPA} = 2 - \left( \sqrt{7} - 2 \right)x \]
\[ = 2 - .6458x \]
\[ = 1.5. \]

(62) \hspace{1cm} (63) \hspace{1cm} (64)

The function intersects AFC = 1.5. See Figure 4.

So more broadly, we have,

\[ \text{AVPA} = 2 - \left( \sqrt{b^2 + 3ac} - b \right)x. \]

(65)

See Figure 5.

This function y (= AFC), identifies all the changes in average (linear) cost from the point where diminishing returns begins (x = 1/3) to the NRMI at zero costs (x = 1) and between y = 2.67 and y = 0. It identifies the demand curve for input x (e.g., labor) given different wage rates.

There is a limit to the economically meaningful range of linear cost functions. The x-value on the vertex of the AVP function, b/2c, is coordinate with the point on the TVP function which coincides with the function y = 2.5x. Any slope greater than 2.5 lies outside every point on the TVP function in quadrant I and so is uneconomical. In other words the vertex on the MVP function, which is 2.67, the point at which diminishing
returns begin, is outside the scope of normal economic considerations. The relevant x-values, then, often called stage II, lie between 1 and 1/2 (not 1/3) in the above example. (More generally, the line tangent to TVP which identifies this maximum is \( y = [a + b^2/4c]x \). Of course, an even narrower economic limitation is defined by the slope of a line extending from the origin to the peak of the TVP function which identifies the maximum return when cost is zero.

The obvious next step is to replace "2" with "a". By varying the value of "a" as a function of AFC, i.e., what amounts in the above procedure to subtracting AFC from the original "a" value, a yet more general equation can be derived of the form:

\[
\text{AVPA} = a_0 - (\sqrt{b^2 + 3ac} - b)x. \tag{66}
\]

This is an important application because it allows students to see the usual stages of the production function and the demand curve for factors but from a linear average value production equation, not from the marginal value product curve. In terms of theory, since the value of "a" is the constant portion of residual value remaining after subtracting cost (AFC), it can be understood as the constant average rent accruing to the fixed factor as a result of hiring more of the variable factor. As such, the y-intercept in this view takes on a heightened significance it does not have in mainstream theory.

This function achieves the same results as the MVP curve even though it has no explicit term reflecting diminishing returns. This is accomplished by the effect the average factor cost (AFC) has on the difference between the two slopes involved. As AFC increases, the residual amount, or rent, "a", decreases as does the slope, \( \sqrt{3ac + b^2} \), which contains "a." Likewise, the difference with the other slope, b, also decreases and so does the amount subtracted from \( a_0 \). In other words, the changes of external cost, AFC, not inherent limits of the production function (cx^2 and 3c), drive the changes which in this function give the same pattern of results as the MVP curve. The firm can be seen as striving to maximize this rent by hiring the right amount of the variable factor. Maximization of rent coincides with the maximization of net returns. Though perhaps not the final goal, it is the firm's operational target. All this constitutes a substantially different perspective from that of mainstream theory.

**THE AVERAGE VALUE PRODUCT, INVERTED**

Because the foregoing approach departs from mainstream theory, it may be helpful to get a broader perspective and context on it by examining a variation of the method and some of its formal properties. This is intended for insight rather than actual use in solving an equation, although that can be done.

The model developed so far is based originally on the assumption of uniform coefficients and focused mainly on the constant "a" value on the left side of equation (25) as well as the slope of the LAVP on the right side to find the optimal solution. If we turn our attention to that slope and adjust the expression under the square root sign, it will allow us to find points which break-even with the x-axis and are also identical with the optimal solution which has not been the case before. This form of the AVP function is inverted and rewritten as,

\[
\text{AVPI} = (3ac + b^2) - (x + b)^2. \tag{67}
\]
We note the equation is devoid of both the "a" (y-intercept) in the numerator of the algebraic solution and the "3c" in the denominator of the marginal calculus. But because its solution values can be transformed into either one, it can be understood as the generic form which underlies both. In the algebraic case the transformation is given by the inverted solution, say, break-even point one (B1), and multiplying by "a":

\[ \text{NRMI} = \frac{a}{B1}. \]  

(68)

For the case of equation (33), the inverted function is,

\[ \text{AVPI} = 16 - (x + 2)^2. \]  

(69)

Figure 6 gives the graph of this function. Setting \( y = 0 \) and solving in the usual manner, we get \( x = 2 \) and -6. Inverting this answer \( x = 2 \) which makes it \( \frac{1}{2} \), and then multiplying it by the coefficient "a", which is 2, we get the transformed value of 1 for the original function.

We can also see that the first parenthetical expression of equation (67) defines the vertex as the sum of the y-intercept, 3ac, and the slope of the linear part of the AVP equation, b, squared. (As before, the square root of the vertex is the slope of the right-hand part of equation (50)).

Changes in external costs are easily accounted for in this form. To calculate the effect of a linear cost change we must subtract it from the original amount to get the difference on the vertical axis of symmetry, \( D_y \):

\[ D_y = (3ac + b^2) - (3ac + b^2) \]
\[ = 3ac - 3ac \]
\[ = 3c(a_0 - a). \]

(70)  

(71)  

(72)

Since \( a = a_0 - \text{AFC} \), this becomes,

\[ D_y = 3c (a_0 - (a_0 - \text{AFC}) \]
\[ = 3c (\text{AFC}). \]

(73)  

(74)

Subtracting equation (74) from equation (67) gives\(^3\),

\[ \text{AVPI} = 3c(a_0 - \text{AFC}) + b^2 - (x + b)^2. \]

(75)

In the case of equation (34), the external cost is \( \text{AFC} = 1 \), the vertical difference is therefore 6, and the new vertex of the inverted equation (AVPI) changes from 16 to 10.

\[ \text{AVPI} = 10 - (x + 2)^2. \]

(76)

The right-hand break-even point is 1.1623, and, inverting and multiplying by "a", which is 1, the NRMI is 1/1.1623, or .8604. Likewise, the y-intercept changes from 12 to 6. See Figure 6.
We can also interpret the results, not as before as a difference between slopes, but as a simple form of averaging where the axis of symmetry \(A_x = -b\) is central with variations based on the slope, the square root of the vertex. The break-even values are given by:

\[-b + (-\sqrt{3ac + b^2}) = A_x + (-\sqrt{V}).\]  

(77)

A few other formal relationships deserve mention. For instance, we should note the horizontal distance \(D_x\) between the absolute values of the break-even points on the x-axis is given by,

\[D_x = (\sqrt{V} + A_x) + (\sqrt{V} - A_x)\]
\[= 2\sqrt{V}.\]  

(78)  

(79)

For the example above (with coefficients of 1,2,2), we have,

\[= 2\sqrt{3(1)2 + 4}\]
\[= 6.3246.\]  

(80)  

(81)

This distance between the break-even points can also be defined as the original slope of the LAVP \(S_L\) plus twice the value of the right-hand break-even point:

\[S_L + 2B1 = 2\sqrt{V}\]
\[B1 = \sqrt{V} - S_L/2.\]  

(82)  

(83)

Continuing with the above result, we have,

\[S_L + 2B1 = 4 + 2(1.1623)\]
\[= 6.3246.\]  

(84)  

(85)

In general “\(b\)” can be replaced with “\((1/2)S_L\)”. We also observe that the break-even points sum to \(S_L\):

\[B1 + B2 = S_L\]
\[1.1623 - 5.1623 = -4\]  

(86)  

(87)

The absolute value of their product yields the y-intercept:

\[(B1)(B2) = 3ac\]
\[(1.1623)(5.1623) - 3(1)2\]
\[= 6.\]  

(88)  

(89)  

(90)

This brief and incomplete exercise with the generic form, AVPI, stresses the perennial presence of the slope, \(S_L\), derived from the case of uniform coefficients but remaining even when changes are later introduced. Equally significant is the presentation in terms of an average with variations. This is a different concept of both the solution and business decision-making processes. It claims the optimal value to the production problem
is discovered by considering that which is central, not marginal, and departures from that centrality.

**A NOTE ON THE LONG RUN**

In considering the long run where the fixed factor can be changed, a firm may need to know what change in the internal cost, from the original value of \( c_m \), would compensate for an increase in external cost in such a manner as to retain the same NRMI as in the original case which is 1. This means that equation (51) can be rewritten as,

\[
\sqrt{(3ac + b^2)} - b = a. \tag{91}
\]

Solving for \( c \) gives,

\[
c = \frac{a + 2b}{3}. \tag{92}
\]

Since \( 2b \) equals \( S_m \), this can also be written as,

\[
c = \frac{a + S_m}{3}. \tag{93}
\]

Using equation (34), with coefficients of 1,2,2, means \( c \) must decrease to 1.67 to compensate for the external cost and,

\[
AVP = 1 + 2x - 1.67x^2. \tag{94}
\]

At this level, \( TVP = AVP \) at \( x = 1 \), the NRMI. In other words a \( 1/6^\text{th} \) reduction on the internal cost compensates for a 50% reduction in the \( a \) value, from 2 to 1. (This does not restore the original level of net return which was 2 but it does increase it from 1.07 to 1.33.)

**COMMENT ON THE ANALYSIS**

The foregoing treatment illustrates the broader possibility of using algebra for traditional economic topics. The solution was arrived at without the use of marginal concepts, such as marginal value product and marginal factor cost, but with simple algebra. This makes understanding basic economic principles more accessible for students and business decision makers unless, of course, marginal analysis itself is the basic principle in question. It also directs attention to the possibility that businesses behave in non-marginalist ways suggesting instead that they distribute the effect of diminishing returns over the course of production so as to produce a stable linear average value product while still maximizing net return. In other words, the internal cost imposed by the fixed factor is evenly allocated like overhead in the production process. However, some may object that since an adjustment to the slope is made to accommodate changes in cost, that this is essentially marginalist thinking anyway and so the analysis is not really different. In other words, is any recognition of change necessarily marginalist?

The issue is similar to the textbook treatment of full cost-plus pricing. After considering the advantages and disadvantages of this method most textbooks conclude that
in a world of imprecise and costly information, under certain conditions this method gives a good approximation of profit maximizing prices. Salvatore, for example, writes, "...to the extent that marginal cost is constant or nearly constant over the normal or standard level of output of the firm, marginal cost is approximately equal to (the fully allocated) average cost. Therefore cost-plus pricing would not lead to product prices that are much different from prices based on the MR = MC rule" (2012, p. 510). Indeed, in the analysis of markups based on elasticity of demand, he replaces the term MC with one for average cost. Keat and Young likewise hold that where the marginal cost curve is constant or horizontal it is identical or nearly identical to average cost and in their arithmetic treatment they likewise use average cost to replace marginal cost (2009, pp. 406-407). It is interesting to observe that they say when a firm adjusts its markups to meet new circumstances it is acting "as if it has knowledge of its demand and cost curves; that is, it is acting consistently with marginal pricing" (2009, p. 406).

