A DIFFERENT ANALYSIS OF RURAL DEVELOPMENT'S BUSINESS AND INDUSTRY LOAN GUARANTEE PROGRAM: THE IMPACT ON TAX REVENUE IN OKLAHOMA COMMUNITIES

By

TY ROPE SMITH

Bachelor of Science in Agricultural Economics

Oklahoma State University

Stillwater, Oklahoma

2018

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
May, 2020

A DIFFERENT ANALYSIS OF RURAL

DEVELOPMENT'S BUSINESS AND INDUSTRY

LOAN GUARANTEE PROGRAM: THE IMPACT

ON TAX REVENUE IN OKLAHOMA COMMUNITIES

Thesis Approved:

Dr. Brian E. Whitacre

Thesis Adviser

Dr. Notie H. Lansford

Dr. Mike D. Woods

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Brian Whitacre, and my committee members, Dr. Mike Woods and Dr. Notie Lansford, for all of the countless hours and effort they devoted to the process of this project. I would also like to thank them for all of the guidance and knowledge they gave me throughout my time Oklahoma State.

I would also like to thank my parents for all the love and support they gave me throughout graduate school. They were always willing to listen to me ramble on about whatever issue, big or small, I faced, and provide guidance when I needed it.

Name: TY ROPE SMITH

Date of Degree: MAY, 2020

Title of Study: A DIFFERENT ANALYSIS OF RURAL DEVELOPMENT'S BUSINESS AND INDUSTRY LOAN GUARANTEE PROGRAM: THE IMPACT ON TAX REVENUE IN OKLAHOMA COMMUNITIES

Major Field: AGRICULTURAL ECONOMICS

Abstract: Rural areas across the U.S. have struggled since the Great Recession, with limited employment growth and significant outmigration. Many public- and private-sector programs are focused on generating economic development in rural areas, but few are formally evaluated. One often-overlooked component required for effective rural development is the generation of local sales tax revenue, which helps fund city amenities and services. This research evaluates how one specific rural development public-sector program – the USDA's B&I Loan Guarantee program – impacts sales tax revenue for recipient Oklahoma communities. Sales tax revenue and census demographic data for all Oklahoma communities that charged a sales tax between 2005 and 2015 is meshed with information on B&I loan recipient communities during that time. Multivariate regression and coarsened exact matching (CEM) techniques are used to assess the impacts on sales tax revenues across all 2-digit Standard Industrial Classification (SIC) retail codes. Propensity score matching is also used for a robustness check of the CEM results. The results from the regression models depict mixed results of the impact that the B&I program has on total retail sales (TRS) and TRS per capita in Oklahoma. During time period one (2005-2010), the regression coefficient for the loan amount variable were positive and statistically significant across all models except one. Time period two (2010-2015) results show little to no significant impact of B&I loans on TRS and TRS per capita. The economic environment during these two periods were dramatically different, with the Great Recession occurring during time one and a major oil boom in time two. The results demonstrate that the B&I program and other similar programs may be more vital during tough economic times when tax revenue is crucial for rural communities.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

While most (97%) of the United States' land mass is defined as rural, only 19.3% of the total population resided there as of 2010 (US Census Bureau, 2017). This is significantly smaller than the 54.4% of the US population living in rural areas in 1910 (Ratcliffe et al., 2016). This trend of population decline in rural areas is a result of the growth of urban areas over the past century. Urban and rural areas also differ in levels of – and growth in – employment. Since the 2008-2009 recession the unemployment rate has declined in both geographic areas, but employment growth has been significantly slower in rural locations (Cromartie, 2018). The slower rate of recovery in many rural areas stems from a variety of sources, including a slowdown in traditionally rural-dominated industries such as manufacturing, and lower levels of amenities to attract new residents or businesses.

The need for more economic development in rural areas has been recognized by both the private and public sectors, with many programs devoted to promoting growth in rural economies (Johnson, 2009). However, there is a limited amount of literature evaluating these programs to assess their effectiveness. Most evaluations that do take place are performed by the agencies that administer the programs or by outside evaluators that may not conduct a thorough analysis. Bartik and Bingham (1997) discuss six reasons why rigorous program evaluations are not typically done, including that more complex studies are difficult to do and are usually more costly. These complex analyses not only increase budgetary costs but also require time from administrators and staff to design and implement the collection of data needed for the study. However, Bartik and Bingham (1997) argue that the main reason rigorous evaluations are not performed is that if the outcome is not favorable, the program administrators may face negative political consequences. This results in the need for program evaluations done by researchers without any ties to the agency, using a thorough and well-thought-out methodology as opposed to some of the "in-house" evaluations that often take place (for example, surveying program beneficiaries or summarizing program statistical data). Such surveys or simple descriptive statistics can be easier to manipulate, whereas a more robust analysis that considers the program's true impact on specific economic outcomes will be more useful to policymakers (Bartik and Bingham, 1997).

One of the public programs devoted to the economic development of rural America is the United States Department of Agriculture (USDA) Rural Development (RD) Rural Business-Cooperative Services (RBS). The purpose statement of RBS is to "provide loans, loan guarantees, grants, and payments designed to increase economic opportunity in rural America" (Office of Budget and Program Analysis, 2018). The RBS program with the largest amount of funding is the Business and Industry (B&I) Loan Guarantee program with nearly $1 billion worth of funds designated to it every fiscal year (FY). Table 1 shows the proposed budgets for the major programs of RBS since 2012, and demonstrates the dominance of the B&I program within this line of funding. In the first budget proposals by the president for FY 2018 and FY 2019, the funding for almost all RBS programs were eliminated (shown in Table 1). However, these proposals to cut the funding of RBS programs did not stand, and were later funded in the enacted budgets[1] – likely due to a backlash from both the public and private sectors. One example of the backlash is a letter prepared by the ranking Democratic member of the House Committee on Small Business (Velazquez, 2018). This letter explained how eliminating the B&I program would

[1] The enacted budget amounts are located in Appendix I

significantly cut the loan pool for rural businesses and their communities. This letter also argued that small businesses are the foundation for the U.S. economy, and that they need an environment that will allow them to continue to grow and succeed. Removing all funding from the program would make it more difficult for such businesses to operate in a stable environment and would present them with even more challenges than they already face. She further argued that the lower funding specifically for the B&I program would lead to lower economic growth, higher unemployment rates, and population loss in rural areas (Velazquez, 2018).

Table 1

USDA Rural Business-Cooperative Service Programs Proposed Budget Level of Funding (*in millions of dollars)*

Item	2012	2013	2014	2015	2016	2017	2018	2019	2020
Appropriate Tech. Transfer to Rural Areas	**-**	**-**	**$2**	**$2**	**$2**	**$2**	**-**	**-**	**-**
Bioenergy for Advanced Biofuels	**$105**	a/	a/	$15	$15	$15	$15	-	$7
Biorefinery Assistance Guaranteed Loans	**-**	a/	a/	$124	$254	-	-	-	-
Business and Industry Guaranteed Loans	**$823**	**$821**	**$741**	**$591**	**$758**	**$892**	**-**	**-**	**$1,000**
Intermediary Relending Program	**$36**	$19	$19	$10	$10	$19	-	-	-
Rural Business Development Grants	**$37**	$30	$0	$58	$30	$30	-	-	-
Rural Cooperative Development Grants	**$16**	$13	$55	$0	$6	$6	-	-	-
Rural Economic Development: Loans and Grants	**$43**	$43	$43	$69	$97	$97	-	-	-
Rural Energy for America (Sec. 9007) Loans and Grants	**$211**	$19	$52	$313	$485	$450	$423	$515	$370
Rural Microentrepreneur Assistance Program (sec. 6022): Loans and Grants	**$36**	$22	$22	$43	$42	$37	$10	-	-
Small Socially Disadvantaged Producer Grants	**-**	-	-	$3	$3	$3	-	-	-
Value-Added Producer Grants	**$20**	$15	$15	$11	$10	$11	-	-	-

a/ - subject to reauthorization

These are all proposed budget amounts and not the estimated or enacted amounts

Source: USDA 2012 -2020 budget summary proposals

However, the B&I loan program was reestablished in the proposed 2020 USDA budget, demonstrating that the program has political support again. In fact, the 2020 proposal was increased by $80 million from the FY 2019 enacted amount ($920 million) to $1 billion dollars. This level of funding is expected to assist 433 businesses and create or save roughly 11,000 jobs, which is estimated in the 2020 USDA Budget explanatory notes[2]. The details of the 2020 USDA budget proposal are interesting: It proposes rescinding the mandatory funds provided in the 2018 Farm Bill and the elimination of all other discretionary RBS programs (Table 1). Therefore, while the B&I loan guarantee program currently has significant support and funding, a rigorous program evaluation is needed to document the degree to which it is linked to economic development in rural America. This leaves the door open for future research to investigate the importance of the other RBS programs that were not funded in the 2020 USDA budget.

One important component of rural economic development is the fiscal situation of local governments (i.e. city/county). In the public finance literature, there is a section focusing on "tax adequacy" for state and local governments (Berney and Larson, 1968; Fox, 1986; Alm, 1996; Nelson et al., 2007). The concept of tax adequacy focuses on the governments' ability to provide a consistent level of public services for their constituents, even in economic downturns (Mikesell, 1984; Fox, 1986). Reducing the variability of revenue received is an important part of this idea. Tax revenue is one of the main income sources for governmental revenue, but it is especially vital for local governments. In 2016, roughly 40 percent of local revenue came from property, sales and other taxes (Tax Policy Center, n.d.). A large portion of the other local revenue comes from intergovernmental transfers, with a historical average of roughly 45 percent in rural areas (Felix and Henderson, 2010). However, during the most recent recession, revenue from intergovernmental transfers from state government shrunk, suggesting that local tax revenue is

[2] The expected numbers are estimated based on the amount of funding levels, how many businesses were assisted, and from the number of jobs created in previous years.

even more vital for rural America (Felix and Henderson, 2010). If intergovernmental funding is cut by the state government, there is a need to offset the cuts to prevent budget shortfalls in local governments. Property taxes in nonmetropolitan counties remained relatively strong during the economic downturn because of the stable nature of rural real estate, but there is still a need to find more avenues of revenue to support the local government (Felix and Henderson, 2010).

Several studies from the finance literature have examined income versus sales tax revenue in terms of tax adequacy, and their resulting impacts on economic growth. Holcombe and Sobel (1995) and Sobel and Wagner (2003) examined the cyclical variability of the state income and sales taxes using data from two distinct periods. State revenue data from 1949 to 1989 was used in Holcombe and Sobel (1995), while Sobel and Wagner (2003) used 2001 state revenue data. Both articles found that sales tax revenues are more consistent over time compared to income tax, and are therefore very important to the state's tax adequacy. Helms (1985) examined how state and local tax rate increases affected economic growth when the revenue is used in different ways. His state-level results from pooled time series and cross-sectional data showed that funding transfer payments significantly *slowed* economic growth. However, if the revenue is used to fund public services, the enhanced services may more than counterbalance the disincentive effects of the associated taxes (Helms, 1985). Thus, increasing local tax revenue (and using it appropriately) is an important component of economic growth.

Even though increasing tax revenue is not a listed goal or objective for any specific USDA RD program, a different, but still important avenue of research is to examine whether the B&I program expands total retail sales, which helps support local governments[3]. The Holcombe and Sobel (1995) and Sobel and Wagner (2003) articles demonstrate the importance of sales taxes

---

[3] The goal of the B&I program is to promote the creation of rural businesses to secure start-up capital, finance business expansion, and create jobs. B&I also improves the quality of life for residents, and it helps economic development in rural areas. These goals support the USDA strategic goal 4. (Office of Budget and Program Analysis, 2019)

for reducing the variability in *state* budgets. However, little analysis has focused on the impacts of specific economic developments programs to local tax coffers.

This research will evaluate the B&I Loan Guarantee program by using multivariate regression to examine tax revenue changes in Oklahoma communities after receiving a B&I loan guarantee. The dataset is pruned using coarsened exact matching to ensure appropriate comparisons between treated and control groups. A logical hypothesis is that an increase in tax revenue will be observed, but to my knowledge no formal evaluation of this type has taken place on this or other similar programs. This research aims to address this void in the literature. As a robustness check on the results, propensity score matching is also used to match communities receiving a B&I loan guarantee to similar communities that did not.

CHAPTER II

REVIEW OF LITERATURE

Background Information

The B&I loan guarantee program is currently administered by RD's Business and Cooperative Programs. However, the B&I loan program was created before RD was even established. B&I was organized under the Farmers Home Administration, a former program of the USDA, in 1972 by the Consolidated Farm and Rural Development Act (Cowan, 2016). The Rural Development Agency was later conceived by the Agricultural Reorganization Act of 1994, which established the three current sectors of RD: Rural Business-Cooperative Service (RBS), Rural Housing Service (RHS), and Rural-Utilities Services (RUS) (Cowan, 2016).

