A GUIDE TO THE AMERICAN COMMUNITY SURVEY (ACS) FOR THE RURAL RESEARCHER: UNPACKING THE CONCEPTUAL AND TECHNICAL ASPECTS OF USING ACS FOR SCHOOL-COMMUNITY RESEARCH

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Kristie Nicole LeBeau

December 2020

ABSTRACT

This paper explores the relationship between the quality of a school (or lack thereof) and the economic vitality of rural communities in Indiana. Indiana's Department of Education (IDOE) has a robust accountability system that provides the proxy for school quality, and the American Community Survey (ACS) provides the economic and demographic data for communities. However, ACS data presents obstacles for rural researchers to navigate due to the small population sizes of rural communities. Therefore, this paper addresses both the conceptual and the technical details of using ACS-type data for rural research. I offer suggestions to assist rural researchers in navigating ACS-type data and bring attention to the *whys* behind the results, pushing researchers to embrace these questions so that they might propel future studies that give rural schools and communities a voice in education policy.

BIOGRAPHICAL SKETCH

Kristie LeBeau obtained her master's degree in the field of Development Sociology from the College of Agriculture and Life Sciences at Cornell University. Her research focuses on the nuances of the relationship between rural schools and communities. Kristie earned her undergraduate degree in Sociology from Saint Mary's College in Notre Dame, Indiana, in the Spring of 2018, graduating *Summa Cum Laude* as a valedictorian of her class. Kristie has been a member of Alpha Kappa Delta, the International Sociology Honors Society since the Spring of 2017. Her senior comprehensive titled, "Support a School, Preserve a Community: A Content Analysis of a Newspaper's Portrayal of a Rural School-Community Relationship" served as another contribution to her work in this field and was a finalist in the North Central Sociological Association President's Undergraduate Paper Competition. Kristie aims to continue this work by exploring the multiple roles a school plays in a rural community, focusing on the state of Indiana, her home state, with the goal of aiding rural leaders as their schools and communities face the challenges of changing contexts.

ACKNOWLEDGMENTS

I would like to thank my committee members for their guidance on this project. First and foremost, I would like to thank Professor John Sipple, my chair, for challenging me throughout my graduate work and fueling my passion for working with rural schools. Next, I would like to thank Professor Scott Peters whose passion to improve public engagement in higher education motivates me as a scholar. I would also like to thank Dr. Warren Brown and Jan Vink for their assistance in helping me navigate the American Community Survey when I hit roadblocks. The assistance and guidance of these people allowed this project to come to fruition.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

In discussions surrounding rural communities, the relationship between the school and the community is rarely left out. Rural researchers have illustrated the connection between the two, citing the role of the school as part of the local identity (Tieken, 2014; Autti & Hyry-Beihammer, 2014), the role of school leaders as community leaders (Budge, 2006; Harmon & Schafft, 2009; Zuckerman, 2020), as well as the role the school plays in contributing to the economic vitality of a community (Lyson, 2002; Sipple, Francis, & Fiduccia, 2019). As such, examining schools and communities in relation to each other is not only important but also necessary for rural researchers.

However, the causal direction of that relationship is less clear, and perhaps points in more than one direction. In the rural school-community literature, discussions surrounding school closure and economic downturn often point to the school as a contributor to community vitality (Lyson, 2002). In this way, the causal relationship might be assumed to be in the direction of the school impacting the community vitality. On the other hand, as the scope of education literature is expanded (including urban, suburban, and rural areas) and seen with a different lens, the view of the school-community relationship can be flipped and examined for the way that family and community characteristics impact the presence of a school and the opportunities and achievement of students within schools. Coleman (1966) was the first to shed light on this type of relationship by pointing to the positive connection between a student's family background, the peers with whom she attends school, and student achievement, and since then, many researchers have followed suit. More recently in the era of standardization and accountability, researchers have built on this connection by pointing to the positive linkage between family and community characteristics and standardized test scores (Tienken, 2017). Accordingly, the local effects of a community have the ability to influence how a student performs in a classroom and shape post-graduation opportunities. Therefore, no matter what the direction of the relationship is in any given scenario, the connection between schools and communities points towards the conclusion that the two must be examined together. Casto and colleagues (2016) capitalize on this point and make the call to rural researchers and beyond to expand their thinking around education policy to be more inclusive of the ways that education policy is interlaced with community policy.