However, this congenial attitude, though quite common, is not held by all authors. Thomas and Maurice argue vigorously that this method does not work (2011, pp. 610-614). After voicing their distaste for taking up valuable textbook space to deal with the topic, its continued use requires some attention even while they assert sophisticated firms no longer use this technique. In their truncated arithmetic treatment they do not, as in so many other cases, replace MC with AC and relate the method to the elasticity of demand to show that the markup can be consistent with marginal pricing. Instead, they spend time emphasizing what they believe are both the theoretical and the practical flaws of this technique. "Any method of pricing, whether it is cost-plus pricing or some other rule for setting price, will be generally unreliable for finding the profit-maximizing price and output if the technique is not mathematically equivalent to setting marginal revenue equal to marginal cost" (2011, p. 613). Apparently, they do not believe there are any circumstances in which such an equivalency holds. Instead, they conclude it is "a poor choice for making decisions" and they assert flatly that "[p]rofit-maximizing firms do not employ cost-plus pricing methodology" (2011, p. 613).

So, when is a technique "mathematically equivalent" to the MC = MR rule, or as in the case above, when MFC = MVP? Does this refer to the result only or to the pathway to that result? In the foregoing analysis the result is mathematically the same as in the standard textbook treatment, but the method, a mere technique, and the model, the assumptions upon which the method is founded, are not equivalent. Would Thomas and Maurice be content to use this method? One suspects not, for even here the different model implies a different view of firm behavior which is inconsistent with mainstream economic thinking. Instead of seeing firms making their production and pricing decisions by comparing small marginal changes in revenues and costs (or, value product and factor costs), they are seen here as making those decisions based on the average performance where the internal and external costs are distributed evenly throughout the production process. The method implies that firms are just as interested in stability and continuity (of production and pricing) as they are in adjusting to cost changes. The method assumes that many small changes are mutually cancelling and are not worth tracking. When changes are necessary, the method can handle them. Nor is there any requirement for change to be marginal, that is, per one unit, rooted in the calculus of infinitesimals.
IMPLICATIONS FOR THE THEORY OF THE FIRM

This approach also has implications for the behavior of the firm's decision-makers. To understand this aspect we need to look more closely at how the model works, emphasizing the difference between the algebraic and calculus-based methods. This can best be done by again comparing the two solution equations but this time with view to the nature of the firm's decision-making process:

\[
\begin{align*}
\text{Algebra} & \quad x = \frac{a}{\sqrt{3ac + b^2}} - b \\
\text{Calculus} & \quad x = \frac{1}{3}
\end{align*}
\]

(95) (96)

The first equation discounts the constant average (gross) returns, "a", of the fixed factor, by the difference in the slopes of the two linear functions. The optimal variable factor amount is that quantity which complements this difference and is identified by the intersection of these two functions, i.e., the amount of x hired that leads to the same amount of constant average value product. Another way of putting it is to say that the product of the optimal solution and this difference in hiring rates must equal the constant value of “a” (after deducting AFC). When this occurs, net returns are, coincidently, maximized.

The second equation discounts the numerator, which is derived from the MVP and shows the rate of change in the TVP, by a term, 3c, which reflects diminishing returns. (Such a term is conspicuously absent from the algebraic approach.) The optimal solution is the amount of the variable factor needed to make the tangent lines on the respective TVP and TFC functions parallel, or, in equally familiar terms, where the marginal value product of the last unit hired just earns its marginal factor cost.

An example will make this clear. Consider the following AVP function,

\[4 + 3x - \frac{1}{3}x^2 = 0.\]

(97)

Solving for x, we have,

\[x = 10.17891 \text{ or } -1.17891.\]

(98)

See Figures 7 and 8.

Using the algebraic method, we have,

\[4 + 3x = \sqrt{((4)(3)(1/3)) + 9}x\]

\[x = 6.60555128.\]

(99) (100)

In Figure 9 both the left and right sides of the equation (99) now yield the same gross average returns. With this value of x, the y value is 23.816665 which yields gross total returns of 157.32213. If we want to calculate the net returns, we subtract the third factor from the gross returns:

\[
\begin{align*}
\text{TVP} & = 157.32213 - \frac{1}{3}(6.60555128)^3 \\
& = 61.24817.
\end{align*}
\]

(101) (102)
Also in Figure 9, the gross average returns of area 2 are given as the product of “a” times “x” i.e., \(4(6.60555128)\) which yields 26.42221.

The important aspects of this economic model, then, are that the solution which maximizes net returns can be found in terms of gross, rather than net, returns. This arises because the area under the MVP can take other shapes, such as squares and triangles, which lend themselves to linear treatment. (See Appendix B for more on this method of calculation.) It also means that the downward-sloping portion of the MVP curve reflecting diminishing returns is not necessary to find the solution.

Figure 10 illustrates these points further. The upper functions, expressed as average gross returns, converge at the x value that maximizes net returns. Beyond this point net returns are reduced. To be more explicit, it is the difference in y values of these functions that decreases until, at the intersection, it is zero, thus indicating that net returns have been maximized.\(^4\)

The lower function, whose slope is itself the difference between the slopes expressed in the denominators of equations (51) and (95), is a line proceeding from the origin to the line \(y = ax\) given by the equation,

\[
a = 0.60555128x. \tag{103}
\]

Solving for \(x\) again gives 6.60555128 as the optimal amount (after which net returns decrease) and corresponds exactly with the intersecting point on the upper part of the graph. The firm’s decision to hire a certain amount of the variable input factor can be understood as being motivated by a desire to maintain the constant average rent for the fixed factor which increases with increases in the quantity of variable input until the that average (“a”) is reached. (Cf. equation 66 above.)

**SUMMARY**

The foregoing is offered as an alternative method and model to the traditional cubic value product function of marginal analysis. The approach results in the same net return-maximizing input answer and still identifies the stages of production but does so by using simple algebra instead of marginal concepts based on the use of differential calculus. However, it achieves this result without relying on the concept of diminishing returns and expresses the results merely as gross returns. The derivation of the linear, adjusted, and inverted forms of the model show some measure of versatility in application which should be useful in teaching and research, though clearly more needs to be done in this regard. As a model of firm behavior it implies, among other things, that businesses select the amount of the variable factor so as to complement the fixed factor by focusing on the constant “a” rather than on comparisons of small changes of input and output.
ENDNOTES

1 This equation (51) can also be used to solve for the profit maximization of the textbook case where total revenue is a quadratic function of the form TR = ax - bx^2 coupled with a cubic total cost curve, TC = cx^3. The formula differs from equation 51 only in that in the denominator the sign is “+” on the “b” term: x = a/[(3ac + b^2) + b].

By setting “c” to zero, the solution is “a/2b,” which is the same result derived from marginal analysis. This is a more general method than the one proposed in an earlier paper. (See Ancil, 2011.)

2 Under the square root sign, the “a” used here is the result of subtracting the value of AFC from a_0, the original y-intercept on AVP, and is not to be confused with the y-intercept used as the constant on the adjusted average value product function (AVPA).

3 Care must be taken in any interpretation of the coefficients (a, b, c) of the inverted equation (AVPI) and the use of the given formulae. These coefficients are different from those of the initial average value product (AVP) which are a_0, b_0, c_0. Thus,

\[ AVPI = 3c_0(a_0 - AFC) + b_0^2 - (x + b_0)^2 = 3c_0(a_0 - AFC) + b_0^2 - [x^2 + 2b_0x + b_0^2] .\]

The coefficients in brackets are now a_i = b_0, b_i = 2b_0, and c_i = 1. As a result we can write the vertex of AVPI is,

\[ V_i = a_i + b_i^2 \text{ and } a_i + b_i^2 = 3ac_0 + b_0^2, \quad \text{where } a = a_0 - AFC.\]

4 It may be helpful to draw an analogy from elementary calculus. A tangent line to a curved function is defined in terms of the converging points of secant lines. As the points come closer together, the differences of their slopes diminish, until the secant line coincides with the tangent line and is tangent at one point of the curve.

REFERENCES


APPENDIX A  CRITICAL CONDITIONS FOR AVPA

The function in the main text has three critical attributes that deserve comment:

1. When \( \sqrt{b^2 + 3ac} = 0 \), then the AVPA is at its maximum at \( y = a_0 + b^2/3c \). This corresponds to the peak of the MVP curve. Under the square root sign, the minimum value it can take is when \( b^2 = -3ac \). At that point the value under the square root sign is zero and the function is at its peak. For equation (33), that means when \( a = -2/3 \), it implies that the most AFC can be is 2.67 which is, like the MVP, also the vertex of the AVPA function, the point at which diminishing returns begins. (Any value less than this (< -2/3) will yield a negative result.)

One can proceed along this fashion where values of \( x \) are to the left of \( x = 1/3 \), but the “a” values correspondingly become greater than -2/3 replicating the y-values, lying to the right of \( x = 1/3 \). This works so long as one uses the negative roots of the function for \( x \) values to the left of \( 1/3 \). The point is that the “a” value can’t be below this limit, or any combined values of “a” and “c” such that \( 3ac > b^2 \).

2. From equation (66) it is clear that when \( \sqrt{b^2 + 3ac} - b = 0 \), AVPA intersects \( y = a_0 \). What is not immediately clear is the \( x \) value required since the term becomes zero when either “a” or “c” is zero. That value is \( x = 2b/3c \). While there is more than one way to solve for this, it is useful to set the AVPA equal to the (algebraically derived) MVP and thereby emphasize the importance of the “2/3” constant with changeable coefficients:

\[
\begin{align*}
\text{a}_0 - (\sqrt{b^2 + 3ac} - b)x &= a_0 + 2bx - 3cx^2 \\
- (\sqrt{b^2 + 3ac} - b) &= 2b - 3cx \\
x &= (2b + \sqrt{b^2 + 3ac} - b)/3c \\
-2b/3c + [\sqrt{b^2 + 3ac} - b]/3c.
\end{align*}
\]

Replacing \( x \) above with the right-hand expression of equation (51), this result can be solved again and rewritten as,

\[
x = 2b/3c + a/\sqrt{b^2 + 3ac} + b.
\]

Where \( b = c \), as in this case above, and \( a = 0 \), the constant is simply 2/3 and the mathematical range is from \( x = 1/3 \) to \( x = 3/3 \), or “mean” \( x \) value of \( 2/3 + 1/3 \). If the focus is just on the value of “a” at zero, the same result is attained from solving the right side of equation (A-1).