In the early years of RBS, direct loan funding was available, but since fiscal year 2002 there has been no funding for direct loans. This has resulted in the B&I program consisting of loan guarantees only (Cowan, 2016). Since there are no more direct loans available, only lenders with sufficient experience, the legal lending authority, and financial strength to operate a successful loan program can apply for loan guarantees. These lenders include federal or state-charted banks, savings and loans, farm credit banks, and credit unions (Rural Development, 2017). The amount of guarantee on the loan depends on the size of the loan. Loans of $5 million or less are guaranteed up to 80%, a 70% loan guarantee is offered for loans between $5 million and $10 million, and loans exceeding $10 million to $25 million receive a maximum of a 60% guarantee (Rural Development, 2017). The limit for most loan guarantees is $10 million but the administrator of RBS can grant loan guarantees up to $25 million at his or her discretion (Office of the Comptroller of the Currency, 2018). The typical loan guarantee size was roughly $3.5 million for FY 17 for the US and was roughly around $3.7 million for Oklahoma (Oehler and Lewis, 2017; Wiles[4], 2019). Some of the industries receiving loan guarantees in Oklahoma are manufacturing, construction, oil and gas companies, convenience stores, restaurants, and medical services, which are all high-risk industries but are important to improve economic development[5].

Proponents of the B&I program argue it is vital because locally owned banks are struggling in rural communities, and many do not have abundant resources to fund significant investments in economic development projects. A recent analysis shows that rural locations comprise only 10% of all small loans to businesses, and that over 30% of rural counties do not have a local bank as of 2017 (up from only 13% in 1995) (Ensign and Jones, 2019).

Lenders use the B&I program to increase lending capacity, because most rural banks do not have ample amounts of capital to loan large amounts of money to one business. There is also significant risk involved with loaning large amounts of money to just one borrower because of the default rates of small businesses. The rate of small businesses that survived their first year from 2005 to 2017 was roughly 79%, but this number falls off to roughly 50% for survival to five years or longer, and only about one-third of businesses survive 10 years or longer (Office of Advocacy, 2018). A report from Voight and Campbell (2017) found that roughly one of six loans from the Small Business Association (SBA) defaulted between 2006 and 2015. An analysis by Tetreault (2019) found a similar percentage for the average default rate of SBA loans in Oklahoma from 2010 to 2018 (13.3%). The B&I program aims to reduce the risk for a bank lending to business in rural areas. The delinquency rate for B&I was roughly three to 10 percent from 2000 to 2017, notably lower than those from SBA loans (Johnson, 2009; Oehler and Lewis, 2017). The B&I

[4] Brian Wiles is the Program Director for RBS in the state of Oklahoma.
[5] These examples of industries come from the dataset provided by the Oklahoma RD state office.

program also aims to reduce the failure rates of businesses as well. Rupasingha, Crown, and Pender's (2019) analysis found that 0.30 of 1000 businesses are predicted to fail two years after receiving a B&I loan, while non-recipient businesses with similar characteristics have a predicted failure rate of approximately 3 per 1000 businesses.

One important caveat when discussing rural development efforts is how "rural" is defined. RD's definition for "rural areas" is somewhat different from the definitions used by other governmental offices. RD defines the eligible areas for the B&I program as any *town or city* with a population of 50,000 or less (Rural Development, 2017). [6] This differs from the U.S. Census Bureau's traditional definition of "rural" as a city with population less than 2,500. As the USDA Economic Research Service explains, there are multiple ways to define a rural community (Ratcliffe et. al, 2016). The Office of Management and Budget (OMB) defines metropolitan areas as central *counties* with one or more urbanized areas densely-settled urban entities with 50,000 or more people, as compared to the city level definition of the census and RD. OMB also includes outlying counties economically tied to core counties that have 25% of workers living in the county commuting to the central counties, or where 25% of the employment comes from the central counties, as metropolitan. Nonmetro counties are located outside the boundaries of the metro areas (Cromartie and Parker, 2019). Government programs use alternative definitions of "rural" or "nonmetropolitan" to designate entities that are eligible for funding (Cromartie and Bucholtz, 2008).

Literature Review

There is a limited amount of literature focusing on evaluating the B&I loan guarantee program, but there have been several studies assessing other USDA and federal economic development

[6] RD also allow the borrower's headquarters may be in a larger city as long as the project location is located in an eligible area. The lender can be located anywhere, and some projects may be funded in rural and urban areas under the Local and Regional Food System Initiative (Rural Development, 2017).

efforts. One of the more recent studies of an RD program is Rupasingha, Pender, and Wiggins' (2018) evaluation of the Value-Added Producer Grant Program (VAPG). Rupasingha et al. (2018) used survival analysis to document that businesses receiving a VAPG between 2001 and 2013 were less likely to fail, and that the likelihood of the failure decreased as the amount of the VAPG increased. This same study also used multivariate regression analysis to see if the VAPG affects employment growth and found that businesses receiving a VAPG employed on average five more workers than nonrecipients in the same period. This increase in employment was substantial since a participating business on average only employed 14 workers at the time of the grant (Rupasingha et al., 2018). Another study, by Janeski and Whitacre (2009), examined the short- and long-term impact of RD's Water and Sewer Infrastructure Program. They used multivariate regression and average treatment effects on different economic growth measures between Oklahoma communities that received/did not receive funding. In their analysis, they found no evidence of any short-term (less than 10 years) impacts, but in the long term (10 to 20 years) median house values were notably higher (between five and 13 percentage points) in the recipient communities (Janeski and Whitacre, 2009). A third study, by Conley and Whitacre (2019), evaluated the Community Development Block Grant program (CDBG) using similar evaluation techniques and focused on similar economic growth measures as Janeski and Whitacre (2009). However, the evaluation found that rural Oklahoma communities receiving a CDBG did not have a higher level of economic growth over the short and long run compared to similar communities that did not receive funding (Conley and Whitacre 2019).

One study of a non-USDA program is Whitacre, Shideler, and Williams' (2016) evaluation of Oklahoma's Quality Jobs program, which used similar techniques as the previous studies mentioned. The results from their study revealed there was no evidence that Oklahoma communities participating in the Quality Jobs program had economic growth compared to Oklahoma communities not participating the in program (Whitacre, Shideler, and Williams

2016). However, when comparing Oklahoma communities participating in the program to Kansas communities (where no Quality Jobs program exists) there was statistical evidence of median household income growth (Whitacre, Shideler, and Williams 2016). Another non-USDA program evaluation study is Rupasingha and Wang's (2017) analysis of the Community Reinvestment Act (CRA) small business loans. The CRA seeks to address the limited amount of capital available to minority and low-income neighborhoods as well as to small businesses and farms (Rupasingha and Wang, 2017). Rupasingha and Wang (2017) found that at the county level, the lending practices of CRA loans have a positive effect on small business growth (in terms of number of establishments). They used data from a panel of 3050 counties over the period of 1996 – 2010[7].

Hansen and Kalambokidis (2010) assessed the impact of the Job Opportunity Building Zone (JOBZ), a tax-free-zone economic development program, in Minnesota. The main goals of JOBZ were to increase new investments and job creation in the zones designated by the program. In their assessment, Hansen and Kalambokidis (2010) determined at the county level there was little evidence of economic growth during the first three years of the program, but there was significant evidence of positive job creation.

These articles generally demonstrated that the programs evaluated had some type of positive result. However, several articles with various types of econometric analysis have found little to no positive impact from other types of government programs. Holtz-Eakin and Schwartz (1995) explored the existence of productivity spillovers from the state highway system, because it is implied that public infrastructure is linked to large productivity effects from public spending. They found no evidence of increased spillover productivity, despite the fact of the highway systems are designed to have interstate linkages. Evans and Karras (1994) used panel data from

[7] Rupasingha and Wang (2017) used county level data instead of individual business level data because they did not have access to individual borrower data. Also, they did not use employment numbers as their dependent variable because the employment data are withheld to avoid disclosing data of individual companies; instead, they used establishment data as the dependent variable.

the lower 48 states to determine to what extent government capital and services positively impact private production. In their results, there was only significant evidence for governmental education services being productive, and no findings of productivity related to other governmental activities and services (such as highway, sewer and sanitation, and police and fire services). Thus, not all academic research supports the hypothesis that government program expenditures are effective tools for economic development.

A program with similar characteristics as the B&I is the SBA's 7(a) and 504 loan programs, which was evaluated by Craig, Jackson III, and Thomson (2007).[8] In their analysis, Craig et al. (2007) used a panel data set with approximately 2,200 local market observations, and found a small positive relationship between the level of SBA-guaranteed lending in a local banking market and future per capita income growth. However, the results do not necessarily document a causal relationship because Craig et al. (2007) do not know if the SBA loan guarantee program is directly contributing to growth or if the programs are just acting as a proxy for small business lending.

To my knowledge, there have only been two studies that focused on the B&I program specifically. The first study was Johnson (2009), who focused solely on the impact of the program on county employment for loans obligated from 1986 to 2003. The analysis found through ordinary least squares regression and propensity score matching that there was an increase in employment of 3% to 6% in county employment-per-capita, and a 3% to 5% decrease in county earnings-per-worker growth over the two years immediately following the loan (Johnson, 2009). Thus, the program seemed to increase jobs, but also resulted in lower average earnings – and as such the effects on total county earnings was indeterminate. The most recent

[8] Both the 7(a) and 504 loan guarantee programs, but their maximum amount to guarantee is $1,000,000 and $1,300,000 respectively. Both programs can fund businesses in rural and metro areas, unlike the B&I, which only lends to businesses in rural areas.

study by Rupasingha, Crown, and Pender (2018) used similar economic analysis techniques as previous studies in this area. One unique feature of Rupasingha, Crown, and Pender's (2018) research is their linking of the National Establishment Time-Series (NETS) data to the B&I program administrative data to create a database of 4,361 B&I recipient business between 1990-2003. The NETS database is not openly accessible to the public, but its inclusion of individual businesses allows for a more pointed analysis. Rupasingha, Crown, and Pender (2018) found significant evidence that a business receiving a loan survived longer and created more jobs than non-recipient businesses. They also found that increasing the size of the loan decreased the risk of failure. However, there was no significant impact on business employment growth from increasing the size of the loan (Rupasingha, Crown, and Pender, 2018).

Both previous B&I studies focus on job and economic growth at the county and business level, but they do not evaluate the impact the program had on tax revenue of a community. Both studies also use older (pre-2008) data on the B&I program, whereas this paper uses data from 2005-2015. Tax revenue is vital for rural America as noted in the previous discussion of tax adequacy. A study by Walzer, Blanke, and Evans (2018) discussed the factors affecting retail sales in Illinois non-metro cities with population below 50,000 people, and offered five suggestions of how to build or revive the commercial areas. Walzer et al. (2018) also argued that taxes to generate revenue for local services are crucial for stabilizing populations in areas where outmigration is a common occurrence.

A scan of the relevant literature suggests there are very few studies that focus directly on tax revenues associated with specific programs. Of the two studies found, both analyzed the Small Business Development Center (SBDC) counseling activities and the performance improvements of long-term clients, which should hypothetically result in a generation of incremental tax revenue increase. The most recent article is the Chrisman (2017), which appears to be an updated study of Chrisman and Katrishen (1994). The two articles used the same methodologic approaches in their analysis – however, it should be noted that they are not regression-based. The approach was to calculate a weighted average of each tax figure based on the proportion of clients served by each state to the total clients served by the SBDC, and assumptions about how the jobs and sales supported by their program impacted local tax revenues (Chrisman, 2017; Chrisman and Katrishen 1994). Both of the articles took necessary steps to ensure the respondents were representative of the population. Notably, the articles also examined other economic impacts, such as jobs and sales, and estimated that the SBDC program increased sales and jobs during both times analyzed.

Chrisman and Katrishen (1994) found that the performance improvements generated roughly $2.61 in incremental tax revenue for every dollar expended on the entire SBDC program in the US. Chrisman (2017) found that approximately $2.42 of tax revenue was generated from every dollar spent on SBDC programming. However, these results should be taken with caution because they are only estimates based on responses to a questionnaire from SBDC clients and did not use any actual tax revenue data.

CHAPTER III

DATA AND MATERIALS

DATA

The data used in this paper is obtained at the place level, as designated by the US Census Bureau, for the state of Oklahoma. The US Census Bureau explains that a place is any area representing officially incorporated governments such as cities, towns, villages, municipality, township, community, populated place, neighborhood, postal place/zip code, populated place, or boroughs (Census Reporter, N.D., and Ratcliffe, N.D.). The number of places in a state can vary over time as new cities arise or some places are annexed into others. However, the analysis only focuses on Oklahoma places consistently listed by the Census between 2000 and 2015.

