Regional Shadow Economy: Friend or Foe?

Authored by John T. Sulik

The University of Akron

May 8, 2015

Abstract

The purpose of this study is to look into the effects of the shadow economy on economic growth at the regional level of development. At the time this study was conducted, shadow economy research had been primarily aimed at looking at cross-national effects. To date, there has only been one other study that has looked into the shadow economy at the regional level, however only estimations of the variable were conducted; no regression analysis has been completed. A Dummy Variable Least Squares regression and the Arellano Bond Two-Step Estimation are used to look at the shadow economy, as well as other common indicators of development, and their effect on economic growth while controlling for geographic location of the state as well as time. The results of the regression show that all variables, including the shadow economy size, increase economic growth at a significant level of less than one percent, supporting the hypothesis of this study.

Table of Contents

1. Introduction

One of the hardest topics for economists to study is the shadow economy. The shadow economy includes all economic activity in a nation that is unreported, therefore it does not add to the official economy. Sometimes the shadow economy is referred to as the underground economy or the informal sector of an economy. The shadow economy can be comprised of anything from unreported labor (paying a babysitter with cash) to illegal activities, such as drug trade or tax evasion (Schneider and Enste, 2013). Schneider and Enste (2000) define the shadow economy as:

"…unreported income from the production of legal goods and services, either from monetary or barter transactions, hence all economic activities that would generally be taxable were they reported to the tax authorities."

Schneider and Enste (2000, pg. 79).

This topic is important to economics because the shadow economy encompasses a great variety of fields. Firms in any market may operate under the radar of government regulation and several shadow economy activities are illegal and contribute to crime. Without official report of these activities, potential tax revenue for the government is lost and the effect of most tax policies is underrated or overrated from lack of information (Schneider and Enste, 2000).

Because the activity in this area is unreported, economists must use other methods of measurement to estimate the size of the shadow economy. Given all of these restrictions, finding the effects of the shadow economy on other factors of the economy is rather difficult. However, new methods have arisen in recent years that are considered to be much more accurate in measuring the shadow economy than before, many of which are attributed to Freidreich Schneider (2004). The Dynamic Multiple Indicators Multiple Causes (DYMIMIC) approach is revered as the most accurate way to measure the shadow economy due to its inclusion of every other estimation method (more on the DYMIMIC estimation in the following section). Several studies have used this approach to observe the shadow economy's effect on economic development.

This study poses the question and aims to answer, is there a significant effect on the official economy by the shadow economy at the state level of measurement and, if so, what direction does this effect follow? At the present, there is only one study by Wiseman (2014) that has observed state level shadow economies, however the study only looked into correlations (i.e. government size, taxes, market freedom, etc.) with the measured shadow labor economies, not any actual effects of the shadow economies. The literature is divided among the effects that the shadow economy has on economic development. Some economists, such as Soldatos (1996) and Ginsburgh (1985), suggest that increased shadow economic activity can be beneficial and some, such as La Porta and Schleifer (2014) or Eilat and Zinnes (2000), say that it can be hazardous to a nation. This study will examine the effect of increasing shadow economy sizes on the regional growth within the 50 United States, building heavily on the prior literature for the shadow economy.

2. Lit Review

A. Methods of Shadow Economy Estimation

A detailed review of the common methods of measurement is necessary to understand how economists derive shadow economy variables and use them effectively in modeling relationships.

The electricity consumption method uses overall electricity consumption within a geographic area and compares it proportionately to official GDP for that area. The idea is that both the shadow economy and official economy must use electricity to operate. Thus, electricity use outside what is necessary for the reported GDP is accounted for by the shadow economy. The labor force method is based on the assumption that a negative change in the labor force participation rate is related to an increase in shadow economic activity. The explanation for this is that as job availability in the official economy declines and unemployment rises, individuals will start losing hope of finding official employment and look to unofficial employment in the shadow economy for a means of income (Wiseman, 2013). Another method of estimating the shadow economy is through the monetary indicator method. This method involves estimating income in the shadow economy by looking at money spent on consumption and available money supply. Any discrepancies are attributed to the shadow economy (Soladots, 1996).

The DYMIMIC approach to measuring the shadow economy takes the changes over time of several different indicators and the changes over time of several different causes of the shadow economy and puts them together into a single variable. To do this, they create a model using the chosen indicators and causes to compute the shadow economy size for each individual observation at each individual time period. Most DYMIMIC approaches tend to use electricity consumption and labor force participation rates as two of the indicators for the measurement. Other common variables used in a DYMIMIC approach include direct and indirect taxation, governmental regulation burdens, monetary indicators (such as overall money supply not attributed to Federal Reserve activity) and displacement of production factors (decline in official energy use of other sectors, decrease in real income, etc.) (Schneider, 2004).