The first term of equation (A-5) can be replaced with \( (4/3)A_a \), the axis of symmetry for the AVP, and the entire equation understood as centering around \( A_a \) with variations leading from the peak of AVP to the x-axis by way of the second term, the difference in slopes scaled by 3c. This can also be seen as a measure of the difference between NRMI when \( a = 0 \) and when \( a > 0 \). Adding the difference in slopes (again, scaled by 3c) to \( (4/3)A_a \) gives the NRMI.

The value “2b/3c” can also be seen as the difference between the AVP and AVPA axes of symmetry added to the AVP axis of symmetry, which becomes,

\[
b/c - b/3c - 2b/3c.
\]
There are other observations along this same line. For example, the solution can also be found when \( a = 0 \) by taking the ratio of the axes of symmetry for the AVPA (MVP), \( A^M \), and AVP, \( A^\Lambda \), curves:

\[
A^M/A^\Lambda = (1/3c)1/2b = 2b/3c. \tag{A-7/8}
\]

Though when "a" is zero the solution of the remaining equation is straightforward (i.e., \( x = b/c \)), it is also merely the right-hand break-even point, B1, not the NRMI. That is found by multiplying B1 by 2/3,

\[
\text{NRMI} = (2/3)b/c = 2b/3c. \tag{A-9/10}
\]

3. When \( x = b/2c \), the AVPA value corresponds to the peak of the AVP function. This occurs when the value of "a" is

\[
\begin{align*}
a &= a_o - AFC \\
&= A_s^2c \\
&= -(b/2c)^2c \\
&= b^2/4c. \tag{A-11/14}
\end{align*}
\]

where \( A_s \) is the axis of symmetry for the AVP function and \( a \), \( b \), and \( c \) are the coefficients on the AVP as before. The constant "a_o" is the y-intercept and AFC is the average factor cost.

More generally this can be written as,

\[
\text{AVPA} = a_o + (1/2)bx. \tag{B-15}
\]

And using the axis of symmetry (\( x = b/2c \)) we have the value of \( x \) at the (AVP) vertex giving us the final result,

\[
\text{AVPA} = a_o + b^2/4c. \tag{B-16}
\]
APPENDIX B  CALCULATION OF AREAS

Another method to calculate the results of the main text based on the simple areas of Figure 9 rather than on the curved area under the MVP function can also be derived. We see that,

(area(1) and area(2))2/3 = area of MVP.  

And

\[ (65.4500 + 26.4222)2/3 = 91.8722(2/3) \]
\[ = 61.2481, \]

so that

\[ \text{gross returns} - (x^*)(y^*)/3 - 2/3 (y^* - a)x^*/2 + (2/3)a \]

Thus, both rectangular and triangular areas are accounted for and the equation simplifies to an expression of net returns (NR),

\[ NR = (1/3)x^*(y^* + a). \]

The diminishing returns effect (DR) is expressed by the third term, \( cx^3 \), and is the difference between total returns (TA) and net returns (NR),

\[ DR = TA - NR \]
\[ = x^*y^* - (1/3)x^*(y^* + a) \]
\[ = 2/3(x^*)(y^* - a/2). \]

The TVP can also be calculated in this way:

\[ TVP = (1/3)x^*(y^* + a) \]

\[ = (1/3)(6.60555128)(23.81665384 + 4) \]
\[ = 61.24817. \]

We can also calculate the TVP if we let the slope,

\[ m = \sqrt{3ac + b^2} \]

and let,

\[ y^* = x^*m. \]

Then,

\[ TVP = (1/3) x^*(x^*m + a). \]
Substituting equation (51) for \( x^* \) gives,

\[
\text{TVP} = \frac{1}{3}[a/(m - b)][am/(m - b) + a] \quad (B-15)
\]

\[
= \left( \frac{a^2}{3} \right) \left( m + (m - b)/(m - b)^2 \right) \quad (B-16)
\]

\[
= \left( \frac{a^2}{3} \right) [2m - b/(m - b)^2] \quad (B-17)
\]

In the present case this gives,

\[
\text{TVP} = \frac{16(2\sqrt{13} - 3)}{3(\sqrt{13} - 3)^2} \quad (B-18)
\]

\[
= 61.24817. \quad (B-19)
\]

To arrive at the solution y-value for AVP, we may, of course, use the original equation for AVP, or we may simply divide the above equation by \( x^* \) and get,

\[
\text{AVP} = \frac{a/3(m/(m - b) + 1)}{\frac{4}{3}(\sqrt{13}/(\sqrt{13} - 3) + 1)} \quad (B-20)
\]

\[
= 9.27222. \quad (B-21)
\]
Figure 7

TVP = 4x + 3x^2 - x^3/3

MVP = 4 + 6x - x^2

AVP = 4 + x - x^2/3

Quantity of Input

$
Figure 8

AVP = 4 + 3x - x^2/3

MVP = 4 + 6x - x^2

--- Series1

--- Series2

Quantity of Input
Figure 10

Area A = Area B
UNLEASHING LEVIATHAN: PUBLIC GOODS UNDER INVOLUNTARY TAXATION

Johnnie B. Linn III
Concord University

ABSTRACT

Wage rates are forced down when government extracts taxes from firms. Labor will migrate from taxed firms to firms under an anarchistic regime unless government can generate a public good that will overcompensate for the fall in wages. A particular public good, law enforcement, provides positive feedback to the efficacy of the tax collectors who fund government because they have recourse to it against firms. A tax rate above which Leviathan can compete with anarchy can be determined for given production and force technologies. In the United States, the current before-taxes tax rate is about three times the threshold tax rate needed for Leviathan. The ratio of public goods volume to non-public goods outlays is about 2.27. This ratio can be maintained at higher tax rates but at the cost of further reduction in the wage. At a sufficiently high tax rate for a given ratio, Leviathan can no longer compete with anarchy.

INTRODUCTION

This paper is the sequel to a previous work, in which Linn (2007) explored conditions under which competitive firms could exist in the absence of government, a condition which is described in Thomas Hobbe’s Leviathan as bellum omnium contra omnes, or the “war of all against all”. Here, Leviathan, or the state, is introduced. Its income is generated solely through involuntary taxation, in a levy against firms which is resisted by those firms by the use of force. Conditions are identified under which Leviathan can be a going concern, meaning when it can improve upon the anarchistic model.

The model of the state to be developed takes as given the type of anarchistic equilibrium derived in the previous work and will take a form that can be shown to arise from that equilibrium.

Force is violence, stealth, or fraud, or defenses to the same, employed by two or more agents against each other with the purpose of acquiring, controlling, or defending something of value. For mathematical convenience, a ratio rule of force is assumed, under
which winnings are exhaustive and the share of winnings to the \( i \text{th} \) agent employing force \( F_i \) is

\[
\phi_i = \frac{F_i}{\sum_j F_j}
\]  

(1)

This rule is used in Ozenne’s (1974) model of bank robbery, where a bank robber’s share of the total amount taken from a bank is in proportion to that robber’s share of effort expended to the total amount of effort expended against that bank.

Hirschleifer (1989) shows that in a ratio model of force, no stable equilibria are possible where force is used by no one.

As in the anarchistic equilibrium, firms produce force and output, using labor that is employed either as workers or guards. Unemployed private individuals constitute “outsiders” who use force but who do not act collectively on the margin. The quantity of outsiders at a particular time is considered to be a fixed parameter, like a price.

Government force is introduced as tax collectors who, like those of New Testament times, forward funds to government that they have extracted from private individuals. Output is divided among firm, outsiders, and tax collectors according to their relative proportions of force employed. The government is not replacing guards, as we have in Grossman (1997), but is fighting them. The output of each firm is fought over in an arena that includes one firm, outsiders, and tax collectors. The number of arenas is sufficiently large that the firms face a purely competitive output market. The tax collectors’ only function is to generate revenue. Other ways that government might use force, such as police or military, are assumed not to be for the purpose of raising revenue.

THE GENERAL MODEL

As before, the own-force elasticity of a user’s share of winnings is

\[
\frac{F_i}{\phi_i} \frac{\partial \phi_i}{\partial F_i} = 1 - \phi_i.
\]  

(2)

and the cross-force elasticity of winnings is

\[
\frac{F_i}{\phi_j} \frac{\partial \phi_j}{\partial F_i} = -\phi_i, \quad i \neq j.
\]  

(3)

The production functions for the firm’s output and force in elasticized form are

\[
\frac{L}{y} = \alpha,
\]  

(4)

and

\[
\frac{G}{F_G} = \beta,
\]  

(5)

where \( L \) is the quantity of workers and \( G \) is the quantity of guards. As explained in the Appendix of the original paper, Marshallian stability requires a value of \( \beta \) greater than unity. The upper bound for \( \alpha \) in a purely competitive output market is unity.

58
The production function for the government’s force in elasticized form is

$$\frac{T}{F_T} \frac{dT}{d\beta} = \beta,$$

(6)

where $T$ is the number of tax collectors. To simplify the mathematics, the assumption is made that the force elasticity for tax collectors is the same as that for guards.

**CASES TO BE CONSIDERED**

Two polar scenarios will be considered. At one extreme, the base of the tax is the winnings of those individuals whose presence engenders the argument of the need for government—the outliers. At the other extreme, the base of the tax is the income of those individuals who benefit from the presence of government—the firms. To further simplify the mathematics, the government will not be constrained to balance its budget on the margin; instead it will direct the tax collectors to impose a tax on the base at a constant after-tax rate $t$ defined on the range $[0, \infty]$. The balanced-budget condition for the government is imposed only at the equilibrium. The after-tax shares of winnings to firms, tax collectors, and outliers respectively are represented as $\phi$, $\phi_T$, and $\phi_O$.