US Census Bureau's American Factfinder website was used to obtain all of the demographic characteristics of all places in Oklahoma for 2005, 2010, and 2015. In order to get the 2005 demographic characteristics, a linear approximation was used to get the average of the 2000 and 2010 official census numbers. The official count from the census was used for the 2010 data, and the American Community Survey (ACS) 2013-2017 five year estimates were used to produce the 2015 data from American Factfinder.

One challenge of working with this data is the inconsistency in the number of places recorded from year to year. In 2005 there were 689 total places in Oklahoma; this increased to 730 in 2010 and moved up to 743 in 2015. We break these places into five different size categories in Tables 2, 3 and 4. The number of places in each size category varies over time.

This variation could be because new places sprung up, some may have been annexed into other towns, and some places were dissolved. The population in the places also fluctuates up and down over the years, which results in some of the towns moving from one size category to another.

The main variable of interest is the location and amount of the B&I loan guarantees issued, which the Oklahoma Rural Development state office graciously provided. During 2005 to 2015 there were a total of 114 loan guarantees issued – 55 of which were issued during the 2005 to 2009 time period and 59 that took place between 2010 and 2014. Figures 1 and 2 show the population size and the amount of the loan received of the loan recipient, respectively. However, one issue with the figures is that some places received multiple loans in one or both time periods. During time period one (2005-2009), only 40 different places in Oklahoma received the 55 loans issued, and only 37 different places received the 59 loans made during time period two (2010-2014); this is shown in Tables 2, 3 and 4. The loan data used in the Tables 3 and 4 are the same (loans made during 2010-2014), but the distribution of the loans by size categories are different[9].

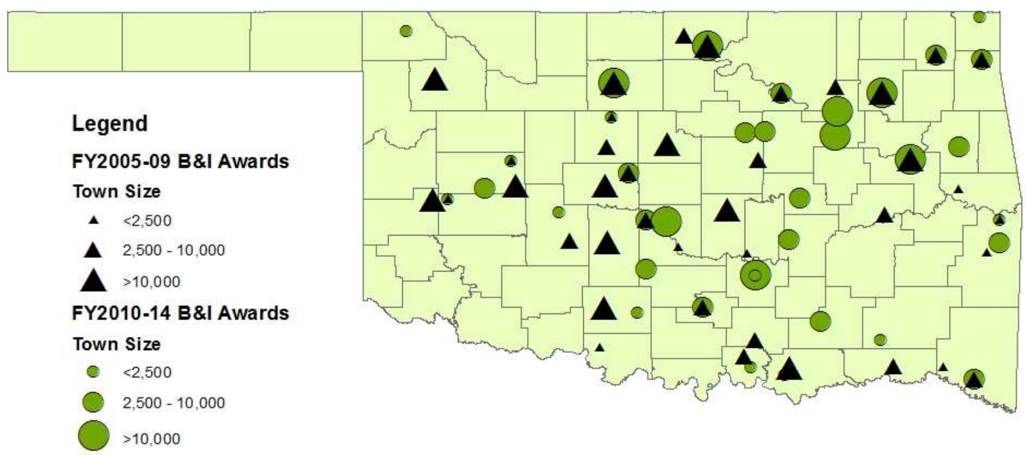
The 2,500 – 9,999 size category received the most loans with 21 loans for each time period, and the < 500 category received the least with only one loan in the 2010-2014 period. Table 2, 3, and 4 show some basic descriptive statistics for Oklahoma places in 2005, 2010, and 2015, broken out by whether or not they received a loan during the 2005 – 2009 or 2010 – 2014 period. The average population is higher in both 2005 and 2010 for towns that received a loan (versus nonrecipients), but not all of the averages are statistically different (Table 2 and 3). However, in Table 4 when the 2015 population is used, the average population with loans is lower than the population without loans in the 500-2,499 category, but the rest of the size categories have similar results as Tables 2 and 3. Furthermore, population is higher and

[9] The distributions of the loans by size category changes between Table 3 and 4 because the populations used for the breaks are different (2010 in Table 3, 2015 in Table 4). The populations of the towns are changing over time, causing towns to switch size categories.

significant for recipient cities across all size categories in Table 3, but only significant for the > 50,000 category in Table 4.

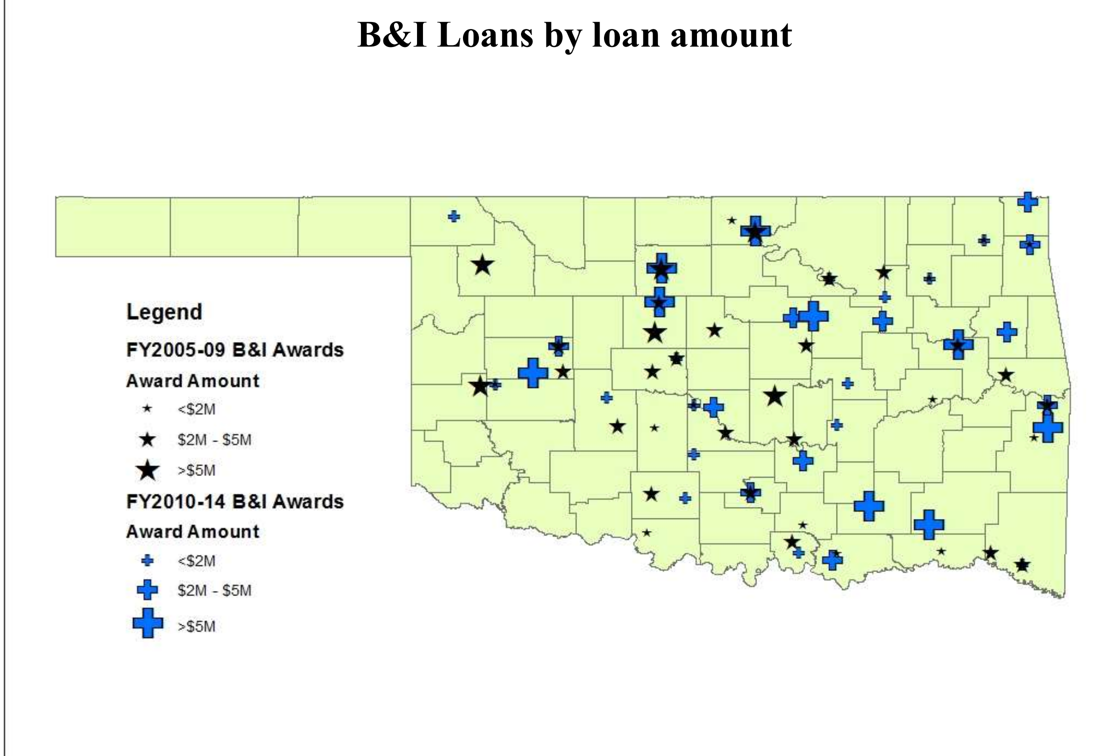
Figure 1

B&I Loans by size of the town



Source: B&I data provided by USDA RD

Figure 2



Source: B&I data provided by USDA RD

Median household income (MHI) is another variable in the basic descriptive statistics tables. Most of the MHI values for loan recipients are lower than those for nonrecipients, but few are statistically significant. However, of the three values that are statistically significant two have a lower MHI in recipient places than nonrecipients (Tables 2, 3, and 4). The fact that more recipient places had lower MHIs may be due to the application requirements for the B&I program[10].

[10] The B&I priority scoresheet awards more points to locations with high poverty levels, which is likely correlated with lower MHI.

Table 2

2005 Basic Descriptive Statistics									
		Number of Places		2005 Population			2005 Median Household Income		
Place size	total loans	With Loans	No Loans	With Loans	No Loans	P-value	With Loans	No Loans	P-value
<500	0	0	347	–	215		–	$31,957	
500 - 2,499	13	10	201	1,479	1,073	0.005***	$29,042	$31,707	0.143
2,500 - 9,999	21	16	74	4,917	4,536	0.245	$34,174	$33,712	0.554
10,000 - 49,999	19	13	21	20,976	20,856	0.487	$34,226	$41,324	0.026**
>50,000	2	1	6	392,478	159,522		$37,303	$46,009	
Total	55	40	649						
Total places in Oklahoma	689								

* represents statistically different means at 10% significance, ** 5% significance, *** 1% significance

Table 3

2010 Basic Descriptive Statistics									
		Number of Places		2010 Population			2010 Median Household Income		
Place size	total loans	With Loans	No Loans	With Loans	No Loans	P-value	With Loans	No Loans	P-value
<500	1	1	378	497	209		$45,192	$38,601	
500 - 2,499	19	11	210	1,430	1,094	0.018**	$39,972	$35,876	0.094*
2,500 - 9,999	21	16	71	5,366	4,559	0.069*	$39,342	$38,803	0.550
10,000 - 49,999	13	6	29	26,698	19,019	0.039**	$40,221	$44,045	0.246
>50,000	5	3	5	360,9943	77,315	0.015**	$42,765	$54,307	0.0824*
Total	59	37	693						
Total places in Oklahoma	730								

* represents statistically different means at 10% significance, ** 5% significance, *** 1% significance

Table 4

2015 Basic Descriptive Statistics									
		Number of Places		2015 Population			2015 Median Household Income		
Place size	total loans	With Loans	No Loans	With Loans	No Loans	P-value	With Loans	No Loans	P-value
<500	0	0	391	–	200		–	$41,893	
500 - 2,499	16	9	208	1,088	1,103	0.466	$44,354	$42,584	0.337
2,500 - 9,999	25	19	73	5,119	4,562	0.148	$42,058	$44,472	0.288
10,000 - 49,999	9	5	29	22,337	20,185	0.310	$45,186	$49,645	0.267
>50,000	9	4	5	300,293	81,672	0.053*	$49,686	$59,982	0.104
Total	59	37	706						
Total places in Oklahoma	743								

* represents statistically different means at 10% significance, ** 5% significance, *** 1% significance

An important part of this analysis on city tax revenue is the number of local retail businesses. Thus, a third data source used in this research is the zip code business pattern data (ZCBP), which includes the total number of retail businesses for all zip codes in Oklahoma. This data is gathered at the zip code level, which does not match the place-level demographic characteristics. This requires an extra step of mapping each zip code to the designated place level to keep the data consistent. In order to map the zip codes to the place level, a conversion from the Missouri Census Data Center (2016) allocating the percentage of each zip code to the associated place level was used. This results in some places being allocated a percentage (non-whole numbers) of a retail business rather than a whole number.

As Tables 5, 6, and 7 show, the average total number of retail business increases as place size increases, which is an expected outcome because having more people in a community is generally associated with more businesses. The average total number of retail businesses is higher for loan recipients than nonrecipients in both 2005 and 2010, and the averages are all statistically different at the 5% level (Table 5 and Table 6). In 2015, this same pattern holds true, with all differences at least statistically significant at the 10% level. The noted difference between recipient/nonrecipient communities in terms of population, MHI, and number of businesses suggest that it will be important to use a technique such as coarsened exact matching to assure the estimation of a "true" counterfactual.

The Oklahoma Tax Commission (OTC) is the source for the total retail sales (TRS) tax data. The TRS data used in this research is SIC code based (retail trade are codes 52-59) from 2005 to 2015, and it is collected at the place level for all retail sectors. The OTC gathers data on the amount of tax revenue each place has collected in each time period. Then the tax revenue for each place is divided by the place-level sales tax rate to get the TRS expenditures for each place (i.e. dollars spent in that town). The TRS data for each place is then divided by the population to obtain the TRS per capita data. Tables 5, 6, and 7 show that in 2005, 2010, and 2015, total retail sales (TRS) per capita in 2010 real dollars are all higher for loan recipients compared to nonrecipients, but not all averages are statistically significant[11].

In this research we are also interested in the *change* in TRS from 2005 to 2010 and 2010 to 2015, which is also displayed in Table 5, 6, and 7. The change in TRS for the two time periods are all higher for places receiving a loan, except for 10,000 – 49,999 size category in Table 7, but not all average changes are statistically higher than those that occurred in nonrecipients cities. This offers some support for the hypothesis that B&I loans can result in positive increases in TRS.

[11] In 2005 and 2015, there is not an average TRS per capita with loans for the <500 place size because the category did not receive a loan during the 2005-2010 time period.