In rural research, a common way of doing this is by carrying out rich, qualitative case studies of communities and establishing the relationship between the two through the lens of the people who experience it first-hand (ex. Corbett, 2008; Tieken, 2014; Carr and Kefalas, 2009). These types of studies are owed great credit for contributing to the conceptual underpinning of the nuances of school-community relationships. However, if the goal, as outlined by Casto et al. (2016) is to examine and influence the education policies that impact communities so that they might be able to establish a more "community aware" education policy, then quantitative work is often necessary to discover larger geographic patterns. This can be done by combining different levels of data, pulled from a variety of data repositories (Sipple, Fiduccia, & LeBeau, forthcoming). A useful source of community data for the United States is the American Community Survey (ACS), which continuously samples communities in order to provide data estimates from year to year for the United States Census Bureau (USCB). When it comes to education data, there are multiple national and state level sources that generate data at the school and individual level. However, with these data sources comes obstacles as well, especially for rural researchers.

This is where the purpose of this paper is realized. Initially, I set out to explore the relationship between a rural community's economic vitality and the quality of the school (or lack of) within the community. I still intend to provide a thorough conceptualization of this type of relationship—exploring the nuances of education policy and how it impacts the way school leaders interact with their students and the larger community. However, my experience with this project calls for a more in depth conversation surrounding the technical navigation of this type of data. With a focus on rural communities in Indiana, I turned to ACS data for community level data and the Indiana Department of Education (IDOE) for school level data. Indiana's DOE assigns letter grades ranging from A-F to every public school in the state based on student performance, growth, and college and career preparation. This provides a useful source of school-level data from the state. However, when collecting community-level data from the ACS, I quickly realized that the level of error within the ACS data created an obstacle with my research that was not easily fixed, interfering with the validity of my results. Because ACS data is made up of samples of data that produce estimates for the larger population, the small population size of rural communities make it difficult to collect results that are representative of the population—resulting in a high level of error. Therefore, the purpose of this paper is two-fold—providing a conceptualization of the relationship between education policy and community vitality, in addition to the technical details of how ACS-type data can be utilized and navigated by rural researchers to carry out this type of research. I plan to utilize my own experience using ACS data for rural school-community research as an instructive lesson in the strengths and weaknesses of such research. I aim to lay out the logic as to why ACS-type data is useful for rural researchers while also drawing attention to the obstacles that this data presents for rural researchers so that they may be more fully realized and navigated. I rely on a set of guidelines intended to guide urban planners using ACS data with high levels of error (Jurjevich et al., 2018) and suggest that these guidelines should also be utilized by rural researchers. Moreover, I argue that, albeit its obstacles, ACS data serves as a useful resource for rural researchers and should be utilized with intention so that it may be used as a launch pad for future studies that can delve deeper into the intricate details of *why* certain patterns and trends emerge.

I plan to do this in four parts. First, I will give a background of the nuances of ACS data in an attempt to present its data collection and presentation methods in a manner that can be understood and transferable for rural researchers looking to add ACS data to their research toolkit. Next, I will dive into the project example, first giving a general overview of the purpose of the study followed by a conceptual underpinning of the project. After laying this foundation, I will lay out the methodology and then take a deep dive into the obstacles and intricacies that presented themselves throughout—using them to illustrate ways to approach and overcome these hang-ups that might otherwise discourage someone from using the ACS data. Lastly, I will discuss the implications of this study for rural researchers and educational policymakers, exploring the potential explanations of the results and analyzing how the results from this type of data exploration can be used as a launch pad for future studies examining the interconnectedness of schools and communities in education policy.