**Base of the tax is outlier income**

When the base of the tax is outlier income, the quantity of tax collectors is fixed because the number of outliers from which they extract income is fixed, and amount they extract from them is fixed. The value of $t$ therefore does not appear in any first-order conditions and the zero profit equilibrium condition for the firm is the same as in Linn (2007), or

$$\bar{\phi} = \frac{a+\beta-1}{\beta}.$$  

(7)

The shares of winnings for the outliers and tax collectors is in proportion to the first and second terms in the brackets of Equation (8) respectively and their sum is invariant.

$$\bar{\phi}_T + \bar{\phi}_O = \left(1-\frac{a}{\beta}\right) \left[\frac{1}{(1+t)} + \frac{t}{(1+t)}\right]$$

(8)

If the arenas are not too tiny—each having less than two individuals using force—the competitive equilibrium with government will exhibit a higher wage than the anarchistic equilibrium, and workers will abandon the anarchistic arenas. We can see this by starting from the special case where the government has hired precisely one, and its first, tax collector from the pool of outliers in an arena. There is no effect on the firm; it is facing a combined force having one less outlier and one tax collector, and one tax collector alone cannot gain from having a force elasticity greater than unity. The effect, then, of hiring precisely one tax collector out of a pool of outliers is neutral. We can go on to infer that having more than one tax collector will make an arena with government more attractive than an anarchistic one. Tax collectors then employ their
greater force elasticity, and fewer tax collectors than outliers would be needed to produce the same amount of force opposing the firm. The government thus generates a surplus, which can be devoted to production of a public good.

The amount of social surplus generated under this tax regime would be small. Liun (2007) shows that outlier income in the anarchistic equilibrium is small to start with. At best, government can capture and redistribute no more than what all outliers had under anarchy, so government does not grow sufficiently monstrous to merit the name of Leviathan. We turn, therefore, to the alternate case in which the base of the tax is firm income.

**Base of the tax is firm income**

For a constant after-tax rate $t$ on firm income net of outliers, the force production of the tax collectors is in fixed proportion to that of the firm, and the firm’s share of winnings takes the form, from Equation (1),

$$
\phi = \frac{F_G}{F_G + (1+t)F_U},
$$

(9)

where $F_G$ and $F_U$ are the quantities of force produced by the firm’s guards and by outliers, respectively. From Equation (2), the firm’s own-force elasticity of its winnings is

$$
\frac{\frac{\partial \phi}{\partial F_G} \cdot \phi}{\phi} = 1 - \phi(1 + t).
$$

(10)

The profit maximizing function of the firm is

$$
\pi = \phi Y - wL - wG.
$$

(11)

The wages of all forms of labor are assumed to keep the same ratio with each other. Subscripts for $w$ will not be carried in the equations because a given wage can be linked with its factor from the context. The firm’s first-order condition for its workers is

$$
\frac{\partial \phi Y}{\partial L} - w = 0,
$$

(12)

and for its guards is

$$
\frac{\partial \phi Y}{\partial G} = \frac{\phi(1 + t)}{\phi} - w = 0.
$$

(13)

By inserting Equations (12) and (13) into Equation (11), we find the firm’s competitive equilibrium profit to be

$$
\tilde{\phi} Y [1 - \alpha - \beta (1 - \tilde{\phi})(1 + t)] = 0.
$$

(14)

and its equilibrium share of winnings is
\[ \bar{\phi} = \frac{\alpha + \beta - 1}{\beta(1+e)} \]  

(15)

The tax collectors' share of winnings is solely at the expense of the firm as can be seen from calculating the winnings of outliers:

\[ \bar{\phi}_y = 1 - \bar{\phi}(1 + t) = \frac{1 - \alpha}{\beta}, \]

(16)

which is the same as if there had been no tax.

For constant values of \( \alpha \) and \( \beta \), the second-order sufficiency condition for the firm is satisfied in the neighborhood of the competitive equilibrium, as demonstrated in the Appendix.

In each of Equations (12) and (13), the wage, or net marginal product, of labor is forced down compared to what is found in the anarchistic equilibrium. So how can arenas with government compete with arenas without it? An appropriate public good must be found whose value will more than compensate for the cost of government.

Selecting an appropriate public good

We must find a public good that can be applied initially at the scale of one arena. Thus, a public good such as a military force to defend a collection of arenas against attack by a foreign government would be infeasible, as anarchistic arenas being defended would be free-riding on the others.

One possible public good is protection of an arena from organized force-using groups—other than the tax collectors themselves—that could operate in that arena. In other words, the government establishes itself as the arena’s organized force-only-using monopoly. The assumption need not be made that the government is not a rent-seeker. Powell and Stringham (2008) review papers that show that government may arise to serve its own interests but nonetheless create a situation that is preferred by all to anarchy.

Another possible public good is protection of property against fires. Anderson (ca. 1980) describes how fire departments were first formed by fire insurance companies to protect buildings whose owners had paid for fire insurance. The buildings bore the appropriate “fire mark”, a token indicating evidence of payment.

We presume that a public good exists that can be produced at the arena level, and we proceed to the problem of funding the good.

GENERATING A GOVERNMENT SURPLUS

Let us suppose that the government budget at equilibrium is given by

\[ \bar{\phi}_T = b(w^T + w^F_0 + rK_0), \]

(17)

where \( b \) is a coefficient whose value, it is hoped, is greater than unity, \( T_0 \) represents "inert" government workers—those who are not tax collectors and who produce goods and services that are not pure public goods, and \( rK_0 \) is the amount paid to financial capital for
the government non-public goods. The labor and capital outlays for the public goods are not explicitly stated in Equation (17) because they are to be paid from the government surplus.

The difference between $b$ and unity is the government surplus expressed as a percentage of tax collector and non-public goods outlays. If the government surplus were to be used to purchase a purely private good to be distributed equally over the total population of the arena, the government surplus per capita for the whole population would be a fraction of the government surplus per government worker. But if the surplus is used to purchase a pure public good, there would be no diminution of the per capita value of the surplus over the total population.

Figure 1 is a combination of two graphs that show the relationship between the government surplus per capita and the quantity of the public good produced. The right hand side shows the optimum quantity of the public good. The non-public goods outlays include the compensation of tax collectors and the slope of the non-public goods outlays line is the wage. The horizontal line linking the two sides of the figure shows the government surplus required for the optimal quantity of the public good. Optimization of the public goods does not necessarily imply that government surplus is also maximized, but government planners might lose track of the distinction. The magnitude of the government surplus is also the public good's contribution to gross domestic product.

If the value of the public good is exactly its contribution to gross domestic product, the quality of life for every member of the population is given by the expression $bw$ and the test for viability of a government regime is whether the value of $bw$ is greater than that of $w$ in the anarchistic regime. If the value of the public good is greater than its contribution to gross domestic product, the $bw$ test is a sufficiency condition.

FINDING THE PARAMETER ESTIMATES

In estimation of the parameters using modern data, we must keep in mind that taxation has evolved from the times of bare emergence from anarchy. Firms do not "fight" tax collectors now; it is better to say that firms use part of the protective services they employ to safeguard information about their tax reduction strategies, including fraud, while the tax collectors employ suitable countermeasures, among which is recourse to law enforcement.

Also, in considering modern data, a sizeable proportion of taxes and transfer payments involve individuals. In this model, taxes on labor income are imputed to be internalized by firms in paying higher wages than what workers would have asked for in an anarchistic, tax-free economy. Likewise, outliers extract winnings not only from firms but also from workers' earnings, and labor's losses to outliers will be internalized by firms.

In Equation (17) we replace $\bar{\nu} Y$ with its equivalent in terms of the firm's winnings and obtain

$$\bar{\tau} \bar{\nu} Y = b(wT + wT_0 + rK_0)$$  \hspace{1cm} (18)

Then, from Equation (12),

$$b = \frac{wL}{a(wT + wT_0 + wK_0)}$$  \hspace{1cm} (19)
To find the statistic for $\alpha$, we combine Equations (12), (13), and (16) to obtain

$$\frac{\alpha}{1-\alpha} = \frac{wL}{wG}$$  \hspace{1cm} (20)

The left-hand side of Equation (20) is defined over the interval $[0, \infty)$ and will be used as the statistic for estimating the confidence interval for $\alpha$.

For estimation of $\beta$ we combine Equations (20) and (16) and obtain

$$\beta \frac{G}{G} = \frac{wG}{wL+wg}$$  \hspace{1cm} (21)

When $\beta$ is a constant, the force outputs are exponential functions of the force effort inputs. We then use Equation (1) to obtain

$$\frac{t\beta}{\phi} = \frac{(cT)^\beta}{\phi},$$ \hspace{1cm} (22)

where $c$ is a coefficient of the comparative potency of tax collectors and guards. Unlike the factor ratios used in Equations (19), (20), and (21), the quantities here are the raw factor quantities, not their compensations. Though labor hired as tax collectors or guards can be regarded as homogenous, the value of $c$ is greater than unity because tax collectors facing guards have recourse to government force while guards do not. The value of $c$ is calibrated to generate a $t$ that matches the observed value of $t$.

THE THRESHOLD CONDITION FOR GOVERNMENT VIABILITY

If the government drives down the equilibrium wage, it can remain a viable competitor in the market if it can generate a public good whose value, when combined with the money wage, exceeds the wage that would prevail in the anarchistic markets.

Taxation causes the wage to fall by a factor $1/(1+t)$ compared to the anarchistic equilibrium. Then, for government to be competitive with anarchy it is sufficient that the value of $b$ be greater than $(1+t)$. From Equation (19), we find the threshold condition

$$b = \frac{twL}{a(wT + wT_0 + rK_0)} \geq 1 + t,$$

or

$$\frac{t}{1+t} \geq \frac{a(wT + wT_0 + rK_0)}{wu}.$$  \hspace{1cm} (23)

The expression on the left-hand side is the tax rate expressed as a percentage of the before-taxes tax base and is defined on the interval $[0,1]$. For a given level of taxation, there is an upper limit on the compensation of government employees, as would be expected, and for a given compensation of government employees, a minimum tax level is needed.
The government surplus is sensitive to government non-public good outlays and interest payments to capital. To maintain a constant government surplus multiplier, the government would have to raise taxes to accommodate additional non-public goods and capital expenditures, but the higher taxes would drive down the wage, and there is a maximum tax rate at which the government could no longer compete with the anarchistic equilibrium.

A NOTE ON THE COMPENSATION OF CAPITAL

Except for compensation of financial capital, the data below show labor compensation only. The magnitudes of \( \alpha \) and \( \beta \) are calculated below by use of labor compensation ratios as if labor is the sole contributor to national income. We will assume that the ratios of physical capital compensation to labor compensation are constant across occupation categories. We can then accept the labor data as indicative of the whole economy without having to include capital data explicitly.

THE DATA

Table 1 shows the data and their sources. The years 2002, 2007, and 2012 for the United States are used.