B. Influences of the Shadow Economy

After solving the issue of measuring the shadow economy, economists face the issue of actually measuring its effect on the official economy. A study by Soldatos (1996) looked into the shadow economy’s influence on official market operations. Soldatos used a monetary indicator model to measure the shadow economy and estimated the effect the shadow economy has on official money markets. The study used a growth model (more specifically two: Solow’s Growth Model and Tobin’s Model of Growth) to estimate if any growth in the shadow economy resulted in a change in the GDP of the official economy. The results indicate that an increase in the shadow economy, as a result of a decline in the money supply of the official economy, actually led to increased output in the official economy (Soladots, 1996). Soldatos explains this phenomenon in that people take their money from the official economy and invest it in potentially more lucrative businesses outside of the official economy. The investment promotes progress of the whole economy, but will not be measured officially. Conversely, La Porta and Schleifer (2014) found that a decline in the size of a shadow economy results in faster economic growth. They used the percentage of self-employed workers as a proxy for the size of the shadow economy and found that as percent change in self-employment declined (suggesting a decline in the shadow economy labor market because people moved from the self-employment of the shadow economy to official employment with companies), that per capita GDP increased rapidly. This can be attributed to increased wages in the official economy or due to increased job availability in the official economy, thus attracting more workers from the shadow economy (La Porta and Schleifer, 2014). This result contradicts the finding by Soladots and may be due to differing samples, measures, etc. and demonstrate the differences of effects found in prior studies. Several other studies show how the relationship is even more complicated.

The shadow economy is known to have a reverse-causal relationship with economic development based on other prior literature. That is, studies have found that the shadow economy influences economic development and economic development also influences the shadow economy. Eilat and Zinnes (2000) observed the shadow economy's effect on countries in the transitional phase of development. They used a modified version of the electricity consumption method model to measure the size of shadow economies and use a dynamic panel regression to find the effects of the shadow economy on economic development. Their findings demonstrate that a fall in economic development due to recessionary periods leads to an increase in the level of shadow economic activity. However, additional estimation resulted in the finding that a general rise in the shadow economy always related to a decline in economic activity. Eilat and Zinnes (2000) attribute the shadow economy's influence on the official economy as a result of underestimated macro policies and insufficient tax revenues. The official economy, on the other hand, can affect the shadow economy through increased job availability and wage targeting policies (Eilat and Zinnes, 2000). Schneider and Enste (2000) confirmed these findings, although they used a collection of countries at all levels of development, not simply transition countries. To further complicate this relationship, Schneider (2004) ran a study to estimate the two-way causality effects of the shadow economy on economic development for different levels of development of countries. He runs the estimations separately for OECD countries and for a set of low income countries according to World Bank classification. Using a DYMIMIC modeling approach to measure the shadow economy and a dynamic panel model to run his regressions, Schneider found that the relationship between the two variables was opposite for the two groups of countries. Shadow economy and economic growth were negatively correlated for developing nations, which he reasoned by a greater number of unemployed persons and greater amounts of corruption in developing nations that leads to an influx of employment in the shadow economy. The variables were positively correlated for developed nations, which Schneider (2004) attributes as a shadow economy consisting primarily on spillover labor and illegal pleasurable consumption such as drugs that both add to the whole economy rather than take away from it. Based on the literature, the relationship between the shadow economy and the official economy is clearly complex and intricate.

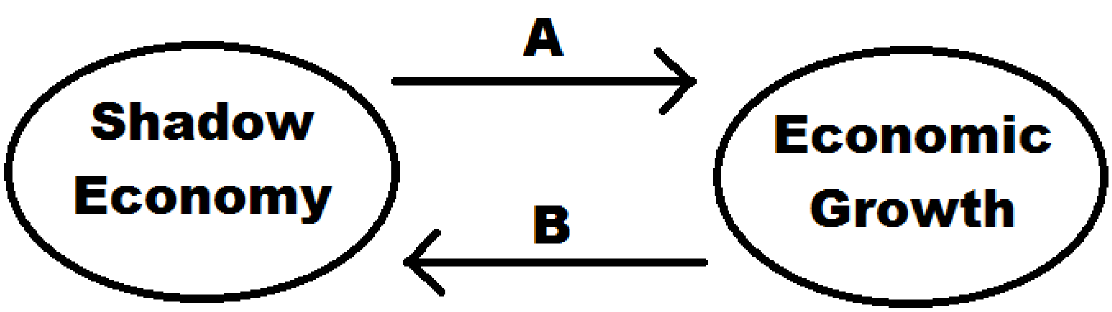
The studies reviewed to this point have solely estimated the effects of the shadow economy for varying nations. Recently, Travis Wiseman (2013) has looked into the correlations between the shadow economy and other economic variables at the regional level. He used his own version of the DYMIMIC model (included the electricity consumption, labor force participation and taxation variables) to measure the shadow economies of the 50 United States and then used correlation calculations to see if any relationship may exist between the estimated shadow economies and other common economic variables. His results demonstrated that all of the shadow economies were significantly positively correlated to tax and social welfare burdens. This suggests that as the governments increase taxes or collections of social programs such as social security, the shadow economy will grow to maintain current income levels by avoiding the new costs.