Table 5

2005 Retail Descriptive Statistics									
	2005 Total number of Retail Businesses			2005 TRS Per Capita (2010$)			2005-2010 TRS Change (In millions of dollars)		
Place size	With Loans	No Loans	P-value	With Loans	No Loans	P-value	With Loans	No Loans	P-value
<500	0.00	0.92		$0	$4,869		$0	$0.271	
500 - 2,499	8.70	5.94	0.007***	$6,592	$5,796	0.180	$1.799	$0.977	0.022**
2,500 - 9,999	34.22	25.60	0.067**	$14,021	$9,516	0.002***	$10.532	$8.426	0.294
10,000 - 49,999	131.44	95.59	0.038**	$15,743	$13,305	0.050*	$41.810	$53.341	0.750
>50,000	1842.24	636.14		$17,723	$13,388		$445.488	$299.350	

* represents 10% significance, ** 5% significance, *** 1% significance

Table 6

2010 Retail Descriptive Statistics									
	2010 Total number of Retail Businesses			2010 TRS Per Capita			2010-2015 TRS Change (In millions of dollars)		
Place size	With Loans	No Loans	P-value	With Loans	No Loans	P-value	With Loans	No Loans	P-value
<500	3.00	0.84		$7,990	$6,509		$7.083	$0.641	
500 - 2,499	11.04	5.29	0.000***	$7,711	$5,649	0.034**	$4.069	$2,003	0.015**
2,500 - 9,999	33.78	23.43	0.007***	$13,570	$9,771	0.006***	$16.189	$9.679	0.039**
10,000 - 49,999	144.16	90.99	0.015**	$15,163	$13,768	0.224	$71.050	$63.586	0.380
>50,000	1447.08	236.95	0.011**	$14,919	$12,378	0.060*	$1,320.0	$203.669	0.031**

* represents 10% significance, ** 5% significance, *** 1% significance

Table 7

2015 Retail Descriptive Statistics									
	2015 Total number of Retail Businesses			2015 TRS Per Capita			2010-2015 TRS Change (In millions of dollars)		
Place size	With Loans	No Loans	P-value	With Loans	No Loans	P-value	With Loans	No Loans	P-value
<500	0	1.02			$10,424		$0	$0.660	
500 - 2,499	7.500	5.115	0.040**	$10,690	$6,682	0.008***	$4.604	$1.987	0.007***
2,500 - 9,999	31.058	22.038	0.009***	$14,255	$10,734	0.017**	$14.070	$9.636	0.096 *
10,000 - 49,999	128.661	91.525	0.073*	$16,151	$15,100	0.372	$41.300	$63.586	0.166
>50,000	1184.09	259.901	0.038**	$16,072	$13,105	0.037**	$1,050.0	$203.669	0.061 *

* represents 10% significance, ** 5% significance, *** 1% significance

Figure 3 shows the average TRS per capita from 2005 to 2015 for places in all five of the size categories, regardless of whether they received a B&I loan. The TRS per capita for all size categories follows a general trend of increasing from 2005 to 2015. However, the change from 2009 to 2010 was a decrease for nearly all sizes, which is expected due to the Great Recession. Further investigating the TRS impacts of B&I loans, Figure 4 shows the TRS per capita of loan recipient communities five year prior to and five year after receiving a loan guarantee during time period one. Some cities appear to experience an increase in TRS per capita after receiving a B&I loan, while others do not. The results in Figure 4 suggests that a more rigorous analysis is needed to further investigate the true impact of the B&I program has on TRS and TRS per capita of places in Oklahoma.

Figure 3

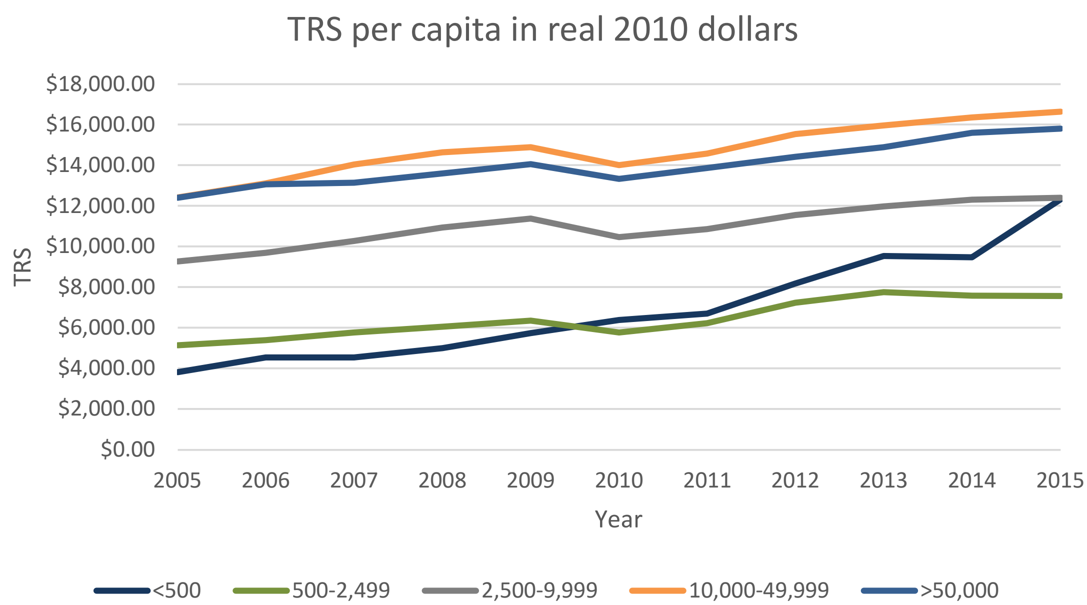
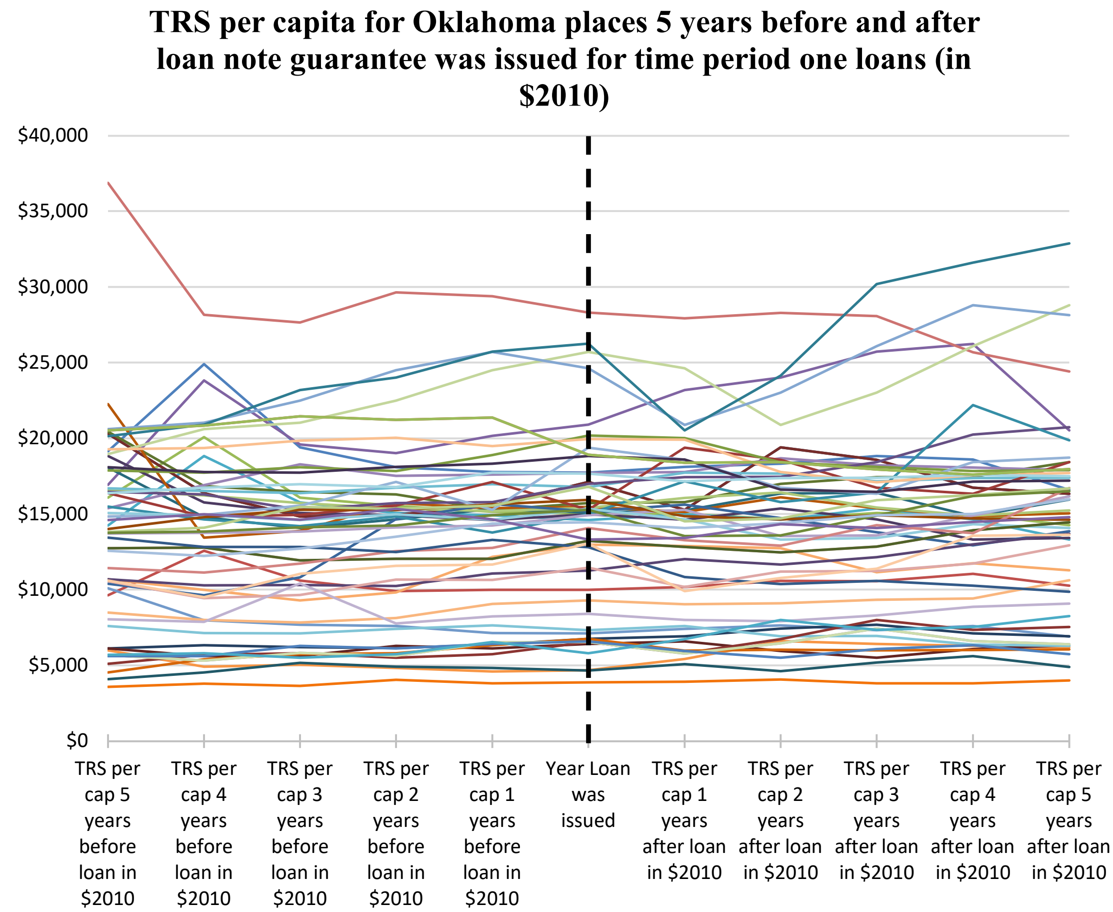


Figure 4



This research focuses on places in Oklahoma with populations greater than or equal to 500 people and less than 50,000 people.[12] One reason the minimum population size is limited to 500 is because of the large difference between total observations and observations with TRS data. In 2005 there were 347 total census observations with <500, but only 186 places had TRS data. Similarly, in 2010 there were 378 total census observation with only 188 places recording TRS data – likely because many lacked a sales-tax collecting business. Therefore, limiting the data to places with only 500 people or more will help deal with this data inconsistency. Further, as Table 3, 4, and 5 show, only one town with less than 500 people received a B&I loan during 2005-2015.

[12] A business usually cannot receive a B&I loan guarantee if located in a place with a population of 50,000 or more people. However, the location of the loan can be deemed eligible by the USDA eligibility map, which could be the reason there are four loans for places with population > 50,000 in tables 2, 3, and 4.

The remaining size categories are relatively consistent with the number of observations reporting TRS and total observations and only a few places switch size categories (i.e. moving to a higher or lower population grouping) across the years of analysis.

METHODS

The main challenge with this research is estimating the counterfactual – i.e. the tax revenue a city with a B&I recipient would have earned if that business did not receive a B&I loan. This is a similar issue that Rupasingha, Crown and Pender (2018) faced in their analysis of the B&I program. For example, if a business needed a loan guarantee but it did not receive one, there may have been negative repercussions to that business that impacted local retail sales (for example, going out of business or losing employees). Data on this alternative reality does not exist, so I utilize econometric techniques to try and answer the research question at hand. First, simple linear regression is used to get preliminary results that estimate the impact of a loan on tax collections over a specific period before running a first-differenced linear regression model that focuses on changes over time. Next, coarsened exact matching is used to prune the dataset and ensure better matches between the treated and control groups. Finally, the last step is to rerun the chosen regression models with the pruned dataset.

Simple Linear Regression

Ordinary least squares (OLS) regression is the first econometric technique used in this research to analyze the data. The OLS regression equation used is:

$$Y_t = X_t\,\beta_t + L_{t-1}\gamma_t + \epsilon \qquad (1)$$

where Y_t is the TRS (and TRS per capita) in time period *t,* X_t are the control variables (population, median household income, number of retail businesses, percent black, percent with a bachelor’s degree, percent of population with income below poverty, and local tax rate),

L_{t-1} represents total loan amount for each place made over the prior period 2005-2009 (or 2010-2015), and ϵ is the error term of the model, for all time t. γ is the parameter of focus in this research, with a positive and significant γ implying that the loans are having an impact.

First-differenced Linear Regression

The second step of the analysis is to use a first-differenced specification of the OLS regression because I am also interested in explaining the *changes* in retail sales between years. The first-differenced OLS regression equation used in this research is:

$$\Delta Y_t = \Delta X_t \, \beta_t + \Delta L_t \gamma + \epsilon \qquad (2)$$

where ΔY_t is the change of TRS (or TRS per capita) in time period *t, t* is either 2005-2010 or 2010-2015, ΔX_t are the changes in control variables (ΔMHI, Δ poverty, Δ population, Δ # of businesses, etc.) in time period *t*, ΔL_t is the change in loan variable, only if the loan takes place in time period *t*, and ϵ is the error term.

A place receiving a loan guarantee may not see an impact on TRS the first year the loan guarantee is issued. Therefore, this research is going to analyze the change of TRS from 2005 to 2010 and 2010 to 2015 by looking at cumulative loans granted during the two periods of analysis. The analysis is most interested in the estimate of γ – i.e. whether the loans had a statistical impact on changes in total retail sales.

Coarsened Exact Matching

In order to overcome the identification challenge (matching of the treated and control groups), Rupasingha, Crown and Pender (2018) constructed counterfactuals using the outcome of untreated businesses, which is also the approach used in this research. Creating counterfactuals helps determine what would have happened to the places' TRS if a business had not received a loan guarantee. While the regression-based approaches above use all observations, the differences between treated and control groups for variables like the number of businesses and/or population (Tables 2-7) may be problematic.

To account for these differences, Coarsened Exact Matching (CEM) is utilized to reduce the dataset into control and treated groups that are more similarly matched. CEM is a relatively new econometric technique, with one of the first demonstrations coming from Blackwell et al. (2009). CEM is used for estimating causal effects when covariates have differences between treated and control groups (Blackwell et al., 2009). The control group is the counterfactual of the dataset. Other techniques, such as propensity score matching (PSM), have tried to address this. However, Iacus, King and Porro (2011) explain that PSM and other matching methods called "Equal Percent Bias Reducing" (EPBR) are only designed to satisfy weaker properties because most of the time the three main assumptions of EPBR are not satisfied in observational data[13]. EPBR methods only improves for the mean imbalance (main effects) of the variables in the dataset, but does not guarantee any level of imbalance reduction of the data by making assumptions that are normally unverifiable about the data set (Iacus, King, and Porro, 2011; Iacus, King and Porro 2012).