Data for outlier winnings are inferred from reported losses due to robbery and crimes against property in the Crime in the United States reports (Federal Bureau of Investigation, 2014a, 2014b, and 2014c). Losses due to robbery for 2002 are found at Federal Bureau of Investigation (2014a), at p. 32 for robbery and p. 42 for crimes against property. Losses for 2007 are found at Federal Bureau of Investigation (2014b). Losses for 2012 are found at Federal Bureau of Investigation (2014c).

Losses due to fraud are not included in the Federal Bureau of Investigation reports.

This paper will exclude outlier winnings due to fraud and costs of countermeasures to fraud in the calculations involving force. The OES category for protective services, 33-0000, is confined to occupations that protect persons and property, and contains no occupations specifically combating fraud, nor, for obvious reasons, occupations perpetrating fraud. Occupations that combat insurance fraud are found in OES category 13-1030 for claims adjusters, appraisers, examiners, and investigators, but countermeasures for no other types of fraud appear to have corresponding OES categories.
The best approach, then, is to use the available FBI data for all outlier winnings, and the available OES data for all guard outlays.

IMPLICATIONS OF THE RESULTS

The value of $\alpha$ is 0.9936 with a standard deviation of 0.00008, very close to the value of unity that would be seen in the conventional competitive firm that returns all the value of its output to its factors. The value of $\beta$ is 2.03 with a standard deviation of 0.39. The value of $\alpha$ is more precisely estimated than the value of $\beta$ because the latter is sensitive to outlier winnings, which are hard to measure.

The magnitude of the government surplus multiplier is 2.27 with a standard deviation of 0.13. The average before-taxes tax rate, with net domestic product as the base, is 0.319. The minimum before-taxes tax rate for government viability is 0.154. In the range of increase of the tax rate from 0.154 to 0.319, the government exhibits an approximate unit elasticity in the size of the multiplier with respect to the tax rate.

The maximum before-taxes rate at which the government could remain competitive with the anarchistic equilibrium, keeping its current government surplus multiplier, is 0.559. Currently, Leviathan sits in a wide range of lower and upper tax rate limits for viability, but increased non-public good outlays or interest payments could narrow that range.

REFERENCES


APPENDIX

The elements of the Hessian matrix are

\[
H_{11} = \frac{-\alpha \phi Y}{L^2} + \frac{\alpha^2 \phi Y}{L^2}; \\
H_{12} = H_{21} = \frac{\alpha \beta \phi [1 - \phi (1 + \epsilon)] Y}{L G}; \\
H_{22} = \frac{-\beta \phi [1 - \phi (1 + \epsilon)] Y}{G^2} + \frac{\beta^2 \phi [1 - \phi (1 + \epsilon)]^2 Y - \beta^2 \phi^2 [1 - \phi (1 + \epsilon)] Y}{G^2}. \tag{A3}
\]

A quantity \( \phi Y \) can be extracted from each row of the determinant and brought outside as a multiple. Likewise, the inverses of \( L \) and \( G \) can be extracted by two row operations and two column operations, respectively. Also, a quantity \( \alpha \) can be extracted from the first row and a quantity \( \beta \) from the second row. Also, the quantity \([1 - \phi (1 + \epsilon)]\), which is the outliers’ share of winnings, can be extracted from the second row. All these multiples are positive. The determinant has become

\[
\Delta = \frac{\alpha \beta \phi^2 Y^2}{L^2 G^2} \left[ 1 - \phi (1 + \epsilon) \right] \begin{vmatrix} \alpha - 1 & \beta [1 - \phi (1 + \epsilon)] \\ \alpha & \beta [1 - \phi (1 + \epsilon)] - \beta \phi - 1 \end{vmatrix}. \tag{A4}
\]

At the competitive equilibrium, we substitute in the firm’s and outliers’ shares of winnings from Equations (15) and (16) and obtain

\[
\Delta = \frac{\alpha \beta \phi^2 Y^2}{L^2 G^2} \left[ 1 - \phi (1 + \epsilon) \right] \begin{vmatrix} \alpha - 1 & 1 - \alpha \\ \alpha & -\alpha - \beta \phi \end{vmatrix}. \tag{A5}
\]

We see that the diagonal terms are negative. In the final step, we find that

\[
\Delta = \frac{\alpha \beta \phi^2 Y^2}{L^2 G^2} (1 - \alpha) \gamma \beta \phi \left[ 1 - \alpha - \beta \phi \right] > 0, \tag{A6}
\]

so the second-order sufficiency conditions are met.
Figure 1. Government Funding for the Public Good for a Given Wage.
<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>SOURCE</th>
<th>2002</th>
<th>2007</th>
<th>2012</th>
<th>DESIGNATOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net Domestic Product</td>
<td>BEA NIPA Table 1,17,5 line 3</td>
<td>$9,318.1 billion</td>
<td>$12,216.0 billion</td>
<td>$13,701.7 billion</td>
<td></td>
</tr>
<tr>
<td>Government Current Receipts</td>
<td>BEA NIPA Table 3.1 line 1</td>
<td>$2,967.0 billion</td>
<td>$4,001.8 billion</td>
<td>$4,259.2 billion</td>
<td></td>
</tr>
<tr>
<td>Ratio of Government Current Receipts to Net Domestic Product</td>
<td></td>
<td>0.318</td>
<td>0.328</td>
<td>0.311</td>
<td>$t / (1 + t)</td>
</tr>
<tr>
<td>After-Taxes Tax Rate</td>
<td></td>
<td>0.462</td>
<td>0.488</td>
<td>0.451</td>
<td>$t</td>
</tr>
<tr>
<td>All Non-Military-Specific Labor Income</td>
<td>OES-00-0000</td>
<td>$4,534.74 billion</td>
<td>$5,466.87 billion</td>
<td>$5,965.87 billion</td>
<td></td>
</tr>
<tr>
<td>Non-Military-Specific Gov't Labor Income</td>
<td>NAICS 99-9000 component of OES 00-0000</td>
<td>$414.09 billion</td>
<td>$456.08 billion</td>
<td>$526.70 billion</td>
<td></td>
</tr>
<tr>
<td>All Private-Sector Labor Income</td>
<td>All Non-Military-Specific Labor Income less Non-Military Specific Gov't Labor Income</td>
<td>$4,120.65 billion</td>
<td>$5,010.79 billion</td>
<td>$5,540.09 billion</td>
<td></td>
</tr>
<tr>
<td>Protective Services Income</td>
<td>OES 33-0000</td>
<td>$99.77 billion</td>
<td>$119.65 billion</td>
<td>$138.10 billion</td>
<td></td>
</tr>
<tr>
<td>Non-Military Public Goods Income</td>
<td>NAICS 99-9000 component of OES 33-0000</td>
<td>$73.38 billion</td>
<td>$86.95 billion</td>
<td>$102.50 billion</td>
<td></td>
</tr>
<tr>
<td>Guard Income</td>
<td>Protective Services Income less Non-Military Public Goods Income</td>
<td>$26.39 billion</td>
<td>$32.70 billion</td>
<td>$35.60 billion</td>
<td>$w_G</td>
</tr>
<tr>
<td>Worker Income</td>
<td>All Private-Sector Labor Income less Guard Income</td>
<td>$4,094.26 billion</td>
<td>$4,978.09 billion</td>
<td>$5,504.89 billion</td>
<td>$w_L</td>
</tr>
<tr>
<td>Tax Collector Income</td>
<td>OES 13-2081</td>
<td>$3.17 billion</td>
<td>$3.39 billion</td>
<td>$3.67 billion</td>
<td>$w_T</td>
</tr>
<tr>
<td>Non-Public-Good Non-Military-</td>
<td>Non-Military Specific Gov't Labor Income less</td>
<td>$340.71 billion</td>
<td>$369.13 billion</td>
<td>$424.20 billion</td>
<td>$w_T_0</td>
</tr>
<tr>
<td>Specific Gov't Labor Income</td>
<td>Non-Military Public Goods Income</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------------------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>Gross Gov't Investment</td>
<td>BEA NIPA Table 3.1, line 34</td>
<td>$443.6 billion</td>
<td>$592.2 billion</td>
<td>$619.0 billion</td>
<td></td>
</tr>
<tr>
<td>Gross Gov't Investment for Defense</td>
<td>BFA NIPA Table 3.9.5 line 19</td>
<td>$98.7 billion</td>
<td>$152.7 billion</td>
<td>$165.1 billion</td>
<td></td>
</tr>
<tr>
<td>Gov't Net Interest Paid</td>
<td>BEA NIPA Table 3.1 line 22 less line 9</td>
<td>$288.2 billion</td>
<td>$380.8 billion</td>
<td>$523.7 billion</td>
<td></td>
</tr>
<tr>
<td>Non-Defense Gov't Capital Outlay</td>
<td></td>
<td>$633.1 billion</td>
<td>$820.3 billion</td>
<td>$977.6 billion</td>
<td></td>
</tr>
<tr>
<td>Ratio of Non-Military-Specific Labor's Income to Net Domestic Product</td>
<td>0.487</td>
<td>0.448</td>
<td>0.435</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Military-Specific Labor's Share of Non-Public-Gov't Capital Outlay</td>
<td>$308.3 billion</td>
<td>$367.5 billion</td>
<td>$425.7 billion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robbery Losses</td>
<td>FBI</td>
<td>$0.593 billion</td>
<td>$0.588 billion</td>
<td>$0.414 billion</td>
<td></td>
</tr>
<tr>
<td>Property Crime Losses</td>
<td>FBI</td>
<td>$16.6 billion</td>
<td>$17.6 billion</td>
<td>$15.5 billion</td>
<td></td>
</tr>
<tr>
<td>Outlier Winnings Excluding Fraud</td>
<td></td>
<td>$17.1 billion</td>
<td>$17.8 billion</td>
<td>$15.9 billion</td>
<td></td>
</tr>
<tr>
<td>Non-Military-Specific Labor's Share of Non-Public-Gov't Total Outlay</td>
<td>$652.0 billion</td>
<td>$739.6 billion</td>
<td>$853.5 billion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Military-Specific Labor's Share of Government</td>
<td></td>
<td>$1,444.9 billion</td>
<td>$1,792.8 billion</td>
<td>$1,852.8 billion</td>
<td></td>
</tr>
</tbody>
</table>