This study aims to build upon the research conducted by Wiseman (2013) by taking the estimated shadow economies and finding any effects they may have on the economic development of the individual states. Because of the relatively new venture into state level shadow economies, this will be the first study to observe any effects from this level of shadow economy to my knowledge. The hypothesis of this study is drawn from that of Schneider (2004) for developed nations and states that if the size of the shadow economy of a given state increases, then the economic growth of that state will increase as well. Because the state shadow economy measurements make up the United States (one of the most developed nations in the world), it is presumed that the relationship found by Schneider (2004) for developed nations will hold true.

That is, if the whole developed nation has a positive relationship between its shadow economy and economic growth, then the individual parts of that nation should have this relationship as well.

3. Theoretical Overview

Schneider and Enste (2000) have combined several prominent theories on the shadow economy to create a single model. In general, most of the theories they took into account had implemented the shadow economy into pre-existing macroeconomic models and observed its effect inside these models. Almost all prior studies had found a significant relationship between the shadow economy and economic growth, but the direction of this influence was contradictory. The following diagram illustrates the relationships found in the literature:



The letter A in the diagram represents the shadow economy's effect on economic growth and letter B represents the effect of economic growth on shadow economy. The study by Markus Adam and Victor Ginsburgh (1985) found a significant positive form of relationship A when studying the shadow economy of Belgium. Another study by Loayza (1996) found that relationship A actually has a negative direction, rather than positive, although the relationship is still significant. In Schneider's independent study (1998), he discovered that there is a positive significant relationship B rather than a relationship A. Because of these varying results, Schneider and Enste (2000) concluded that a significant relationship exists between the shadow economy and economic growth, but the direction of this relationship is still difficult to determine thus creating issues of two-way causality in several econometric analyses.

Methodology

For this study, the regional Solow model used by Holtz-Eakin (1992) is combined with a fixed effects model to create a basis to measure development. The model from this study is as follows:

$$GSP_{it} = \alpha_0 + \beta_1 GSP_{it-1} \alpha_1 educ_{it} + \alpha_2 inv_{it} + \alpha_3 LFPR_{it} + \varepsilon_{it}$$

GSP stands for Gross State Product per capita, which is a measure of state productivity. This measure is no longer collected but has been replaced with a measure of Gross Domestic Product per capita at the state level. Both measures are collected by the Bureau of Economic Analysis and the BEA has stated that the measures are almost identical. The *GSP* of t-1 indicator variable is a lag of the GSP per capita measure. The *educ* variable is a measure of educational attainment (taken as a proportion of the population of people age 25 or older that had a bachelor's degree or higher) that stands as a proxy for human capital. The *inv* variable is a measure of capital investment, which was computed using an advanced statistical equation of Holtz-Eakin's design (1991). The *LFPR* variable is the labor force participation and the ε is the error term. The *i* and *t* subscripts stand for the individual states and the given years, respectively.

When forming the model for this study, the investment variable used by Holtz-Eakin (1992) was omitted due to data availability. There are currently no available measures of capital investment at the state level. For this reason, Holtz-Eakin used an advanced statistical computation to find capital investment and the computation was beyond my current level of ability. However, each of the other variables had readily available data and the same measurements were used to collect them as were for the Holtz-Eakin (1992) study. Initially, this study's model used the shadow economy estimates computed by Wiseman (2013) to compute a static panel econometric model via a Dummy Variable Least Squares (DVLS) regression for regional development. The model is constructed as follows:

$$\log GDP_{it} = \alpha_0 + \beta_1 \log GDP_{it-1} + \alpha_1 \log Shdw_{it} + \alpha_2 \log LFPR_{it} + \alpha_3 \log Educ_{it} + \lambda_i + \gamma_t + \varepsilon_{it}$$