CEM is a "Monotonic Imbalance Bounding" matching method that requires fewer to no assumptions than EPBR matching methods, and it is easier to use and understand (Blackwell et al., 2009 and Iacus, King, and Porro, 2012). The basic idea of CEM is to 'coarsen' each variable into strata of treated and control groups that only have the same identifiable values; thus restricting the dataset only to treated and control groups that are similar in their underlying characteristics (Blackwell et al., 2009 and Iacus, King, and Porro, 2012). This approach likely reduces bias. Furthermore, CEM bounds the degree of model dependence and causal estimation

[13] The three main assumptions of EPBR are: "X (a k-dimensional data set) is drawn randomly from a specified population X, the population distribution of X is an ellipsoidally symmetric density or a discriminant mixture of proportionally ellipsoidally symmetric densities, and the matching algorithm applied is invariant to affine transformations of X" (Iacus, King, and Porro, 2012)

error by ex-ante user choice, does not require extra steps to prune the data, and meets the congruence principle (Blackwell, 2009)[14].

The data is coarsened on certain demographic characteristics likely to influence TRS (population, MHI, percent population is black, percent with a bachelor's degree, percent of population with income below poverty, and local tax rate). Once the dataset is pruned, the same regressions discussed earlier will be used.

Robustness Check

Since CEM is a relatively new methodology, Propensity Score Matching (PSM) is also used to ensure the results are robust. PSM is similar to CEM, where a counterfactual of the untreated group is constructed, but PSM's method of constructing the counterfactual is different than CEM's approach. PSM uses a logistic regression to estimate the likelihood of a community being treated (in this case, a place receiving a B&I loan or not) (Dehejia and Wahba, 2002; Whitacre, Shideler, and Williams, 2016). This is different than CEM's approach of creating "bins" of identifiable values, and then matching control and treated groups with similar bin structures. The PSM approach results in a propensity score reflecting the likelihood of a community receiving a loan guarantee; each community is put in treated and control groups based on whether or not they *actually* received a loan. Once the propensity scores are estimated, alternative matching methods are used to match control communities with otherwise similar treated communities (i.e. with similar propensity scores). In this research, five different matching methods (nearest neighbor, kernel default, kernel normal, kernel biweight, and radius caliper 0.01) are used to compare to the results of the CEM methodology. The nearest neighbor approach matches the treated group with the nearest neighbor(s) (i.e. the five closest matches to each propensity score) within the control

[14] "The congruence principle states that the data space and analysis state should be the same," and if the method does not meet this standard the results are often strange and/or counterintuitive (Blackwell, 2009).

group, with or without replacement (Caliendo and Katrishen, 1994). The nearest neighbor approach in this research matches the nearest five neighbors with replacement. Matching with replacement reduces the bias and increases the average quality of the matches (Caliendo and Katrishen, 1994). Kernel matching is a nonparametric matching estimator that uses weighted averages of the control group to construct the counterfactual to match to treated communities (Caliendo and Katrishen, 1994). As such, all control group observations are used, but those with closer propensity scores are given more weight. The three different types of kernel functions used in this research to construct the counterfactuals are the default (Epanechnikov), biweight, and the normal (Gaussian) kernel. The radius caliper matching is similar to nearest neighbor except that it matches all available controls to within the distance of the caliper (Caliendo and Katrishen, 1994). The caliper distance in this research is 0.01. Again, PSM is not the main methodology used for this research, rather it is just a robustness check of the CEM results.

CHAPTER IV

RESULTS

The results of the simple OLS and first-differenced regression models with and without the CEM specification for the two different time periods of interest are presented in Table 8, 9, 10, and 11. Table 8 analyzes the impact on 2010 TRS for places in Oklahoma receiving a loan during time period one (2005-2009), whereas table 10 focuses on 2015 TRS and loans received in time period two (2010-2014). Table 9 and 11 perform similar regressions, but changes the dependent variable to TRS per capita. The results of the regression models show mixed results for the impact that B&I loans have on TRS and TRS per capita, but demonstrate stronger impacts in the earlier (2005-2010) period.

The first two sets of columns in Table 8 use equation 1 to estimate the TRS. The simple OLS column uses the unpruned data whereas the simple OLS with CEM weights columns use the pruned data. The L_1 imbalance measure demonstrates that CEM is reducing the amount of imbalance in the data. After controlling for demographic characteristics in 2010, both simple OLS models find that the loan amount variable is significant. In the simple OLS *without* CEM weights the loan variable is significant at the five percent level, and it is significant at the one percent level *with* the CEM weights, when the dependent variable is TRS 2010 (Table 8). Furthermore, in the first-differenced models (third and fourth set of columns) the loan amount variable is statistically significant at the one percent level without CEM weights, and at the five percent significance level with the CEM specification (Table 8). Interpreting the loan amount coefficient of the simple OLS with CEM weights on the 2010 TRS suggests that a one percent increase in loan amount would result in 0.0233% higher TRS collected. As expected, the R-squared values are smaller for the first-differenced models than they are for the basic OLS models. Additionally, the R-squared values increases slightly in the OLS with CEM weights and drastically increases in the first-differenced model with the CEM-pruned specification. This suggests that using CEM results in a more balanced dataset, as expected. The *L1* statistic is the measure of imbalance of the data. The *L1* statistic for the whole dataset is only reduced from 0.9963 to 0.9020 (Table 8). Even though the L1 statistic for the whole dataset is only slightly reduced, the *L1* statistic for each individual variable is reduced, which lowers the imbalance of the variables (Appendix J). However, the one negative of the CEM specification is that it drastically reduces the number of observations in the data set in the regression models (for example, going from 307 to 67 in the basic OLS models in Table 8).

Other variables generally show their expected signs, suggesting that the model is behaving according to economic theory. One variable hypothesized to have a significant impact on TRS is the number of retail businesses. As expected, the number of retail businesses variable is significant at least at the five percent level in both simple OLS models, and in both first-difference models it is at the one percent significance level (Table 8). However, the interpretation of the retail businesses coefficient differs between the OLS and first-differenced models. In the simple OLS model with the CEM specification when the number of retail businesses increases by one, TRS would increase by 0.0095%. However, in the first-differenced CEM model TRS would decrease by 0.1348% when the number of businesses increases by one (Table 8). The hypothesis for this negative impact on TRS is that the poorly run businesses are going out of business over time, leaving only more efficient stores.

A third variable hypothesized to have an influence on TRS is the percent of population with a bachelor's degree or higher. In time period one when TRS is the dependent variable, the bachelor's degree coefficient is positive and significant at the five percent level in both basic OLS models. However, when the first-differenced models are used, the education variable is not significant in either (Table 8).

Table 8

Pre and post CEM regression with *Log of TRS 2010* as dependent variable, 2010 control variables

	Basic OLS		Basic OLS with CEM weights		first difference OLS		first difference with CEM weights	
Number of Observations	307		67		262		61	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Log of 2010 population	1.1969	0.000***	1.0791	0.000*	6.6908	0.000***	14.8330	0.000***
Log of Median Household Income 2010	-0.3695	0.049**	0.9862	0.001***	1.0156	0.511	-2.4489	0.444
Total retail business 2010	0.0041	0.011**	0.0095	0.001***	-0.0751	0.001***	-0.1348	0.001***
Percent of population black 2010	-1.2635	0.002**	-2.5425	0.054*	-1.9595	0.827	-8.1178	0.795
Percent of population with bachelors or higher in 2010	1.4888	0.043**	2.1966	0.041**	-4.4003	0.449	13.0271	0.251
Percent of population with income below poverty 2010	0.3193	0.538	3.6309	0.000***	3.8917	0.226	9.7134	0.202
Log of amount of loan received in 2005 – 2009	0.0185	0.010**	0.0233	0.000***	0.0975	0.000***	0.0671	0.013**
Rate 2010	21.4616	0.000***	-38.9517	0.000***	-35.0834	0.157	5.1941	0.933
constant	10.1085	0.000***	-1.9024	0.523	14.0816	0.000***	14.3525	0.000***
R-Squared	0.8747		0.9677		0.2116		0.4296	
L1	0.9963		0.9020		0.9963		0.9020	

*Note: * indicates the following *** = p<.01, ** = p<.05, * p=<.1*

L1 is the measure of imbalance of covariates between treatment and control groups (smaller L1 implies better balance)

Table 9 applies the same independent variables and focuses on the same time period as Table 8, but the dependent variable is changed to TRS per capita. The results in Table 9 are similar to those in Table 8. Both simple OLS models demonstrate that the amount of loan received is positive and significant at the five percent level, but in the first-differenced models only the CEM specified model is significant, at the five percent level (Table 9). The interpretation of the loan variable in the first-differenced with CEM specification is that a one percent increase in loan amount would increase the TRS per capita by 0.0048% (Table 9). Additionally, the R-Squared value increases drastically in both models when the CEM pruned dataset is used.

Focusing on the number of retail businesses variable when TRS per capita is the dependent variable, it is significant at the five percent level in both simple OLS models, but not significant in either first-differenced models (Table 9). Switching focus to the percent of the population with a bachelor's degree, the results of the education variable are similar to the number of retail businesses. Both simple OLS models are significant at the five percent level, and the first-differenced with the CEM specification is significant at the ten percent level.

Table 9

Pre and post CEM regression with *Log of TRS per capita 2010* as dependent variable, 2010 control variables

	Basic OLS		Basic OLS with CEM weights		first difference OLS		first difference with CEM weights	
Number of Observations	307		67		303		67	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Log of 2010 population	0.1969	0.000***	0.0791	0.300	-0.2204	0.206	-0.0135	0.965
Log of Median Household Income 2010	-0.3695	0.049**	0.9862	0.001***	0.2004	0.278	-0.0310	0.908
Total retail business 2010	0.0041	0.011**	0.0095	0.001***	0.0016	0.577	-0.0024	0.474
Percent of population black 2010	-1.2635	0.002***	-2.5425	0.054*	0.4336	0.694	6.9051	0.009***
Percent of population with bachelors or higher in 2010	1.4888	0.043**	2.1966	0.041**	-0.8924	0.209	1.8644	0.058*
Percent of population with income below poverty 2010	0.3193	0.538	3.6309	0.000***	0.1430	0.708	1.0133	0.121
Log of amount of loan received in 2005 – 2009	0.0185	0.010**	0.0233	0.000***	0.0013	0.588	0.0048	0.035**
Rate 2010	21.4616	0.000***	-38.9517	0.000***	0.0041	0.999	-6.3353	0.202
constant	10.1085	0.000***	-1.9024	0.523	0.1246	0.000***	0.1308	0.004***
R-Squared	0.3860		0.7299		0.0161		0.2452	
L1	0.9963		0.9020		0.9963		0.9020	

*Note: * indicates the following *** = p<.01, ** = p<.05, * p=<.1*
L1 is the measure of imbalance of covariates between treatment and control groups (smaller L1 implies better balance)

This research also analyzed the impact of TRS (2010) for a place receiving a loan when the previous time period demographic characteristics (2005) were used to coarsen the dataset and run the regression models. The basic idea is to be able to “predict” later TRS using earlier period characteristics. The same approaches and techniques used above in the “same year” (2010 TRS and 2010 demographic characteristics) regression models are applied here in the “different year” regression models. The results in time period one of the “different year” regression models are similar to the “same year” regression models. The loan variable is statistically significant at least at the five percent level in all four models when 2010 TRS is the dependent variable (Appendix A). In the “same year” regression models, the loan variable was significant in all four models (Table 8). Switching the dependent variable to 2010 TRS per capita, only the basic OLS models are significant (Appendix B) in the “different year” regression models. This is different than the “same year” regression models where both the basic OLS models and the first-differenced *with* the CEM specification is significant (Table 9).

The results for the “same year” regression models in time period one shows that the main variable of interest, the log of the amount of loan received variable, is significant in seven of the eight models (Table 8 and 9), while only being significant in six of the eight “different year” models (Appendix A and B). These results support the hypothesis that places receiving loans see their TRS expenditures increase.

The second time period regression coefficients and p-values are depicted in Tables 10 and 11. All of the independent variables used in period two are the same as in time period one except the 2015 values are used instead of 2010. Here, the loans are those made during 2010-2014. However, the regression results in time period two are quite different than in time period one. The OLS and first-difference models *without* the CEM specification both display significant loan coefficients at the five percent level when TRS is the dependent variable. However, neither coefficient is significant when the pruned dataset is used (Table 10). The results of the total number of retail businesses and the percent of the population with a bachelor’s degree or higher differ from the loan variable. The number of retail businesses variable is significant at the one percent level in both OLS models. However, the education level variable is significant at the five percent level in the basic OLS *without* the CEM specification, and it significant at the ten percent level in the first-differenced *with* the CEM specification.