$ rK_0$

$ wU$
<table>
<thead>
<tr>
<th>Current Receipts</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gov't Surplus Multiplier</td>
<td>2.22</td>
<td>2.42</td>
<td>2.17</td>
<td>b</td>
</tr>
<tr>
<td>Protective Services Employees</td>
<td>OES 33-0000</td>
<td>2,993,490</td>
<td>3,087,650</td>
<td>3,207,790</td>
</tr>
<tr>
<td>Non-Military Public Goods</td>
<td>NAICS 99-9000 component of OES 33-0000</td>
<td>1,798,900</td>
<td>1,834,700</td>
<td>1,910,240</td>
</tr>
<tr>
<td>Employees</td>
<td>Guards</td>
<td>Protective Service Employees less Non-Military Public Goods Employees</td>
<td>1,194,590</td>
<td>1,252,950</td>
</tr>
<tr>
<td>Tax Collectors OES 13-2081</td>
<td>69,320</td>
<td>65,750</td>
<td>65,560</td>
<td>T</td>
</tr>
<tr>
<td>Tax Collector Advantage</td>
<td>12.16</td>
<td>14.17</td>
<td>13.71</td>
<td>c</td>
</tr>
<tr>
<td>Statistic for $\alpha$</td>
<td>155.14</td>
<td>152.24</td>
<td>151.79</td>
<td>$\alpha / (1 - \alpha)$</td>
</tr>
<tr>
<td>Statistic for $\beta$</td>
<td>1.69</td>
<td>1.96</td>
<td>2.45</td>
<td>$\beta$</td>
</tr>
<tr>
<td>Worker Output Elasticity ($\alpha$)</td>
<td>0.9936</td>
<td>0.9935</td>
<td>0.9935</td>
<td>$\alpha$</td>
</tr>
<tr>
<td>Threshold Before-Taxes Rate for Government Viability</td>
<td>0.158</td>
<td>0.148</td>
<td>0.157</td>
<td></td>
</tr>
<tr>
<td>Upper Limit for Before-Taxes Tax Rate with Current Multiplier</td>
<td>0.549</td>
<td>0.587</td>
<td>0.539</td>
<td></td>
</tr>
</tbody>
</table>
A TRADE NETWORK WITH A SINGLE TRADER UNDER ASYMMETRIC INFORMATION

Sabri Yilmaz
Penn State University-Harrisburg

ABSTRACT

Buyer and seller interactions are analyzed with intermediaries called traders using a network structure. Goods are traded in the market through those networks. Each seller and buyer is linked to a trader through a network but not directly linked to each other. Our main contribution to the literature is that we introduce asymmetric information on the valuation of goods by sellers and buyers. Our model basically consists of a two-stage game with incomplete information. First, the trader offers bid prices to the sellers and ask prices to the buyers. Second, trade takes place when the sellers and/or buyers accept the trader’s offers. The trader tries to maximize his profit. We analyze even network structures with the same number of buyers and sellers tied to one trader. When the penalty is almost zero, we find that trader never sets the same bid and ask prices to the seller and the buyer, respectively.

INTRODUCTION

Buyer and seller interactions play an important role in the network literature which includes studies by Rubinstein and Wolinsky (1987), Li (1998) and Kranton and Minehart (2001). These interactions have been investigated and modeled in the previous literature such as Blume et al. (2007) analyzing buyer-seller networks through intermediaries. These intermediaries play the role of middlemen, which we call traders in our model. The goods are traded in the market through those networks. Examples of goods traded through buyer-seller networks include the Japanese electronics industry and the Italian textile industry, mentioned in Nishiguchi (1994) and Lazerson (1993), respectively. Nishiguchi (1994) discusses the problematic cases in the theories of the contractual relations of the Japanese subcontracting system. He explains the positive effect of subcontracting on the improvement of Japanese economy especially for the automobile and electronics industry. Besides that, Lazerson (1993) emphasizes that the success and the development in these industries stem from the well-established and the specialized network
structures between small and medium-sized textile firms in Modena. The efficient structural form in both cases leads to improvement in the regional and the whole economy.

Our ultimate goal is to find the conditions for the maximization of the trader’s profit under asymmetric information. The conditions are dependent upon the valuations of the traded goods by the agents and the probabilities of having those valuations. Moreover, these conditions reveal the maximum profit for the trader or when he chooses not to make any offers that lead him to be worse off. In the framework of our model, each seller and buyer are linked to a trader through a network. However, there are no direct links between sellers and buyers. Our main contribution to the literature is that we introduce asymmetric information on the valuation of goods by sellers and buyers. A buyer or a seller knows his valuation for the good but not anyone else’s. Also, the trader does not know the buyer’s or the seller’s valuations.

The model is a two-stage game with incomplete information. The trader’s goal is to maximize his profit. In the first stage, trader offers bid prices to the sellers and ask prices to the buyers. In the second stage, trade takes place between the trader and the sellers/buyers who accept the offers made by the same trader. Initially, we assume that each seller has only one unit of the good. The expected payoff to the trader is computed at the end of the second stage. We calculate the differences between the accepted ask and bid prices by the buyer and seller, respectively. Next, we multiply them by the corresponding probabilities of the valuations of the good for the seller and buyer accepting the offers. One should note that trader will end up with the good itself if he gets the good from the seller but the buyer does not agree to buy the good from him due to the buyer’s valuation. In that case, his profit from this type of trade will only be his own valuation of that good minus the bid price he pays off. If the seller does not accept the bid offer by the trader, then the trader is not allowed to promise to sell and his payoff will be zero. When the number of the buyers whom the trader has an agreement with exceeds the number of the sellers accepting trader’s bid offers, he will be subject to a penalty which is a value greater than zero. The implementation of the penalty concept is used in order to avoid possible cheating to gain more profit by the trader. This penalty was first introduced by Blume and et al. (2007). However, they focus on the trading networks with complete information in their study.

Next, we discuss the closely related literature: Blume et al. (2007). Authors construct a model using a game in which sellers, buyers and traders engage in trade on a graph that allows the buyers and sellers to tie to the traders. The model emphasized in the present paper is related to their model on trading networks. One difference in the setup is that they deal with complete information networks at which traders know the valuations of the goods by the sellers and buyers while we deal with asymmetric information. Also, buyers and sellers do not know each other’s valuation. In this setup, traders do not know the valuations of the sellers and the buyers. While they mainly focus on the equilibrium concept, we would like to concentrate on the offer schedules by the trader to maximize his expected payoff. Hence, authors show that the resulting game always yields a subgame perfect Nash equilibrium and that all equilibria give an efficient allocation of goods. Our main contribution to the literature at this point is that we deal with the trader’s welfare maximization problem under asymmetric information.

Rubinstein and Wolinsky (1987) basically analyze the market with sellers, buyers and middlemen in which any trade between buyers and sellers can occur directly with each other or indirectly through the middlemen. Their study investigates the steady state results with a constant number of different type of agents over time. The matching process is
defined by the probabilities that any agent of type i has of meeting an agent type j at a
given period. Authors establish a framework to analyze the activity of middlemen and
determine the extent of his activity endogenously. The main difference between this paper
and our study is that they concentrate on the distribution of gains from this trade to each
party which takes place in this trade without any network structures. On the contrary, we
are interested in the formation of networks in any trade setting.

Moreover, Condorelli (2009) deals with asymmetric information in his model
where the author is interested in the equilibrium concept of trading outcomes influenced by
the networks. He establishes a dynamic bilateral trading model with networks consisting of
multiple traders taking a part in a market for one good. He finds out that the traders staying
in latter positions in this chain of agents pay smaller amount than the previous agents
although those players have a lower chance of purchasing the good as it travels from one
trader to another in this chain. Our model is different when compared to the model by
Condorelli (2009) because only one agent initially has a good to sell and anyone connected
to him can become a “trader” in his model whereas there are multiple agents who have
goods to sell and there is a fixed set of traders in our model. Li (1998) describes the private
information about the quality of consumption goods as an instrument to motivate the role
of middlemen. It is discussed that an agent endogenously chooses to become a middleman
through investment in a technology of verifying quality. In our model, we do not have any
investment related to the good itself; instead we investigate the conditions at which the
middleman maximizes his profit under private information about the sellers and the buyers.

Sarma et al. (2007) focus on market models at which transactions are carried out
through largely unknown intermediaries or brokers who have an important role in the
setting. Buyers would like to buy goods at minimum prices through brokers. Moreover,
Nermuth et al. (2009) study a model with price competition. Authors emphasize that their
model could be considered as a network model of price competition. Finally, Khoury et al.
(2009) investigate a model for route distribution and using large scale networks where
distribution of the route is done by advertising its prefix to the direct neighbors in the
network. This is one example of possible applications into engineering.

Another application of trade networks is introduced by Yilmaz (2014), at which,
author investigates the effects of flipper activities on the real estate market outcomes for
Atlanta. Flippers act as traders, i.e. flipper buyers-flipper sellers, in his setting. The
suburban area is found to be more profitable for flippers as target areas under a similar
bidding/asking process as we have in our paper.

There are other buyer-seller network papers without traders, i.e. Kranton and
Minehart (2001). Authors find that buyers and sellers are able to maximize their welfare
through those formed network structures. One of the reasons why the buyer-seller networks
emerge is that the competition can improve when there are multiple links formed with the
trader. Also, In Kosfeld (2003), author mentions that Corominas-Bosch (1999) investigated
the network effect on competition in a network model where prices are determined through
a bargaining strategy instead of an English auction. Besides, there is no uncertainty on the
buyer valuation of the seller’s good. Their proposed network structure also reveals each
player’s bargaining power. This model is tested through an experiment by Charness,
Corominas-Bosch and Frechette (2001). Authors basically emphasize that the buyers use
the advantage of the bargaining power through this network. One should note that there are
no middlemen in this experiment. In conclusion, they argue that the attitude of the agents in
this formation is consistent with the theoretical predictions although the exact equilibrium
point is failed. One possible reason for that is the insufficiency in learning by which the
buyers and sellers develop a social link for a better bargaining strategy. Furthermore, the
experiments on dictator games imply that the bargaining power is less effective in those
types of games as mentioned in Roth (1995) and Camerer and Thaler (1995). Fehr and
Schmidt (1999) introduce the considerations of fairness and equity in order to explain this
phenomenon. Another network application in international trade, Rauch (1999), introduces
a network structure in differentiated products. His analysis supports that common colonial
ties have a significant role for differentiated products when compared to goods traded on
organized exchanges in buyer-seller matching games.

In the earlier literature, Myerson (1977) first introduced the network structure
dealing with the cooperation among individuals. Also, Jackson and Wolinsky (1996)
focused on the stability and efficiency economic networks where individuals form or break
links. They emphasized that a stable and efficient network cannot always be formed. Dutta
and Jackson (2001) and Jackson and Watts discussed the similar setup on the formation of
general networks and stable networks, respectively. Finally, Watts (2001) concentrated on
a dynamic model of network structures at which each individual could form or sever links
on the network.