How is ACS data collected and what is it used for?

What is the American Community Survey?

The American Community Survey (ACS) continuously collects a comprehensive set of data from a sample of the population. It differs from the Census in a few ways. The Census is conducted every 10 years, collects data on the entire population, and collects a short list of information which mainly provides a count of the population. In comparison, the ACS is conducted every month of every year, collects data from a sample of the population, and collects a more comprehensive set of information about a range of topics (education, occupation, housing, etc.).

These differences allow the ACS to provide up to date information to communities every year, making it an important source of information for local leaders.

According to the Census Bureau (2020a), the ACS "provides local and national leaders with the information they need for programs, economic development, emergency management, and understanding local issues and conditions." On a different level, the ACS and the Census are used to determine how just over $675 billion of federal spending is distributed each year (USCB, 2020a). In its call to citizens to respond to the ACS, the Census Bureau (2020b) argues that by responding, "you are doing your part to help your community plan for hospitals and schools, support school lunch programs, improve emergency services, build bridges, and inform businesses looking to add jobs and expand to new markets, and more." In plain terms, the ACS is tasked with a hefty job to provide up to date, accurate information for localities. After all, this data works to provide leaders with the information necessary to assess the state of their communities. However, it is important to know the complexity of ACS data in order to make sure its consumers utilize and display the information accurately.

What is a Census place?

Before delving into the intricacies of ACS data, it is important to clarify what type of geographic unit I am focusing on for this paper—a Census "place." For the purposes of this paper, I use "place" and "community" interchangeably, but the definition of a Census "place" is important to specify. The Census defines a place as a concentration of a population. This means that not everyone lives in a place. For instance, picture a small town's main street. Those who live on the main street and within the other blocks that make up the town center are most likely within the place; however, those that live out of the town's center—perhaps a few miles out of town—are not counted within the place's population count. To help illustrate this, I have included a map of places in central Indiana (Figure 1). All of the tan outlined figures represent concentrations of populations across the region. Residents inhabit areas outside of these places, and for those residents that do not live within the outlined population center, they are not a part of the population count of the place. It might seem counterintuitive for a rural researcher or local leader to be interested in this geographic unit if not all residents are included. However, the Census place data allows rural localities to zoom in on their population center as opposed to other macro units like the county level or school district level, which often includes other towns.

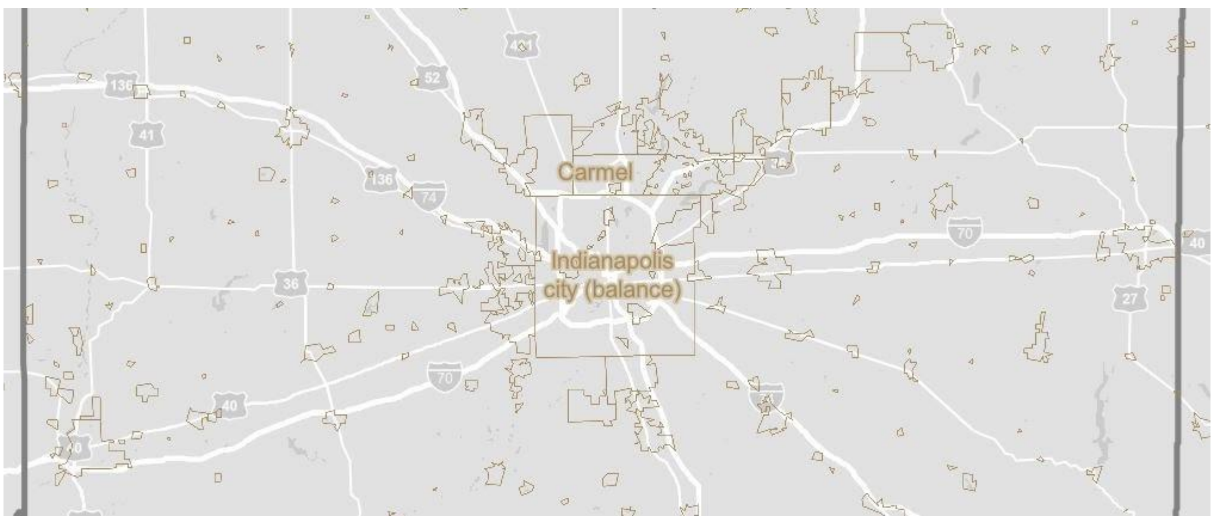

Figure 1. Central Indiana Census Places. Source: American Community Survey. United States Census Bureau.