Table 10

Pre and post CEM regression with *Log of TRS 2015* as dependent variable, 2015 control variables

	Basic OLS		Basic OLS with CEM weights		first difference OLS		first difference with CEM weights	
Number of Observations	304		42		269		36	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Log of 2015 population	1.1644	0.000***	1.4106	0.000***	2.0076	0.037**	1.9082	0.515
Log of Median Household Income 2015	-0.5073	0.048**	-1.1264	0.241	-1.0821	0.074*	0.7322	0.590
Total retail business 2015	0.0051	0.008***	0.0334	0.007***	0.0217	0.390	0.0032	0.969
Percent of population black 2015	-1.7487	0.000***	-13.5702	0.002***	0.5445	0.905	35.3124	0.055
Percent of population with bachelors or higher in 2015	2.0141	0.030**	6.3511	0.103	2.4240	0.309	17.5137	0.060*
Percent of population with income below poverty 2015	0.3559	0.641	-2.5463	0.269	-0.6086	0.691	0.1114	0.973
Log of amount of loan received in 2010 - 2014	0.0208	0.015**	0.0188	0.132	0.0549	0.01**	0.0194	0.590
Rate 2015	30.0396	0.000***	23.5749	0.175	-56.4361	0.07*	186.1157	0.007
constant	11.6638	0.000***	16.6354	0.109	15.1294	0.000***	14.6841	0.000
R-Squared	0.8394		0.9240		0.0784		0.5680	
L1	0.9853		0.9264		0.9853		0.9264	

*Note: * indicates the following *** = p<.01, ** = p<.05, * p=<.1*

L1 is the measure of imbalance of covariates between treatment and control groups (smaller L1 implies better balance)

Furthermore, when the dependent variable of the models are changed to TRS per capita, the simple OLS *without* the CEM weights is the only model where the loan variable is statistically significant (Table 11). The number of retail businesses variable is significant at the one percent level in both OLS model, but not significant in either first-difference models. The education level variable is statistically significant only in the basic OLS *without* the CEM specification as well.

Again, the R-squared values increases when the CEM specification is used. The L1 statistic only reduces from 0.9853 to 0.9264 for the whole dataset (Table 11), but the L1 statistic for almost all variables is reduced after the pruning, which reduces the imbalance of dataset (Appendix K).

Table 11
Pre and post CEM regression with *Log of TRS per capita 2015* as dependent variable, 2015 control variables

	Basic OLS		Basic OLS with CEM weights		first difference OLS		first difference with CEM weights	
Number of Observations	304		42		302		40	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Log of 2015 population	0.1644	0.006***	0.4106	0.028**	-0.7277	0.000***	-0.6273	0.047**
Log of Median Household Income 2015	-0.5073	0.048**	-1.1264	0.241	-0.0100	0.899	-0.0518	0.765
Total retail business 2015	0.0051	0.008***	0.0334	0.007***	0.0044	0.212	0.0126	0.247
Percent of population black 2015	-1.7487	0.000***	-13.5702	0.002***	-0.3687	0.528	2.2883	0.298
Percent of population with bachelors or higher in 2015	2.0141	0.030**	6.3511	0.103	0.0168	0.959	-0.7532	0.540
Percent of population with income below poverty 2015	0.3559	0.641	-2.5463	0.269	0.0787	0.675	-0.1798	0.592
Log of amount of loan received in 2010 - 2014	0.0208	0.015**	0.0188	0.132	0.0021	0.483	0.0064	0.192
Rate 2015	30.0396	0.000***	23.5749	0.175	-3.9119	0.354	-5.2068	0.548
constant	11.6638	0.000 **	16.6354	0.109	0.2238	0.000***	0.1596	0.001***
R-Squared	0.3619		0.7411		0.1206		0.3359	
L1	0.9853		0.9264		0.9853		0.9264	

*Note: * indicates the following *** = p<.01, ** = p<.05, * p=<.1*
L1 is the measure of imbalance of covariates between treatment and control groups (smaller L1 implies better balance)

In order to check the robustness of the second time period results, the demographic characteristics were changed to the previous time period (2010) instead of the same time period (2015). Again, the same techniques and approaches used above in the "same year" regression models are used here in the "different year" regression models. The results of the "different year" regression models for time period two are shown in Appendices C and D. These results are very similar to the "same year" regression models that are in Tables 10 and 11, lending confidence that the results are robust. The loan variable is significant in both the basic OLS and first-differenced models *without* the CEM specification, and significant in the basic OLS *with* the CEM specification when the dependent variable is TRS (Appendix C). Switching to TRS per capita as the dependent variable, the loan variable is statistically significant only in the basic OLS models (Appendix D).

The results for time period two shows that the main variable of interest, the log of the amount of loan received variable, is significant in three of the eight models in the "same year" models, and it is significant in five of the eight "different year" models (Table 10, and 11, and Appendices C and D). Notably, none of the CEM-reduced specifications show any significance of the loan variable in the "same year" models. These results do not support the hypothesis – that B&I loan recipients have higher TRS – as much as the results in time period one.

PSM was also used to check the robustness of the CEM results. The logit models underlying the PSM approach are detailed in Appendix G (2005-2010) and Appendix H (2010-2015). They generally behave as expected, with positive and significant results for total retail businesses in period one (Appendix G), and positive and significant results for total retail businesses and the log of median household income in period two (Appendix H). The p-values of the five different PSM methods used and whether these p-values support or contradict the CEM p-values are displayed in Tables 12 and 13. The "S" and "C" in Tables 12 and 13 denote whether the PSM results support (S) the results of the CEM models or that they contradict (C) the CEM results[15]. The PSM results in Table 12 are uniform for all matching methods. The five PSM methods support the CEM results for three of the four dependent variables, and only contradicts when the dependent variable is the differenced TRS per capita from 2005 to 2010 (Table 12). Here, the PSM suggests no statistical difference between the treated and control groups, but the CEM results (Table 8) suggests otherwise – but only at the 10 percent level. Thus the PSM results largely match those for CEM in the first time period of analysis.

Table 12

Pseudo R2	0.1518

P-value of PSM methods for time period 1

type of matching	Nearest neighbor		Kernel (default)		Kernel (normal)		Kernel (biweight)		Radius Caliper(.01)	
	treated	supp/ contra	treated	supp/ contra	treated	supp/ contra	treated	supp/ contra	treated	supp/ contra
TRS 2010	0.056	**S**	0.020	**S**	0.008	**S**	0.028	**S**	0.016	**S**
TRS per cap 2010	0.047	**S**	0.005	**S**	0.001	**S**	0.010	**S**	0.007	**S**
TRS 0510 diff	0.008	**S**	0.012	**S**	0.012	**S**	0.014	**S**	0.001	**S**
TRS per cap 0510 diff	0.556	C	0.290	C	0.407	C	0.281	C	0.100	C

CEM rejects the null of no B&I impact in time period 1
S – supports the CEM findings for time period 1
C – contradicts the CEM findings for time period 1

The PSM results vary for time period two (Table 13). The nearest neighbor, kernel default, and kernel biweight matching methods only supports the "same year" regressions with the CEM specification when the dependent variables are the differenced 2015 TRS and differenced 2015 TRS per capita (Table 13). This shows that PSM rejected the null of no differences across treated/control groups more often than the CEM specification for this period.

[15] Recall that the CEM results found an impact of the B&I loans in period 1 (at the 10% level), but not period 2.

The kernel normal specification only supports the CEM results when TRS per capita is the dependent variable. However, the radius caliper (0.01) supports the CEM results for three of the four dependent variables. Radius matching only contradicts when TRS per capita 2015 is the dependent variable. Overall, the results of the PSM approach mostly supports the CEM specified "different year" regression models for both time periods, but are more contradictory in time period 2.

Table 13

Pseudo R2 0.1066

P-value of PSM methods for time period 2

type of matching	Nearest neighbor		Kernel (default)		Kernel (normal)		Kernel (biweight)		Radius Caliper(.01)	
	treated	supp/ contra	treated	supp/ contra	treated	supp/ contra	treated	supp/ contra	treated	supp/ contra
TRS 2015	0.083	C	0.057	C	0.003	C	0.079	C	0.181	**S**
TRS per cap 2015	0.011	C	0.012	C	0.001	C	0.017	C	0.026	C
TRS 1015 diff	0.224	**S**	0.227	**S**	0.037	C	0.259	**S**	0.251	**S**
TRS per cap 1015 diff	0.764	**S**	0.960	**S**	0.712	**S**	0.984	**S**	0.639	**S**

CEM <u>fails to reject</u> the null of no B&I impact in time period 2
S – supports the CEM findings for time period 2
C – contradicts the CEM findings for time period 2

Regression models for the entire time period (2015 TRS and TRS per capita as the dependent variables with 2005 control variables and the natural log of loan amounts for places between 2005 and 2014) are depicted in Appendices E and F. The loan variable is statistically significant, at the five percent level or lower, in seven of the eight models for the entire time period. These results are similar results to the "same year" models that had seven of the eight regression models showing significance of the loan variable (Tables 8 and 9). However, one issue with running regression models over the whole time period (2005-2015) is that it masks the differing impacts over the years.

The results from the regression models depict mixed results of the impact that the B&I program has on TRS and TRS per capita in Oklahoma. This could be a result of multiple external factors affecting TRS in Oklahoma communities. During time period one, the Great Recession was taking place in the United States, affecting Oklahoma's (and other state's) economy. As previously mentioned, some of the impact of the Great Recession can be seen with the decrease in TRS per capita for most size categories from 2009 to 2010 (Figure 3). However, during period one the regression coefficient for the loan amount variable was statistically significant across all but one model specification, which suggests that the B&I program and other similar governmental programs may be important to fund during tough economic times. During time period two the results show little to no significant impact of the B&I loans, which could be a result of external factors that are not accounted for in the models. For example, one important industry in the Oklahoma economy is the oil and gas industry, which saw a major boom during the 2010-2015 time period, and likely contributed to local tax collections (Whitacre, et al., 2019).

CHAPTER V

CONCLUSION

The goal of this research is to evaluate the impact of the B&I loan guarantee program on TRS for Oklahoma communities in two distinct time periods. It is important to note that the dependent variable (TRS) is only one possible outcome measure for the B&I program. In fact, it is not even listed in the goals of RDs mission and purpose statement, and as such this evaluation is only a partial component of what the impacts of the B&I program could be. Furthermore, this research aimed to fill a gap in the tax revenue literature by using tax revenue as the dependent variable in the analysis.

In the analysis, the demographic characteristics are controlled for in the OLS models, and then coarsened/pruned in the dataset in order to better match the treated and control groups. Mixed evidence was found in support of the B&I loan guarantee program in the analysis. The key variable, the log of the loan amount a place received, was positively and statistically significant in all but one model specification for time period one (2005-2010), but it was only significant in the second time period (2010-2015) for three model specifications. None of the more robust CEM models found a statistically significant impact of loans in this later time period. One takeaway from this research is the dramatic differences in results between periods of economic downturn (2005-2010) versus a more robust economy (2010-2015). This shows that funding the B&I loan guarantee and other similar programs may be more vital during tough economic times when tax revenue is at risk for rural communities.

These findings should be interpreted with caution because of the limitations in the study and data. One of the limitations of the data is we do not know if all loans were successful or if some loans went into default. If the loans did go into default, it could negatively impact the TRS and TRS per capita of a community because a business is likely closing. The CEM methodology used in this analysis is a powerful tool, and it works best when the imbalance of the dataset is reduced after pruning the data. The robustness checks generally support the findings from the CEM methodology. The two time periods see only a slight reduction of the imbalance of the whole dataset for both time periods, but the individual variables imbalance is reducedSecondly, the number of towns in the pruned dataset is a relatively small dataset, which could also impact the regression model with the CEM weights. Finally, this data is specific to Oklahoma and may not be representative of loan and tax collection scenarios in other states.

REFERENCES

Alm, J. 1996. "What is an 'Optimal' Tax System?" National Tax Journal 49(1): 117-133.

Bartik, Timothy J., and Richard D. Bingham. "Can Economic Development Programs Be Evaluated?" W.E. Upjohn Institute, January 1, 1995. https://doi.org/10.17848/wp95-29.

Berney, R., and A. Larson. 1968. "The Adequacy of State and Local Taxes and Business Location." The Annals of Regional Science 2(1): 103-113

Blackwell, Matthew, Stefano Iacus, Gary King, and Giuseppe Porro. "CEM: Coarsened Exact Matching in Stata." *The Stata Journal: Promoting Communications on Statistics and Stata* 9, no. 4 (December 2009): 524–46. https://doi.org/10.1177/1536867X0900900402.

Caliendo, Marco, and Sabine Kopeinig. "Some Practical Guidance for the Implementation of Propensity Score Matching." *Journal of Economic Surveys* 22, no. 1 (2008): 31–72. https://doi.org/10.1111/j.1467-6419.2007.00527.x.