The paper proceeds as follows. The model is presented in Section 2. We present
our results for the even number of buyers and sellers in Section 3. Finally, we give the
concluding remarks in Section 4.

THE MODEL

In our model, we have a trader who works as an intermediary between buyers and
sellers. We denote the set of traders, buyers and sellers by \{\{t_1, t_2, \ldots, t_n\}, \{b_1, b_2, \ldots, b_m\}, \{s_1, s_2, \ldots, s_l\}\}, respectively. We denote a particular network
by g and the set of all possible networks by G. If a buyer \(b_k\) is linked to a trader \(t_i\) in graph
\(g\), then the link between buyers and traders is represented by \(b_k t_i \in g\) whereas if a seller
\(s_j\) and a trader \(t_i\) have a link in between in graph \(g\), the link is shown by \(s_j t_i \in g\). One
emphasis here is that buyers and sellers are linked to traders; however they are not linked to
each other.

The model is a two-stage game with incomplete information. There is a network,
g through which a trader buys a good from a seller and then the same trader sells the good
to a buyer. In the first stage, traders basically offer bid prices, \(\beta_{ij}\) to the sellers and ask
prices, \(\alpha_{jk}\) to the buyers on the same network with them, where
\(i = 1, 2, \ldots, n; j = 1, 2, \ldots, m; k = 1, 2, \ldots, m\). Initially, each seller has only one unit of the good. In
the second stage, buyers and sellers accept or reject the trader’s offer. If both buyers and
sellers accept the trader’s bid and ask price offers, then the good is sold and the trader’s
payoff will basically be calculated by \(\alpha_{jk} - \beta_{ij}\), where \(\alpha_{jk}\) is the ask price offered by
trader \(t\) to buyer \(k\) and \(\beta_{ij}\) is the bid price offer by the same trader \(t\) to seller \(i\). Each
agent values the given good according to his own preference. Trader \(t\)’s valuation for that

76
good will be $\theta_i \geq 0$ whereas buyer $k$ has a valuation of $\theta_k \geq 0$ and seller $i$ values the
good by $\theta_i \geq 0$. If trader gets the good from the seller, but the buyer does not agree to buy
the good from the trader due to the fact that the ask price is greater than his valuation, then
it remains with the trader himself. In that case, trader's profit just includes his valuation of
that good. Traders are not allowed to promise to sell goods for which sellers refuse the bid
offers. Hence, trader gets a zero payoff since no trade occurs between them. When the
number of buyers whom the trader has an agreement with exceeds the number of sellers
accepting his offer, he is then subject to a penalty, $\rho > 0$.

Let us denote the number of sellers who accept trader $t$'s bid offer and the number of
buyers who accept the same trader's ask offer by $t_i$ and $t_j$, respectively. Then, the
payoff function for trader $t$ is given by the following:

$$
\pi_t = \begin{cases} 
\sum_{j=1}^{t_i} \alpha_j - \sum_{i=1}^{t_j} \beta_i, & \text{if } t_i = t_j \text{ and } ti \in g, tj \in g, \\
\left(\sum_{j=1}^{t_i} \alpha_j - \sum_{i=1}^{t_j} \beta_i\right) + \sum_{k=t_i+1}^{t} (\theta_k - \beta_k), & \text{if } t_i > t_j \text{ and } ti \in g, tj \in g, \\
\left(\sum_{j=1}^{t_i} \alpha_j - \sum_{i=1}^{t_j} \beta_i\right) - \rho \cdot (t_j - t_i), & \text{if } t_i < t_j \text{ and } ti \in g, tj \in g 
\end{cases}
$$

(1)

We impose a restriction in the form of uncertainty on the information about the
valuation of buyers and sellers. The uncertainty works as follows:

Seller $i$ who owns the good has the valuation $\theta_{i1}$ with some probability $p$ and $\theta_{i2}$
with some probability $(1 - p)$. Similarly, buyer $j$ has the valuation of $\theta_{j1}$ with probability
$q$ and $\theta_{j2}$ with probability $(1 - q)$. In our experiment, we restrict $p$ as $0 < p < 1$.
Therefore, the trader's expected payoff under asymmetric information becomes the product
of the corresponding probabilities of accepting the offers by seller $i$ and buyer $j$ times
$(\alpha_j - \beta_i)$ plus the probability that the offer is accepted by seller $i$, but rejected by buyer
$\rho \cdot (t_j - t_i)$ times the net value of the good which remains in the hands of the trader. The above
mentioned net value is computed by subtracting trader's bid offer to the seller from his
valuation of that good, i.e., $(\theta_i - \beta_i)$. In the case of neither seller nor buyer accepting the
offers, the trader's expected payoff turns out to be zero. If neither accepts, it is obvious that
trader ends up with nothing and obtains a zero payoff. We denote the expected payoff for
trader $t$ by $\pi_t$. The penalty comes into play, basically, when the number of accepted offers
by the buyers exceeds the number of accepted offers by the sellers. We assume that the
penalty, $\rho$ is some positive number.

For example, $\pi_t = \tilde{p} \cdot \tilde{q} \cdot (\alpha_j - \beta_i)$ is calculated to be the expected payoff for
trader $t$ when seller $i$ accepts the bid offer with a probability of $\tilde{p}$ and buyer $j$ accepts
the ask offer with a probability of $\tilde{q}$. If $(\alpha_j - \beta_i)$ is positive, our trader will gain from
trade, otherwise, trader \( t \) encounters a loss. On the other hand, if both seller \( i \) and buyer \( j \) reject the corresponding offers with probabilities \( (1 - \overline{p}) \) and \( (1 - \overline{q}) \), respectively, then \( \pi_t \) will become \( 1 \cdot (1 - \overline{p}) \cdot (1 - \overline{q}) \cdot 0 \). So, trader gets zero payoff.

We investigate two cases with the same number of sellers and buyers under network settings with one trader: one seller-one buyer and two sellers-two buyers. In each setting, we assume the valuations of the good are that the seller values the good by 0 with probability \( p \) and by \( \frac{1}{2} \) with probability \( (1 - p) \). In a similar fashion, the buyer values the good by \( \frac{1}{2} \) with probability \( q \) and 1 with probability \( (1 - q) \).

First, we consider the simple case with one seller, one trader and one buyer, where the trader is linked to the buyer and to the seller. The buyer accepts the trader’s offer if it is equal to or below his valuation whereas the seller accepts the trader’s offer if it is above or equal to his valuation of the good. Here, we find that the trader never sets the same bid and ask prices to the seller and the buyer, respectively. The expected profit the trader gets under that scenario is zero, which is always less than the payoffs he ever gets from the other possible offer schedules. We find that the trader receives the maximum expected profit when he offers different bid and ask prices. The other offer schedules with different prices yield maximum payoffs to the trader depending on the probability that the buyer values the good by \( \frac{1}{2} \) and the valuation of the good by the trader. Explicitly, when the probability that the buyer values the good by \( \frac{1}{2} \) is more than \( \frac{1}{2} \), i.e. the buyer is more likely to value the good by \( \frac{1}{2} \), the trader gets the maximum expected payoff by offering the bid price of 0 and ask price of \( \frac{1}{2} \). Otherwise, when the buyer is less likely to have the valuation of the good as \( \frac{1}{2} \), the trader maximizes his own profit through the bid price of 0 and ask price of 1. Also, we find that the seller considers any offer less than \( \frac{1}{2} \) as a low offer in this case. Hence, he is more likely to value his good by \( \frac{1}{2} \) and accepts any offers with bidding at least \( \frac{1}{2} \) for the good in order to receive a higher return. Finally, trader maximizes his expected payoff by offering \( \frac{1}{2} \) and 1 to the seller and the buyer, respectively, if the seller is less likely to value the good by 0. The buyer will also be more inclined to value the good by 1 at the same time due to the condition we obtain for this situation.

**RESULTS: One Seller and One Buyer**

We consider the following network where we have only one seller, one trader and one buyer:

```
seller    -->    trader    -->    buyer
```

Hence, the set notation for the graph of the above network is \( g = \{ s, t_1, t_2, b \} \). Seller \( i \) accepts nothing below \( \min(\theta_{i1}, \theta_{i2}) \) and buyer \( j \) does not accept any offers above \( \max(\theta_{j1}, \theta_{j2}) \). We solve this model using backwards induction, which works as follows:

First, we note that the buyer accepts the trader’s offer if it is equal to or below his valuation while the seller accepts trader’s offer if it is above or equal to his valuation of the good. Trader \( t \)'s goal is to maximize his expected payoff. The trader sets low bid price, \( \beta \) and a high ask price, \( \alpha \) in order to reach his goal. Therefore, he never sets \( \beta \) above...
max(θ₁, θ₂) and accordingly never sets α below min(θ₁, θ₂) since the trader is constrained by the buyer/seller’s acceptance.

Assume the valuations of the good are as follows:

For any seller, valuations are 0 with probability \( p \) and \( \frac{1}{2} \) with probability \( 1 - p \). Similarly, for any buyer, they are \( \frac{1}{2} \) with probability \( q \) and 1 with probability \( 1 - q \).

Recall that \( \beta_{ni} \) represents the bid price which trader \( t \) offers to seller \( i \), and \( \alpha_{iy} \) represents the ask price which the trader \( t \) offers to buyer \( j \). Then, it is obvious that trader \( t \) offers bid prices of \( \beta_{ni} = 0 \) or \( \frac{1}{2} \) and ask prices of \( \alpha_{iy} = 1/2 \) or 1 since \( 0 \leq \beta \leq \max(\theta_1, \theta_2) \) and \( 1 \leq \alpha \geq \min(\theta_1, \theta_2) \) along with the restriction specifying that the bid offer will not be accepted by the seller unless it is greater than or equal to 0 and the buyer will not agree on an ask price that exceeds 1 when \( \theta_1 = 0, \theta_2 = 1/2, \theta_1 = 1/2, \theta_2 = 1 \). Furthermore, since the trader’s goal is to maximize the difference between the ask prices and the bid price, we have:

\[
1 \geq \alpha - \beta \geq \min(\theta_1, \theta_2) - \max(\theta_1, \theta_2) \geq 0.
\]

Hence, under these specifications, trader chooses either \( \beta = 0 \) or \( \frac{1}{2} \) and either \( \alpha = 1/2 \) or 1 as he cannot gain from another offer schedule.