To break it down further, Census places can either be an incorporated place or a Census Designated Place (CDP). An incorporated place is a "legally bounded entity" (Ratcliffe, n.d.) as defined by the laws of the state, and a majority have an active government (ex. cities, towns, villages). The boundaries of incorporated places are adjusted every year. In comparison, a CDP is a "statistical entity" (Radcliff, n.d.) defined by the Census Bureau, and their boundaries are established every 10 years. CDPs do not have active governments but are described as "settled population centers with a definite residential core, a relatively high population density, and a degree of local identity” (U.S. Department of Commerce, 1994, p. 20). CDPs are not as numerous as incorporated places, but they are important geographic units because they allow localities that may not otherwise be recognized by the Census Bureau to tabulate their population count (U.S. Department of Commerce, 1994). Moreover, despite their differences, incorporated places and CDPs provide useful geographic units for researchers to study rural population centers across the US.

Interpreting ACS data

Because the ACS is collected continuously on a sample of the population, the data is released as estimates. There are 1-year and 5-year estimates—1-year estimates are for populations of 65,000 and up, and 5-year estimates are for all areas (USCB, 2009). It is important to clearly understand how to interpret these estimates, as they are not data from a single point in time, but rather, an average across multiple months (60 months for the 5-year estimates). For example, when describing the total poverty estimate for a place, one must clarify that the estimate is the average total poverty estimate over five years of data.

Data is reported as estimates to increase data reliability for the samples collected by pooling a larger amount of data. However, the estimates also present a certain degree of uncertainty for smaller populations due to their inevitably small sample size. To help understand this issue, the Census Bureau provides users with a “Margin of Error” (MOE) for each statistic (USCB, 2020c). MOE’s are based on a 90 percent confidence level (USCB, 2009). For large cities, this MOE alongside the estimate can provide a useful range of values that is representative of the population. However, for small areas, this MOE can vary so much that the estimate is outlandishly misleading.

Because numbers like population counts and occupation figures are difficult to compare at raw value across different places, researchers often use the coefficient of variation (CV) to determine whether or not the estimate is reliable for one place or another. The CV "provides a measure of the relative amount of sampling error that is associated with a sample estimate" (USCB, 2009, p. A-13). This is figured by calculating the standard error (MOE/1.645) and dividing that by the estimate and multiplying by 100 to get the percentage. A low CV indicates a low amount of sampling error relative to the estimate. Spielmann, Folch, and Nagle (2014) reference the National Research Council (NRC) and Environmental Systems Research Institute (Esri) as two sources of guidelines for the interpretation of coefficients of variation. The NRC (2007) states that 10-12 percent or below is a reasonable CV to accept the estimation as reliable. Esri (2014) specifies this further, citing 12 percent or below as "high reliability," between 12 percent and 40 percent as "medium reliability," and anything above 40 percent as "low reliability" (p. 6). Accordingly, although these are rough rules of thumb, any CV above 12 percent should be used with caution, and anything above 40 percent should raise extreme concern.