Chrisman, James J. "The Economic Impact of Small Business Development Center Counseling Activities in the United States: 2015-2016," September 2017, 1–31. https://americassbdc.org/wp-content/uploads/2018/01/Chrisman-2017.pdf

Chrisman, James J., and Frances Katrishen. "The Economic Impact of Small Business Development Center Counseling Activities in the United States: 1990–1991." *Journal of Business Venturing* 9, no. 4 (July 1994): 271–80. https://doi.org/10.1016/0883-9026(94)90008-6.

Conley, Kelsey, Brian Whitacre. "Evaluation of the Community Development Block Grant for Rural Oklahoma Communities" Unpublished thesis Oklahoma State University, 2019, 1-20.

Cowan, Tadlock. "An Overview of USDA Rural Development Programs." *Congressional Research Service*, no. RL31837 (February 2016): 40.

Craig, Ben R., William E. Jackson, and James B. Thomson. "Small Firm Finance, Credit Rationing, and the Impact of SBA-Guaranteed Lending on Local Economic Growth." *Journal of Small Business Management* 45, no. 1 (January 2007): 116–32. https://doi.org/10.1111/j.1540-627X.2007.00202.x.

Cromartie, John. "Rural America at a Glance, 2018 Edition," November 2018, 1-6. https://www.obpa.usda.gov/FY19explan_notes.html

Cromartie, John, and Shawn Bucholtz. "USDA ERS - Defining the 'Rural' in Rural America," June 2008. https://www.ers.usda.gov/amber-waves/2008/june/defining-the-rural-in-rural-america/.

Cromartie, John, and Timothy Parker. "USDA ERS - What Is Rural?," April 2019. https://www.ers.usda.gov/topics/rural-economy-population/rural-classifications/what-is-rural/.

Dehejia, Rajeev H, and Sadek Wahba. "Propensity Score-Matching Methods for Nonexperimental Causal Studies." *The Review of Economics and Statistics* 84, no. 1 (February 2002): 151–61.

Ensign, R. and Jones, C. (2019). "The Problem for Small-Town Banks: People Want High Tech Services." The Wall Street Journal March 2, 2019.

Evans, Paul, and Georgios Karras. "Are Government Activities Productive? Evidence from a Panel of U.S. States." *The Review of Economics and Statistics* LXXVI, no. 1 (February 1994): 1–11.

Felix, Alison, and Jason Henderson. "Rural America's Fiscal Challenge." *The Main Street Economist: Regional and Rural Analysis*, no. 3 (2010): 1–7.

Fox, W. 1986. "Personal Income Tax as a Component of State Tax Structure." The Vanderbilt Law Review 39(4): 1081-1096.

Hansen, Tonya J., and Laura Kalambokidis. "How Are Businesses Responding to Minnesota's Tax-Free Zone Program?" *Economic Development Quarterly* 24, no. 2 (May 2010): 180–92. https://doi.org/10.1177/0891242409359407.

Holcombe, Randall G., and Russell S. Sobel. "The relative variability of state income and sales taxes over the revenue cycle." *Atlantic Economic Journal* 23, no. 2 (1995): 97-112.

Holtz-Eakin, Douglas, and Amy Ellen Schwartz. "Spatial Productivity Spillovers from Public Infrastructure: Evidence from State Highways." *International Tax and Public Finance* 2, no. 3 (1995): 10.

Iacus, Stefano M., Gary King, and Giuseppe Porro. "Causal Inference without Balance Checking: Coarsened Exact Matching." *Political Analysis* 20, no. 1 (2012): 1–24. https://doi.org/10.1093/pan/mpr013.

Iacus, Stefano M., Gary King, and Giuseppe Porro. "Multivariate Matching Methods That Are Monotonic Imbalance Bounding." *Journal of the American Statistical Association* 106, no. 493 (March 2011): 345–61. https://doi.org/10.1198/jasa.2011.tm09599.

Janeski, Ivica, and Brian E. Whitacre. “Long-Term Economic Impacts of USDA Water and Sewer Infrastructure Investments in Oklahoma.” *Journal of Agricultural and Applied Economics* 46, no. 1 (February 2014): 21–39. https://doi.org/10.1017/S1074070800000614.

Johnson, Janna. “Rural Economic Development in the United States: An Evaluation of the U.S. Department of Agriculture’s Business and Industry Guaranteed Loan Program.” *Economic Development Quarterly* 23, no. 3 (August 2009): 229–41. https://doi.org/10.1177/0891242408331026.

Lilliston, Ben. “The State of the USDA: A Quiet Dismantling,” February 2019. https://www.iatp.org/blog/201902/state-usda-quiet-dismantling.

Mikesell, J. L. 1984. “The Cyclical Sensitivity of State and Local Taxes.” Public Budgeting & Finance, 4: 32-39.

Missouri Census Data Center. “Geocorr 2014 - MCDC.” MCDC Data Applications, September 10, 2016. http://mcdc.missouri.edu/applications/geocorr2014.html.

Nelson, R., B. Howe, and P. Dean. 2007. “State Revenue Adequacy and Optimal State Portfolios of Sales Tax Components.” Proceedings of the Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association. Vol 100: 479-485.

Oehler, Kelley, and David Lewis. “Rural-Lenders-Roundtable-Presentation.Pdf,” December 2017. http://colemanreport.com/wp-content/uploads/2017/12/Kelley-Oehler-Rural-Lenders-Roundtable-Presentation.pdf.

Office of Advocacy. “Frequently Asked Questions About Small Business, 2018,” August 2018. https://www.sba.gov/sites/default/files/advocacy/Frequently-Asked-Questions-Small-Business-2018.pdf.

Office of Budget and Program Analysis. “Business and Cooperative Programs.” *2020 USDA Budget Explanatory Notes for Committee on Appropriations*, 2019, 1-61, https://www.obpa.usda.gov/30rbcs2020notes.pdf.

Office of the Comptroller of the Currency. “USDA Rural Development Business and Industry Guaranteed Loans,” September 2018, 4.

Ratcliffe, Michael, Charlynn Burd, Kelly Holder, and Alison Fields. “Defining Rural at the U.S. Census Bureau,” December 2016, 1-8.

Ratcliffe, Michael. “Understanding ‘Place’ in Census Bureau Data Products,” n.d., 23.

Rupasingha, Anil, Daniel Crown, and John Pender. “Rural Business Programs and Business Performance: The Impact of the USDA’s Business and Industry (B&I) Guaranteed Loan Program.” *Journal of Regional Science*, December 13, 2018, jors.12421. https://doi.org/10.1111/jors.12421.

Rupasingha, Anil, John Pender, and Seth Wiggins. "USDA's Value-Added Producer Grant Program and Its Effect on Business Survival and Growth." *Economic Research Report*, no. 248 (May 2018): 38.

Rupasingha, Anil, and Kyungsoon Wang. "Access to Capital and Small Business Growth: Evidence from CRA Loans Data." *The Annals of Regional Science* 59, no. 1 (July 2017): 15–41. https://doi.org/10.1007/s00168-017-0814-9.

Rural Development. "RD-Fact Sheet-RBS-BI Guarantee," July 2017. https://www.rd.usda.gov/files/RD-FS-RBS-BIGuarantee.pdf.

Sobel, R. and G. Wagner. 2003. "Cyclical Variability in State Government Revenue: Can Tax Reform Reduce It?" Tax Analysts State Tax Notes 29(8), pp. 569-76

Tax Policy Center. "What Are the Sources of Revenue for Local Governments?," N.D. https://www.taxpolicycenter.org/briefing-book/what-are-sources-revenue-local-governments.

Tetreault, Tricia. "SBA Lending Statistics State-by-State 2019." Fit Small Business, May 28, 2019. https://fitsmallbusiness.com/sba-statistics-lending/.

US Census Bureau. "What Is Rural America?" The United States Census Bureau, August 9, 2017. https://www.census.gov/library/stories/2017/08/rural-america.html.

U.S. Census Bureau. "Glossary - Census Reporter." N.D. Accessed November 11, 2019. https://censusreporter.org/glossary/.

"What Are the Sources of Revenue for Local Governments?" Tax Policy Center. Accessed July 30, 2019. https://www.taxpolicycenter.org/briefing-book/what-are-sources-revenue-local-governments.

Velazquez, Nydia. "The FY 2018 Presidential Budget: A Broken Foundation for Small Business," June 2018. https://smallbusiness.house.gov/sites/democrats.smallbusiness.house.gov/files/sbc%20fy2018%20budget%20trump%20(final%20version%202).pdf.

Voigt, Kevin, and Caren Weiner Campbell. "1 in 6 SBA Small-Business Loans Fail, Study Finds." NerdWallet, October 3, 2017. https://www.nerdwallet.com/blog/small-business/study-1-in-6-sba-small-business-administration-loans-fail/.

Whitacre, Brian E., David Shideler, and Randi Williams. "Do Incentive Programs Cause Growth? The Case of the Oklahoma Quality Jobs Program and Community-Level Economic Growth." *Economic Development Quarterly* 30, no. 1 (February 2016): 62–74. https://doi.org/10.1177/0891242415614641.

Whitacre, Brian E., Dylan L. Johnston, David W. Shideler, Notie H. Lansford, "The influence of oil and natural gas employment on local retail spending: evidence from Oklahoma panel data." *Annals of Regional Science* (2019) https://doi.org/10.1007/s00168-019-00962-7

Wiles, Brian, program director for Rural Business-Cooperative Services in Oklahoma, personal communication, 2019.

APPENDICES

APPENDIX A

Pre and post CEM regression with *Log of TRS 2010* as dependent variable, 2005 control variables

	Basic OLS		Basic OLS with CEM weights		first difference OLS		first difference with CEM weights	
Number of Observations	301		57		262		57	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Log of 2005 population	1.2422	0.000***	0.8920	0.000***	6.6908	0.000***	15.6097	0.003***
Log of Median Household Income 2005	-0.4522	0.076*	-0.2307	0.651	1.0156	0.511	-1.9480	0.604
Total retail business 2005	0.0024	0.107	0.0134	0.001***	-0.0751	0.001***	-0.0895	0.060*
Percent of population black 2005	-1.4644	0.000***	2.3924	0.070**	-1.9595	0.827	-35.8837	0.145
Percent of population with bachelors or higher in 2005	2.5220	0.005***	8.3316	0.002***	-4.4003	0.449	9.9577	0.467
Percent of population with income below poverty 2005	0.1064	0.890	1.7489	0.192	3.8917	0.226	11.2361	0.107
Log of amount of loan received in 2005 – 2009	0.0168	0.015**	0.0214	0.003***	0.0975	0.000***	0.0775	0.014**
Rate 2005	19.4154	0.000***	1.7748	0.889	-35.0834	0.157	-77.2299	0.187
constant	10.7196	0.000***	10.5309	0.062*	14.0816	0.000***	14.0224	0.000***
R-Squared	0.8836		0.9375		0.2116		0.4132	
L1	0.9962		0.7563		0.9962		0.7563	

*Note: * indicates the following *** = p<.01, ** = p<.05, * p=<.1*

L1 is the measure of imbalance of covariates between treatment and control groups (smaller L1 implies better balance)

APPENDIX B

Pre and post CEM regression with *Log of TRS per capita 2010* as dependent variable, 2005 control variables

	Basic OLS		Basic OLS with CEM weights		first difference OLS		first difference with CEM weights	
Number of Observations	301		57		303		57	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Log of 2005 population	0.2274	0.000***	-0.1019	0.326	-0.2204	0.206	-0.1304	0.793
Log of Median Household Income 2005	-0.6008	0.017**	-0.3509	0.471	0.2004	0.278	0.3547	0.316
Total retail business 2005	0.0026	0.075*	0.0129	0.001***	0.0016	0.577	0.0020	0.667
Percent of population black 2005	-1.4525	0.000***	2.3281	0.064	0.4336	0.694	-0.5858	0.815
Percent of population with bachelors or higher in 2005	2.6941	0.002***	8.6585	0.001***	-0.8924	0.209	0.1727	0.896
Percent of population with income below poverty 2005	-0.0180	0.981	1.7765	0.165	0.1430	0.708	1.5349	0.034*
Log of amount of loan received in 2005 – 2009	0.0172	0.012**	0.0213	0.002***	0.0013	0.588	0.0022	0.487
Rate 2005	20.4660	0.000***	3.9846	0.742	0.0041	0.999	-2.5057	0.678
constant	12.3137	0.000***	11.6196	0.032**	0.1246	0.000***	0.0906	0.078*
R-Squared	0.4050		0.6630		0.0161		0.1136	
L1	0.9962		0.7563		0.9962		0.7563	

*Note: * indicates the following *** = p<.01, ** = p<.05, * p=<.1*

L1 is the measure of imbalance of covariates between treatment and control groups (smaller L1 implies better balance)

APPENDIX C

Pre and post CEM regression with *Log of TRS 2015* as dependent variable, 2010 control variables