Second, we investigate the following cases thoroughly for low values of bid prices and high ask prices in order to maximize the profit for the trader:

1) if \( \beta_{ni} = 0 \) and \( \alpha_{iy} = 1/2 \), then trader’s payoff schedule will be as follows:

When both agents accept the bid and ask offers by the trader, trader gains \( \frac{1}{2} \) with a probability of \( p \) and \( q \), and the same value of \( \frac{1}{2} \) with probabilities \( p \) and \( 1 - q \) due to the fact that the buyer will definitely purchase the good at that ask price and the seller accepts the bid offer of 0 with probability \( p \). Hence, the expected payoff to trader is calculated as follows:

\[
p \cdot q \cdot \frac{1}{2} + p \cdot (1 - q) \cdot \frac{1}{2} + (1 - p) \cdot q \cdot (-p) + (1 - p) \cdot (1 - q) \cdot (-p)
\]

(2)

or

\[
\frac{1}{2} \cdot p - p \cdot (1 - p)
\]

(3)

after simplification.

2) if \( \beta_{ni} = 0 \) and \( \alpha_{iy} = 1 \), then:

Trader gains a value of 1 if both accept with probability \( p \) and \( 1 - q \), respectively. Otherwise, the seller accepts with probability \( p \), however, the buyer does not
accept the offer with probability $q$. In that case, trader ends up with the value of the good minus the price he pays to the seller, so that will lead to $(\theta_r - 0)$.

Therefore, the expected payoff to the trader:

$$p \cdot q \cdot \theta_r + p \cdot (1-q) \cdot 1 + (1-p) \cdot q \cdot 0 + (1-p) \cdot (1-q) \cdot (-\rho) = p \cdot [1-q \cdot (1-\theta_r)] - \rho \cdot (1-p) \cdot (1-q)$$

(4)

3) if $\beta_n = 1/2$ and $\alpha_q = 1/2$, then:

The gain to the trader will be equal to 0 in each case with different probabilities of $p \cdot q, p \cdot (1-q), (1-p) \cdot q$ and $(1-p) \cdot (1-q)$. Therefore, trader's payoff will be equal to 0.

4) if $\beta_n = 1/2$ and $\alpha_q = 1$. Then, the possibilities are:

i) seller accepts with probability $p$ and buyer accepts with probability $1-q$,

ii) seller accepts with probability $p$ and buyer does not accept with probability $q$,

iii) seller accepts with probability $(1-p)$ and buyer accepts with probability $(1-q)$,

iv) seller accepts with probability $(1-p)$ and buyer does not accept with probability $q$.

The corresponding expected payoff to trader becomes:

$$p \cdot 1-q \cdot \frac{1}{2} + p \cdot q \cdot (\theta_r - \frac{1}{2}) + (1-p) \cdot (1-q) \cdot \frac{1}{2} + (1-p) \cdot q \cdot (\theta_r - \frac{1}{2}) = \frac{1}{2} - q \cdot (1-\theta_r)$$

(5)

Here, one should note that the expected payoffs to trader for these four different cases depend on $p, q, \theta$ and the penalty, $\rho$. Now, we let $\rho \sim 0$. The summary of the results is given by the following four propositions:

**Proposition 1:** If $q \cdot (1-\theta_r) > \frac{1}{2}$, then trader bids 0 to the seller and asks $\frac{1}{2}$ to the buyer.

Here, $\frac{1}{2} p > 0$, which is the expected payoff to the trader in this case and $p$ is the probability value.

This result makes sense because $q$, as well, is the probability that the buyer values the good at $\frac{1}{2}$ and the product, $q \cdot (1-\theta_r)$ is quite possible to occur, i.e. higher than 50%, in this first case. Hence, trader will be able to get the highest payoff by offering 0 and $\frac{1}{2}$ to the seller and the buyer, respectively.
\[ \frac{1}{1-p} \cdot \frac{1}{2} < q \cdot (1 - \theta_i) < \frac{1}{2}, \]

**Proposition 2:** If \( \frac{1}{2} < q \cdot (1 - \theta_i) < \frac{1}{2} \), then trader bids 0 to the seller and asks 1 to the buyer.

Here, we again note that, for a probability value, \( p \), we exclude the case with \( p - 1 \), which is implied by the left hand side of the inequality. The product \( q \cdot (1 - \theta_i) \) has to be less than 50\% since the buyer's valuation is 1 with the probability of \( (1 - q) \).

We have such a high spread in this case due to the fact that the probability that the buyer values the good at \( \frac{1}{2} \) is low. Thus, the buyer is more likely to value the good at 1. Similarly, the seller considers any offer less than \( \frac{1}{2} \) as a low offer. So, he is more likely to value his good at \( \frac{1}{2} \) and accepts the offers with bidding at least \( \frac{1}{2} \) for the good to receive a higher return.

**Proposition 3:** Offering \( \beta_i = 1/2 \) and \( \alpha_q = 1/2 \) does not yield the maximum expected payoff to trader unless both \( p < 0 \) and \( q \cdot (1 - \theta_i) > 1 \), which are impossible.

However, \( q \) is a probability value between 0 and 1. Therefore, \( \theta_i \) would become negative. In conclusion, we obtain that the third case never gives us the maximum payoff for the trader.

**Proposition 4:** If \( q \cdot (1 - \theta_i) < \frac{1/2 - p}{1 - p} \), then trader bids 1/2 to the seller and asks 1 to the buyer.

This proposition is feasible since we recall that \( p \) is the probability that the seller values the good \( \frac{1}{2} \). That always allows him to accept the offer of \( \frac{1}{2} \) by the trader, i.e., both with probability \( p \) and \( (1 - p) \). The condition of \( p < \frac{1}{2} \) comes into play at this stage. The remaining, \( q \cdot (1 - \theta_i) > \frac{1}{2} \), follows due to the combination of this argument and the ask price of 1 to the buyer. Moreover, we notice that the two cases might have a common solution set for some values when we compare Proposition 2 with Proposition 4, mathematically. We also recall that, after letting \( p - 0 \), the expected payoff to the trader is \( p \cdot [1 - q \cdot (1 - \theta_i)] + \frac{1}{2} - q \cdot (1 - \theta_i) \) for Propositions 2 and 4, respectively. Depending on those specific values of \( p \), we can conclude that case 2 and 4 might yield the maximum expected payoff to the trader at the same time. In other words, the second and the fourth cases can lead to the maximum expected payoff under the conditions \( \theta_i = 0 \) and \( q < \frac{1}{2} \).

Then, based on the comparison amongst \( p \) and \( \frac{1/2 - q}{1 - q} \), we can decide which one is the maximum, i.e., if \( p > \frac{1/2 - q}{1 - q} \), trader will receive the maximum payoff using the second
offer schedule; or if \( p < \frac{1/2 - q}{1 - q} \), then trader will get the maximum payoff via the fourth offer schedule.

This experiment is considered to be the simplest case of this type of network since it only consists of one seller and one buyer. Therefore, the trader just needs to focus on these two parties and convince them to accept his offers in order to obtain the highest payoff under asymmetric information.

CONCLUSION

We analyze the buyer-seller networks with the traders serving as middlemen under asymmetric information. We investigate a single trader case with one seller and one buyer. The important result in this setting is that the trader chooses to offer different prices to sellers or to buyers when the penalty is almost zero in order to maximize profit. We also prove that the trader receives no payoff when he offers the same price of \( \frac{1}{2} \) to each agent as we expected.

In a latter experiment, we plan to deal with two sellers and two buyers with identical valuations of the good. His goal remains the same as in the current investigation. In addition, we deal with both sets of combinations agreed on his bidding or asking price offers to gain the maximum profits from that trade under asymmetric information.

Another possible extension is the examination of the so-called asymmetric cases, where we have one seller-two buyers or two sellers-one buyer, under the same assumption of asymmetric information. We expect that the outcome will be different at one level in terms of the profit maximization of the middleman. This difference is supposed to be characterized by offering the same ask or bid prices to the other agents in the network and whether these cases might yield a maximum payoff to the trader.

It is also possible and interesting to add at least one more trader into the setup. The competition between the traders would be the new focus area. Moreover, we can consider the problem which focuses on the case how traders can form or sever links with buyers and sellers when both buyers and sellers agree to have links to the traders. Therefore, an analysis on the endogenous networks would be another further step to improve our model of trade networks under asymmetric information.

We can also extend the analysis of our network model to the sequential or the descending/rising-bid auction case as discussed by Kranton and Minehart (2001). When the trader makes the bid and ask offers sequentially, first, the seller accepts or rejects the bid offer. Second, there will be a good for trade between the trader and the buyer if the seller accepts to sell the good he owns. Otherwise, the trader will not be able to offer a product to sell to the buyer. This situation, then, is expected to avoid the penalty case since the trader would be able to know if he would be willing to offer to the buyer or not in advance after the sequential trade between the seller and himself.

ACKNOWLEDGMENTS

I would like to thank Alison Watts for her insightful comments and helpful feedbacks during the preparation of this study. I also would like to thank Sajal Lahiri, Chi-feng Dai, Basharat Pitafi, Henry Hexmoor and Daniel Primont for their helpful
comments. I would like to thank Animesh Ghoshal for his comments at the Illinois Economics Association Conference.

ENDNOTES
1. This paper was mostly completed during my study at Southern Illinois University-Carbondale.
2. In our model, trader chooses the prices. Hence, he has a more strategic position than the other agents in the network and this idea is borrowed from Blume et al. (2007).

REFERENCES


ERRATUM

Johnnie B. Linn III,
Concord University

In Linn (2007), Eq. (25) is improperly set out as

\[ BGEN = \beta = \frac{EC}{UC} + \frac{XC}{EC + XC} \]  \hspace{1cm} (25)

while it should be expressed as

\[ BGEN = \beta = \frac{EC}{UC} + \frac{EC}{EC + XC} \]  \hspace{1cm} (25)

Table 1 shows the original and corrected quantities for BGEN. The corrected statistic has a mean of 2.96 with a standard deviation of 0.19. It meets the one-tailed t-test with two degrees of freedom at the 95% confidence level, rejecting the null hypothesis that \( \beta \) is less than unity.

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>SOURCE</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>DESIGNATOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorrect Statistic</td>
<td></td>
<td>3.71</td>
<td>4.06</td>
<td>4.03</td>
<td>BGEN</td>
</tr>
<tr>
<td>for ( \beta )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Statistic</td>
<td></td>
<td>2.74</td>
<td>3.08</td>
<td>3.05</td>
<td>BGEN</td>
</tr>
<tr>
<td>for ( \beta )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

REFERENCES