For this model, *GDP* is the gross domestic product per capita used as a measure of development. The *GDP* indicator variable is a lag of the dependent GDP by one year. The *Shdw* variable represents the estimations found by the Wiseman (2013) study for each state taken at a per capita level. The *LFPR* variable and *Educ* variable represent the labor force participation rate and education variables respectively, both suggested by Holtz-Eakin (1992). The variable λ is a collection of dummy variables representing each state to control for location. The variable γ is a collection of dummy variables representing each year in the data set to control for time. All variables were taken for each state, i, and for each year, t. Additionally, the indicator variables (other than the dummy control variables) were taken by the logs to compute effects as a percent change for all variables.

After computation of the DVLS regression, a dynamic panel model was conducted via an Arellano Bond Two-Step Estimation (AB) analysis to observe if statistical significance and strength of the model could be improved. The reason for inclusion of the AB estimation is twofold: it corrects for collinearity with the error term and it controls for the two-way causality issues explained earlier. Using the lagged variable as suggested by literature tends to create a correlation between the lagged variable and the error; AB estimation corrects this by measuring differences in the lags of variables, thus removing collinearity inside the error. Additionally, the

AB estimation uses stringent fixed-effects estimations over several lags that remove possibility of outside bias, thus reducing any two-way causality effects. The formation of this model was equivalent to the DVLS regression, however the AB estimation does not display an intercept because it is captured by the parameters of the vectored dummy variables. After computation was completed for the AB estimation, the results of the two models were compared and will be discussed further in the Results section of this study.

Data

The data for this study is taken from a variety of sources and is collected for each of the fifty United States for the years 1997-2008. This time range was chosen to allow the utilization of the shadow economy measurements estimated by Wiseman (2013), which were restricted to the range of 1997-2008. Wiseman noted that the shadow economies of the states decreases over the given time frame, but no other significant outliers exist due to time. The descriptive statistics of variables used can be found in Table 1 with definition and source. The shadow economy variable is the variable of interest for this study and is recycled from the Wiseman (2013) study. The estimation used to find the shadow economy variable is a DYMIMIC model that used employment, electricity consumption and tax revenues as indicator/causal factors. There are two reasons for using the same results to measure the shadow economies: it is the only DYMIMIC model to date used for measuring state level shadow economies and the results of the study are relatively recent and thus can be applied to current economic conditions (last measure taken at 2008).

Other variables to note include the GDP variable, the labor force participation rate and the education variable. The GDP variable is taken as the real GDP per capita in chained U.S. dollars by state. This variable is used as the dependent variable in this study to measure economic development for each state and is also taken with a lag of one year as an estimator of current GDP per capita. The labor force participation rate is calculated using the standard economic formula and was computed by the Bureau of Labor Statistics. The final variable, education, is used as a measure of human capital. According to Holtz-Eakin (1992), an indicator for human capital is necessary when measuring regional economic development and the one used by his study is the same for this study: the percentage of the population age 25 or older that has a four-year bachelor's degree or higher.

TABLE 1: Variable Descriptions with Descriptive Statistics

Variable Name	Description	Source	Mean	Standard Deviation	Min	Max
Gross Domestic Product per capita (*gdpcap*)	Gross Domestic Product by state divided by the state total population, in chained US $	Bureau of Economic Analysis	44063.09	8173.86	28449.00	69965.00
Shadow Economy (*shadow*)	DYMIMIC estimation of the shadow economy by state divided by state total population	Wiseman (2013) study on Shadow Economy	3609.40	528.64	2581.28	5259.54
Labor Force Participation Rate (*LFPR*)	Percentage of people active in the labor force proportionate to the amount of people able to work	Current Population Survey, Bureau of Labor Statistics	64.25	4.03	51.60	73.00
Human Capital, Education(*educ*)	Percentage of persons over 25 years of age that have a bachelor's degree or higher	Current Population Survey, Bureau of Labor Statistics	25.91	4.89	14.60	40.40

Dummy variables for state and year were also created to control for geographic differences and time differences in the DVLS regression. The state of Ohio was selected as a reference variable for the state dummy variable and 2008 was used as the reference variable for the time dummy variable. The year 2008 was used as a reference because it is the most recent available data on the shadow economy.