	Basic OLS		Basic OLS with CEM weights		first difference OLS		first difference with CEM weights	
Number of Observations	306		49		269		46	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Log of 2010 population	1.1512	0.000***	0.8494	0.000***	2.0076	0.037**	5.2382	0.040**
Log of Median Household Income 2010	-0.2067	0.312	0.3229	0.490	-1.0821	0.074*	1.4921	0.284
Total retail business 2010	0.0050	0.004***	0.0204	0.019**	0.0217	0.390	0.0771	0.435
Percent of population black 2010	-1.2769	0.004***	4.7553	0.047**	0.5445	0.905	10.3136	0.500
Percent of population with bachelors or higher in 2010	1.4060	0.079*	2.6759	0.341	2.4240	0.309	0.1639	0.981
Percent of population with income below poverty 2010	0.1027	0.856	2.0114	0.116	-0.6086	0.691	-1.0290	0.774
Log of amount of loan received in 2010 - 2014	0.0172	0.035**	0.0245	0.022**	0.0549	0.010**	0.0007	0.984
Rate 2010	26.2925	0.000***	14.3020	0.367	-56.4361	0.070	21.0293	0.798
constant	8.8539	0.000***	5.2044	0.296	15.1294	0.000***	15.0274	0.000***
R-Squared	0.8517		0.9149		0.0784		0.2263	
L1	0.9927		0.8926		0.9927		0.8926	

*Note: * indicates the following *** = p<.01, ** = p<.05, * p=<.1*
L1 is the measure of imbalance of covariates between treatment and control groups (smaller L1 implies better balance)

APPENDIX D

Pre and post CEM regression with *Log of TRS per capita 2015* as dependent variable, 2010 control variables

	Basic OLS		Basic OLS with CEM weights		first difference OLS		first difference with CEM weights	
Number of Observations	306		49		302		49	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Log of 2010 population	0.1390	0.017**	-0.2236	0.097*	-0.7277	0.000***	-1.1215	0.002***
Log of Median Household Income 2010	-0.3167	0.122	0.3047	0.525	-0.0100	0.899	-0.0178	0.924
Total retail business 2010	0.0053	0.002***	0.0235	0.009***	0.0044	0.212	0.0045	0.736
Percent of population black 2010	-1.0893	0.015**	3.6645	0.132	-0.3687	0.528	2.7892	0.185
Percent of population with bachelors or higher in 2010	1.0155	0.205	2.5213	0.381	0.0168	0.959	-1.2593	0.194
Percent of population with income below poverty 2010	-0.0552	0.923	2.3489	0.075*	0.0787	0.675	0.3826	0.431
Log of amount of loan received in 2010 - 2014	0.0161	0.049**	0.0209	0.054*	0.0021	0.483	0.0008	0.859
Rate 2010	26.5085	0.000***	12.8037	0.430	-3.9119	0.354	12.9068	0.260
constant	10.1415	0.000***	5.9418	0.245	0.2238	0.000***	0.2858	0.000***
R-Squared	0.4050		0.5213		0.1206		0.3595	
L1	0.9927		0.8926		0.9927		0.8926	

*Note: * indicates the following *** = p<.01, ** = p<.05, * p=<.1*
L1 is the measure of imbalance of covariates between treatment and control groups (smaller L1 implies better balance)

APPENDIX E

Pre and post CEM regression with *Log of TRS 2015* as dependent variable, 2005 control variables

	Basic OLS		Basic OLS with CEM weights		first difference OLS		first difference with CEM weights	
Number of Observations	300		83		278		75	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Log of 2005 population	1.1681	0.000***	0.8960	0.000***	3.1036	0.000***	3.5328	0.006***
Log of Median Household Income 2005	-0.2935	0.291	-0.4644	0.330	-0.1580	0.807	0.0578	0.958
Total retail business 2005	0.0030	0.065*	0.0134	0.001***	-0.0037	0.791	-0.0729	0.005***
Percent of population black 2005	-1.4386	0.001***	3.4536	0.031**	-2.4440	0.514	-8.7628	0.243
Percent of population with bachelors or higher in 2005	2.5582	0.009***	8.6460	0.001***	2.1484	0.406	0.8417	0.848
Percent of population with income below poverty 2005	-0.3328	0.694	0.1467	0.926	2.2680	0.183	1.0392	0.768
Log of amount of loan received in 2005 – 2014	0.0222	0.001***	0.0260	0.001***	0.0990	0.000***	0.0513	0.018**
Rate 2005	23.3443	0.000***	6.9448	0.563	-8.6918	0.620	47.1363	0.090*
constant	9.7779	0.001***	13.1910	0.012**	15.1737	0.000***	14.9775	0.000***
R-Squared	0.8612		0.8862		0.2169		0.3093	
L1	0.9919		0.8653		0.9919		0.8653	

*Note: * indicates the following *** = p<.01, ** = p<.05, * p=<.1*
L1 is the measure of imbalance of covariates between treatment and control groups (smaller L1 implies better balance)

APPENDIX F

Pre and post CEM regression with *Log of TRS per capita 2015* as dependent variable, 2005 control variables

	Basic OLS		Basic OLS with CEM weights		first difference OLS		first difference with CEM weights	
Number of Observations	300		83		299		82	
	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
Log of 2005 population	0.1457	0.012**	-0.1391	0.172	-0.2799	0.011**	-0.2999	0.205
Log of Median Household Income 2005	-0.6127	0.027**	-0.6491	0.177	0.1607	0.154	0.2484	0.223
Total retail business 2005	0.0035	0.028**	0.0147	0.000***	0.0051	0.044**	-0.0007	0.881
Percent of population black 2005	-1.2443	0.004***	3.3855	0.035**	-1.1430	0.069*	-0.4128	0.768
Percent of population with bachelors or higher in 2005	2.3938	0.013**	7.6459	0.002***	0.3400	0.455	0.3042	0.710
Percent of population with income below poverty 2005	-0.7174	0.393	-0.0197	0.990	0.2786	0.344	0.3460	0.595
Log of amount of loan received in 2005 – 2014	0.0208	0.001***	0.0227	0.003***	0.0029	0.287	0.0110	0.008***
Rate 2005	24.6529	0.000***	8.6185	0.475	3.6123	0.245	12.4956	0.019**
constant	13.2497	0.000***	15.4027	0.004***	0.3049	0.000***	0.1672	0.028**
R-Squared	0.3410		0.4891		0.0543		0.1448	
L1	0.9919		0.8653		0.9919		0.8653	

*Note: * indicates the following *** = p<.01, ** = p<.05, * p=<.1*
L1 is the measure of imbalance of covariates between treatment and control groups (smaller L1 implies better balance)

APPENDIX G

Logistic regression results for time period one (2005-2010)

Dependent variable is a town receiving a loan 2005 – 2010

Logistic regression		Number of obs	=	301
		LR chi2(**6**)	=	35.24
		Prob > chi2	=	0.0000
Log likelihood =	-98.435431	Pseudo R2	=	0.1518

rec~20052009	Coef.	Std. Err.	z	P>|z|	95% Conf. Interval	
pop_2005	-.0001704	0.0000815	-2.09	0.037	-.0003301	-.0000107
ln_MHI_2005	1.876499	1.485427	1.26	0.206	-1.034885	4.787883
Tot_Bus_2005	0.0489083	0.0146132	3.35	0.001	0.0202671	0.0775496
pct_bac~2005	-4.28565	5.23159	-0.82	0.413	-14.53938	5.968078
pct_inc~2005	4.705972	4.437336	1.06	0.289	-3.991047	13.40299
Rate_2005	7.093151	33.15147	0.21	0.831	-57.88253	72.06883
_cons	-22.72588	15.83885	-1.43	0.151	-53.76947	8.3177

Note: the common support option has been selected
The region of common support is [.04937083, .95023095]
Description of the estimated propensity score in region of common support

Note*

Population 2005 was used instead of the natural log of 2005 population because the balancing property was not satisfied when natural log of 2005 population was used.

APPENDIX H

Logistic regression results for time period two (2010-2015)

Dependent variable is a town receiving a loan between 2010 – 2015

Logistic regression				Number of obs	=	307
				LR chi2(**6**)	=	21.88
				Prob > chi2	=	0.0013
Log likelihood =	-91.685074			Pseudo R2	=	0.1066

rec~20102015	Coef.	Std. Err.	z	P>|z|	95% Conf. Interval	
pop_2010	-.0001395	0.0000762	-1.83	0.067	-.0002888	9.91e-06
ln_MHI_2010	2.114062	1.125521	1.88	0.060	-.0919188	4.320042
Tot_Bus_2010	0.0395799	0.014109	2.81	0.005	0.0119267	0.067233
pct_bac~2010	-4.441529	4.394282	-1.01	0.312	-13.05416	4.171106
pct_inc~2010	3.522476	3.262726	1.08	0.280	-2.872351	9.917302
Rate_2010	48.14733	35.22121	1.37	0.172	-20.88497	117.1796
_cons	-26.67018	12.11538	-2.20	0.028	-50.41589	-2.924468

Note: the common support option has been selected
The region of common support is [.04651842, .64295208]
Description of the estimated propensity score in region of common support

Note*

Population 2010 was used instead of the natural log of 2010 population because the balancing property was not satisfied when natural log of 2010 population was used.

APPENDIX I

USDA Rural Business-Cooperative Service Programs ***Enacted Budget*** Level of Funding
(in millions of dollars)

Item	2012	2013	2014	2015	2016	2017	2018
Appropriate Tech. Transfer to Rural Areas	$2	$2	$2	$2	$2	$3	$3
Bioenergy for Advanced Biofuels	$65	-	$15	$14	$15	$14	$14
Biorefinery Assistance Guaranteed Loans	$0	-	$241	$71	$208	-	-
Business and Industry Guaranteed Loans	**$811**	**$890**	**$958**	**$920**	**$920**	**$922**	**$920**
Intermediary Relending Program	$18	$18	$19	$19	$19	$19	$19
Rural Business Development Grants	$26	$25	$27	$24	$24	$24	$34
Rural Cooperative Development Grants	$9	$6	$6	$6	$6	$6	$6
Rural Economic Development: Loans and Grants	$43	$43	$43	$43	$45	$52	$55
Rural Energy for America (Sec. 9007) Loans and Grants	$61	$13	$129	$88	$200	$342	$342
Rural Microentrepreneur Assistance Program (sec. 6022): Loans and Grants	-	$0	$33	$8	$8	$8	$8
Small Socially Disadvantaged Producer Grants	-	$3	$3	$3	$3	$3	$3
Value-Added Producer Grants	$14	$14	$15	$11	$11	$15	$16

a/ - subject to reauthorization

These are all proposed budget amounts and not the estimated or enacted amounts

Source: USDA 2012 -2020 budget summary proposals

APPENDIX J

Univariate imbalance before and after CEM for 2010 **same year models** (2010 TRS and control variables)

	Before CEM	CEM specification		
			0	1
		All	301	39
		Matched	48	19
		Unmatched	253	20
Multivariate L1 distance:	0.9963		0.9020	
Univariate imbalance:				
	L1		L1	
pop_2010	0.40078		0.2105	
ln_MHI_2010	0.27057		0.2363	
Tot_Bus_2010	0.41715		0.0322	
pct_bach_2010	0.18819		0.1605	
pct_income below poverty_2010	0.18398		0.1301	
Rate_2010	0.18829		0.0263	

APPENDIX K

Univariate imbalance before and after CEM for 2015 **same year models** (2010 TRS and control variables)

	Before CEM	CEM specification		
			0	1
		All	310	33
		Matched	34	11
		Unmatched	276	22
Multivariate L1 distance:	0.9853		0.9264	
Univariate imbalance:				
	L1		L1	
pop_2015	0.31066		0.23972	
ln_MHI_2015	0.21507		0.04545	
Tot_Bus_2015	0.35294		0.36201	
pct_bach_2015	0.25551		0.19426	
pct_income below poverty_2015	0.15625		0.10606	
Rate_2015	0.0864		5.6e-17	

VITA

Ty Rope Smith

Candidate for the Degree of

Master of Science

Thesis: A DIFFERENT ANALYSIS OF RURAL DEVELOPMENT'S BUSINESS AND INDUSTRY LOAN GUARANTEE PROGRAM: THE IMPACT ON TAX REVENUE IN OKLAHOMA COMMUNITIES

Major Field: Agricultural Economics

Biographical:

Education:

Completed the requirements for the Master of Science in Agricultural Economics at Oklahoma State University, Stillwater, Oklahoma in May, 2020.

Completed the requirements for the Bachelor of Science in Agricultural Economics at Oklahoma State University, Stillwater, Oklahoma in 2018.

Experience:

Graduate Research Assistant
Department of Agricultural Economics, Oklahoma State University, 2018-2020

Professional Memberships:

Southern Agricultural Economics Association