Results

After running the DVLS regression, it is clear that the shadow economy has a significant positive effect on economic growth of a state. The DVLS returned over a 99 percent R-squared value, suggesting that almost all variance in the dependent variable is accounted for. The DVLS parameter estimates can be seen in Table 2. The dummy variable estimates are not included in this table, but the complete DVLS with dummy variables can be found in Appendix A. The DVLS regression has several implications on the economic growth of a state. The variable of interest, the shadow economy, shows that a one percent increase in the size of the shadow economy causes a 0.63 percent increase in GDP per capita, the largest increase of all parameters. This effect is positive, supporting the argument that activity in the shadow economy actually helps the economy it is operating within to grow. This can be explained due to the level of development of the 50 United States (which is incredibly high due to the nature of the whole United States economy). As stated earlier, it is likely that the shadow economy of the states is made up of spillover labor that the official economy cannot support and of luxury consumption such as recreational drugs. Although these activities take away from tax revenue and contribute to crime, the transactions made within these industries actually support economic growth and thus provide a positive relationship between the two economies. The lagged GDP per capita variable indicated that a one percent increase in GDP per capita from the prior year led to a 0.4 percent increase in GDP per capita in the current year. The regression shows that a one percent increase in the amount of people over 25 years or older who have a bachelor's degree or higher decreases the GDP per capita by 0.02 percent. This is counteractive to development theory and is also the only variable not significant at the 95 percent level. A one percent increase in the labor force participation rate increases GDP per capita by 0.18 percent. This finding is actually in agreement with the theory of regional development and is statistically significant.

The Arellano Bond Two-Step estimation yielded similar results to the DVLS regression, but held much more significance. The shadow parameter for this estimation yielded that a one percent increase in the size of the shadow economy caused a 0.15 percent increase in GDP per capita. This parameter was still positive, but had much less of an effect than what was found by the DVLS regression, which may be more accurate to the theory because the positive relationship is attributed to spillovers that should have a small influence on economic growth. A one percent increase in GDP per capita from the prior year resulted in a 0.92 percent increase in GDP per capita for the current year. This parameter was much higher than that found in the DVLS, however it is more logical due to the fact that any increase in prior GDP per capita will be retained for the following year's GDP per capita. A one percent increase in the proportion of

TABLE 2: Regression Parameters

Dependent Variable: GDP per capita, by state 2005 chained US dollars

Variable	DVLS	AB Two-Step
Intercept	0.75***	
	(3.96)	
Lagged GDP	0.40***	0.92***
	(15.67)	(479.48)
Shadow	0.63***	0.15***
	(25.07)	(94.76)
Education	-0.02*	0.07***
	(-1.94)	(36.47)
LFPR	0.18***	0.26***
	(4.18)	(34.64)
R-Squared	0.996	N/A
RMSE	0.012	0.346
F-Value	(<0.0001)	(<0.0001)

Note: Values given are coefficients of the respective variables and the numbers in parentheses represent the t-value of the estimated parameters.
***,** and * represent significance at 99%, 95% and 90% respectively.

25 year olds or older that had a bachelor's degree or higher actually caused a 0.07 percent increase in GDP per capita. Using the AB estimation method, the education parameter became positive and significant and now conforms to the theory of regional development, unlike what was found by the DVLS regression. Finally, a one percent increase in the labor force participation rate created a 0.26 percent increase in GDP per capita. When using the AB estimation method, the influence of the labor force participation rate doubled from when using the DVLS regression.

Based on the results found, the Arellano Bond Two-Step Estimation method was more effective at capturing the true effects of the shadow economy on the economic growth of the United States. The model had parameters that were all significant at the 99 percent confidence level and the model overall had an extremely high significance level. Additionally, the education variable corrected itself to conform to the theory with the AB estimation method and the lagged GDP per capita variable had a more reasonable influence on current GDP per capita.

Conclusion

Based on the results of the regressions, it is evident that the shadow economy of the 50 United States has a significant positive effect on regional economic growth as suggested by Schneider (2004). This finding supported the hypothesis of the paper and suggests that at the regional level of development the shadow economy and economic growth hold the same relationship as the national levels. In future progressions of this study, a calculation of capital investment for the states can be used for a more powerful model and to see if the shadow economy effects are absorbed by the capital investment of both industries (both official and unofficial economies must invest in capital). Additionally, the Sargan Test for the Arellano-Bond estimation was over 2,000, suggesting the instruments may not be significant and explains the incredibly high t-values (as seen in Table 2). However, the Sargan Test was created and used for the Arellano-Bover (1995) systems test and outdates the Arellano-Bond estimator used in this study. The variables in this estimation may be overestimated, but the significance found should still be taken into consideration.

This study leads to the idea that public policies aimed at diminishing the shadow economy may be counter-productive to their goals of bolstering the whole economy because the extra income earned in the shadow economy can be redistributed to boost the whole economy. This effect, however, is dwarfed in comparison to the positive effect that would be ascertained from the lost tax revenues to the government. The shadow economy influences only a small fraction of the whole economy whereas the official economy influences almost all of the whole economy. Therefore, the positive relationship found in this study should not be concluded to show the shadow economy has a beneficial effect, but rather that there is simply a positive relationship between the shadow economy and economic growth.

Appendix A:

Full Parameter Estimates of DVLS Regression including Dummy Variables

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	0.74491	0.18806	3.96	<.0001
logshadow	0.62701	0.02501	25.07	<.0001
loggdpcap_1	0.39593	0.02527	15.67	<.0001
logeduc	-0.01989	0.01027	-1.94	0.0533
loglfpr	0.18482	0.04423	4.18	<.0001
alabama	-0.06811	0.00715	-9.53	<.0001
alaska	-0.02734	0.0056	-4.88	<.0001
arizona	-0.01812	0.00549	-3.3	0.001
arkansas	-0.06692	0.00759	-8.82	<.0001
california	0.02493	0.00479	5.21	<.0001
colorado	0.0167	0.00665	2.51	0.0123
delaware	0.03182	0.00576	5.52	<.0001
florida	-0.04409	0.00594	-7.43	<.0001
georgia	-0.00878	0.00508	-1.73	0.0844
hawaii	-0.09144	0.00582	-15.7	<.0001
idaho	-0.03826	0.00726	-5.27	<.0001
illinois	-0.00173	0.00462	-0.38	0.7075
indiana	-0.0462	0.00577	-8.01	<.0001
iowa	-0.04963	0.00684	-7.25	<.0001
kansas	-0.04127	0.00652	-6.33	<.0001
kentucky	-0.04525	0.00657	-6.89	<.0001
louisiana	-0.03435	0.0067	-5.13	<.0001
maine	-0.07303	0.00647	-11.29	<.0001
maryland	-0.05111	0.00594	-8.6	<.0001
michigan	-0.04706	0.0053	-8.88	<.0001
massachusetts	-0.00234	0.00499	-0.47	0.6387
minnesota	-0.01895	0.00687	-2.76	0.006
mississippi	-0.1044	0.00884	-11.8	<.0001
missouri	-0.04284	0.00549	-7.81	<.0001
montana	-0.05072	0.00713	-7.12	<.0001
nebraska	-0.05422	0.00727	-7.45	<.0001
nevada	-0.02294	0.00545	-4.21	<.0001
hampshire	-0.00313	0.0067	-0.47	0.6403

jersey	0.01613	0.00484	3.33	0.0009
mexico	-0.08243	0.00649	-12.7	<.0001
york	0.00412	0.00592	0.7	0.4869
ncarolina	-0.00674	0.00505	-1.33	0.1832
ndakota	-0.08863	0.00802	-11.05	<.0001
oklahoma	-0.0327	0.00631	-5.18	<.0001
oregon	0.05145	0.006	8.58	<.0001
penn	-0.03189	0.00531	-6	<.0001
risland	-0.02325	0.00522	-4.45	<.0001
scarolina	-0.04482	0.0062	-7.22	<.0001
sdakota	-0.06021	0.00816	-7.38	<.0001
tennessee	-0.05224	0.00588	-8.88	<.0001
texas	-0.01033	0.00488	-2.12	0.0348
utah	-0.03223	0.00715	-4.51	<.0001
vermont	-0.07973	0.00792	-10.07	<.0001
virginia	-0.04188	0.00536	-7.82	<.0001
washington	-0.03211	0.00478	-6.71	<.0001
wvirgin	-0.07993	0.01077	-7.42	<.0001
wisconsin	-0.03099	0.00635	-4.88	<.0001
wyoming	-0.03369	0.00645	-5.22	<.0001
yr1997	0	.	.	.
yr1998	-0.13839	0.00717	-19.3	<.0001
yr1999	-0.1204	0.0068	-17.72	<.0001
yr2000	-0.1232	0.0066	-18.67	<.0001
yr2001	-0.13187	0.00605	-21.8	<.0001
yr2002	-0.12235	0.00614	-19.92	<.0001
yr2003	-0.10062	0.00569	-17.69	<.0001
yr2004	-0.08975	0.00548	-16.37	<.0001
yr2005	-0.07986	0.00496	-16.09	<.0001
yr2006	-0.02903	0.00333	-8.71	<.0001
yr2007	-0.00558	0.00274	-2.03	0.0426

Appendix B:

SAS Coding for Data Import, DVLS Regression and AB Two-Step Estimation

```
Proc Import
	datafile="E:\Senior_Proj\Data\shdw_SAS.xlsx"
	DBMS=xlsx
	out=Shdw.shadow;
	sheet=Sheet1;
	run;

Proc Sort data=Shdw.shadow;
	by state year;
	run;

Proc Import
	datafile="E:\Senior_Proj\Data\SASreadyLaborForce.xlsx"
	DBMS=xlsx
	out=Shdw.LFPR;
	sheet=Sheet1;
	run;

Proc Sort data=Shdw.LFPR;
	by state year;
	run;

Proc Import
	datafile="E:\Senior_Proj\Data\gdppercap.xls"
	DBMS=xls
	out=Shdw.gdp;
	sheet=Sheet0;
	run;

Proc Sort data=Shdw.gdp;
	by state year;
	run;

Proc Import
	datafile="E:\Senior_Proj\Data\realgdp_SAS.xlsx"
	DBMS=xlsx
	out=Shdw.gdpwhole;
	sheet=realgdp_SAS;
	run;

Proc Sort data=Shdw.gdpwhole;
	by state year;
```

```
    run;

Proc Import
    datafile="E:\Senior_Proj\Data\pop_SAS.xls"
    DBMS=xls
    out=Shdw.pop;
    sheet=Sheet1;
    run;

Proc Sort data=Shdw.pop;
    by state year;
    run;

Proc Import
    datafile="E:\Senior_Proj\Data\CPS-Educational Attainment\educ_SAS.xlsx"
    DBMS=xlsx
    out=Shdw.educ;
    sheet=Sheet1;
    run;

Proc Sort data=Shdw.educ;
    by state year;
    run;

    Data Shdw.complete0;
    merge Shdw.shadow Shdw.gdp Shdw.LFPR Shdw.educ Shdw.gdpwhole;
    by state;
    shadow=((shdw/100)*gdpcap);
    loggdpcap= log(gdpcap);
    logshadow= log(shadow);
    logeduc= log(educ);
    loglfpr= log(lfpr);
    IF state_code=1 then alabama=1; ELSE alabama=0;
    IF state_code=2 then alaska=1; ELSE alaska=0;
    IF state_code=3 then arizona=1; ELSE arizona=0;
    IF state_code=4 then arkansas=1; ELSE arkansas=0;
    IF state_code=5 then california=1; ELSE california=0;
    IF state_code=6 then colorado=1; ELSE colorado=0;
    If state_code=7 then connecticut=1; ELSE connecticut=0;
    IF state_code=8 then delaware=1; ELSE delaware=0;
    IF state_code=9 then florida=1; ELSE florida=0;
    IF state_code=10 then georgia=1; ELSE georgia=0;
    IF state_code=11 then hawaii=1; ELSE hawaii=0;
    IF state_code=12 then idaho=1; ELSE idaho=0;
    IF state_code=13 then illinois=1; ELSE illinois=0;
    IF state_code=14 then indiana=1; ELSE indiana=0;
```

```
IF state_code=15 then iowa=1; ELSE iowa=0;
IF state_code=16 then kansas=1; ELSE kansas=0;
IF state_code=17 then kentucky=1; ELSE kentucky=0;
IF state_code=18 then louisiana=1; ELSE louisiana=0;
IF state_code=19 then maine=1; ELSE maine=0;
IF state_code=20 then maryland=1; ELSE maryland=0;
IF state_code=21 then massachusetts=1; ELSE massachusetts=0;
IF state_code=22 then michigan=1; ELSE michigan=0;
IF state_code=23 then minnesota=1; ELSE minnesota=0;
IF state_code=24 then mississippi=1; ELSE mississippi=0;
IF state_code=25 then missouri=1; ELSE missouri=0;
IF state_code=26 then montana=1; ELSE montana=0;
IF state_code=27 then nebraska=1; ELSE nebraska=0;
IF state_code=28 then nevada=1; ELSE nevada=0;
IF state_code=29 then hampshire=1; ELSE hampshire=0;
IF state_code=30 then jersey=1; ELSE jersey=0;
IF state_code=31 then mexico=1; ELSE mexico=0;
IF state_code=32 then york=1; ELSE york=0;
IF state_code=33 then ncarolina=1; ELSE ncarolina=0;
IF state_code=34 then ndakota=1; ELSE ndakota=0;
IF state_code=35 then ohio=1; ELSE ohio=0;
IF state_code=36 then oklahoma=1; ELSE oklahoma=0;
IF state_code=37 then oregon=1; ELSE oregon=0;
IF state_code=38 then penn=1; ELSE penn=0;
IF state_code=39 then risland=1; ELSE risland=0;
IF state_code=40 then scarolina=1; ELSE scarolina=0;
IF state_code=41 then sdakota=1; ELSE sdakota=0;
IF state_code=42 then tennessee=1; ELSE tennessee=0;
IF state_code=43 then texas=1; ELSE texas=0;
IF state_code=44 then utah=1; ELSE utah=0;
IF state_code=45 then vermont=1; ELSE vermont=0;
IF state_code=46 then virginia=1; ELSE virginia=0;
IF state_code=47 then washington=1; ELSE washington=0;
IF state_code=48 then wvirgin=1; ELSE wvirgin=0;
IF state_code=49 then wisconsin=1; ELSE wisconsin=0;
IF state_code=50 then wyoming=1; ELSE wyoming=0;
IF year=1997 then yr1997=1; ELSE yr1997=0;
IF year=1998 then yr1998=1; ELSE yr1998=0;
IF year=1999 then yr1999=1; ELSE yr1999=0;
IF year=2000 then yr2000=1; ELSE yr2000=0;
IF year=2001 then yr2001=1; ELSE yr2001=0;
IF year=2002 then yr2002=1; ELSE yr2002=0;
IF year=2003 then yr2003=1; ELSE yr2003=0;
IF year=2004 then yr2004=1; ELSE yr2004=0;
IF year=2005 then yr2005=1; ELSE yr2005=0;
IF year=2006 then yr2006=1; ELSE yr2006=0;
```

```
	IF year=2007 then yr2007=1; ELSE yr2007=0;
	IF year=2008 then yr2008=1; ELSE yr2008=0;
	run;

Proc Panel data=Shdw.Complete0;
	id state year;
	lag loggdpcap (1) /out=shdw.blah;
	run;

Data Shdw.whole;
	set shdw.blah;
		if loggdpcap_1="." then delete;
		run;

Proc Means data=Shdw.whole;
	var logshadow loggdpcap_1 logeduc loglfpr;
	run;

Proc Corr data=Shdw.whole;
	var logshadow logeduc loglfpr loggdpcap_1;
	run;

Proc Reg data=Shdw.whole;
	model loggdpcap=logshadow loggdpcap_1 logeduc loglfpr alabama alaska arizona
arkansas california colorado delaware florida georgia hawaii idaho illinois indiana iowa kansas
kentucky louisiana maine maryland michigan massachusetts minnesota mississippi missouri
montana nebraska nevada hampshire jersey mexico york ncarolina ndakota oklahoma oregon
penn risland scarolina sdakota tennessee texas utah vermont virginia washington wvirgin
wisconsin wyoming yr1998 yr1999 yr2000 yr2001 yr2002 yr2003 yr2004 yr2005 yr2006
yr2007;
	run;

Proc Panel data=Shdw.whole;
	inst depvar exog=(logeduc loglfpr) ;
	model loggdpcap=logshadow loggdpcap_1 logeduc loglfpr /gmm nolevels twostep
maxband=5;
	id state year;
	run;
```

References

Eilat, Yair, and Clifford Zinnes. "The Shadow Economy in Transition Countries: Friend or Foe? A Policy Perspective." World Development 30, no. 7 (2002): 1233-254.

Holtz-Eakin, Douglas. "Solow and the States: Capital Accumulation, Productivity and Economic Growth." National Bureau of Economic Research, no. Working Papers #4144 (1992).

La Porta, Rafael, and Andrei Schleifer. "Informality and Development." Journal of Economic Perspectives 28, no. 3 (2014): 109-26.

Schneider, Friedrich. "Shadow Economies Around The World: What Do We Really Know?" European Journal of Political Economy, 2004, 598-642.

Schneider, Friedrich, and Dominik H Enste. "Shadow Economies: Size, Causes, and Consequences." Journal of Economic Literature 38, no. 1 (2000): 77-114. Accessed January 22, 2015.

Soldatos, Gerasimos T. "Growth and the Interaction Between Money and the Underground Economy." Revue économique, 1996, 73-84.

Wiseman, Travis. "US Shadow Economies: A State-level Study." Constitutional Political Economy, 2013, 310-35. Accessed January 22, 2015.