This is perhaps best illustrated through an example of a few places in Indiana. Figure 2 displays the total population estimate of three places in Indiana roughly within 10 miles of each other—a city, a town, and a CDP. The ACS 2018 5-year total population estimate of Lafayette City, Indiana is 72,444 with a MOE of +/- 230. The coefficient of variation is 0.19%, indicating that the estimate is reliable according to widely accepted thresholds (NRC, 2007; Esri, 2014). In comparison, Dayton town, which has a population of 1,569 and a MOE of +/- 237 has a coefficient of variation of 9.18%. This CV falls into the category of a reasonably reliable estimate, but the jump in percentage illustrates the difference in certainty for smaller sample sizes. This difference is even more pronounced in Buck Creek CDP's population estimate. With a population of 150 and a MOE of +/- 124, the population estimate has a CV of 50.25%, making it a "very unreliable" estimate (Esri, 2014).

	Buck Creek CDP, Indiana		Dayton town, Indiana		Lafayette city, Indiana	
	Total		Total		Total	
	Estimate	Margin of Error	Estimate	Margin of Error	Estimate	Margin of Error
Total population	150	+/-124	1,569	+/-237	72,444	+/-230

Figure 2. 2018 Total population data. Source: American Community Survey 5-year estimates. United States Census Bureau.

Why is it important for rural researchers to understand the intricacies of ACS MOEs and Census places? As previously mentioned, ACS data gives local leaders the necessary information about their communities to make decisions about anything from school programs to construction to the development of new businesses. For small rural towns, using the Census place data allows those communities to zoom into the intricacies of their individual community rather than using macro level data at the county level or higher. However, rural researchers must be cautious when using this data because of the high margin of error that is often associated with their small sample sizes. After surveying and interviewing urban and regional planners, Jurjevich et al. (2018) found that most planners do not understand the statistical uncertainty of ACS data and have a difficult time knowing how to portray that uncertainty to stakeholders. This statistical uncertainty cannot be ignored by researchers who are dealing with some of the smallest populations and must be tactically navigated so as to do justice to the results for local leaders. Accordingly, Jurjevich and colleagues (2018) lay out a set of guidelines for urban planners to follow in order to better understand and articulate ACS results with MOE findings. Other researchers have attempted to develop technical solutions and methods to address the uncertainty of the data, often focusing on Census tracts as a geographical unit (Spielman & Singleton, 2015; Wei, Tong, & Phillips, 2015). This paper does not provide a technical solution, but rather, I am focused on providing a conceptual understanding of how ACS place data can be navigated and utilized by rural researchers.

Therefore, I argue that rural researchers using ACS to study small populations should apply similar guidelines to those suggested by Jurjevich et al. (2018) for urban planners. Rural researchers may be asking different questions and working with different populations than urban planners, but they are both using ACS data to learn more about specific communities and must understand the intricacies of the data they are working with in order to provide accurate results to their stakeholders. The guidelines are laid out as followed:

1. Report the corresponding MOEs of ACS estimates.
2. Include a footnote when not reporting the corresponding MOEs of ACS estimates.
3. Provide context for the level of (un)reliability in ACS estimates; do not be afraid to advise against using data that are clearly unreliable.
4. Reduce statistical uncertainty by collapsing data detail when possible, aggregating Census geographies, and/or using multiyear ACS estimates.
5. *Always* conduct a test of statistical difference when comparing ACS estimates before calculating or reporting any apparent differences in the estimates (Jurjevich et al., 2018).

Every aspect of these guidelines may not apply directly to specific research endeavors by rural researchers; however, conceptually, they can be applied as a sort of checklist to double check one's data and results. I will illustrate how these guidelines can be utilized in my example of using ACS data to examine the school-community relationship in small places in Indiana. Not every point applies, but using these guidelines, in addition to considering other implications of the results, I aim to illustrate how ACS data can be utilized as an effective source of data for rural researchers.

Project example: Purpose and conceptual framework

Outline and purpose

The purpose of this research endeavor was to study the potential relationship between the quality of a school and the economic well-being of a community. I asked the question: How is the quality of a school, as defined by the Indiana Department of Education (IDOE), related to the income level of its community? As I acknowledged before, a question such as this is difficult to determine the causal relationship. However, educational research showing how school performance is related to family and community background (Coleman, 1966; Hill, 2016; Tienken, 2017) and rural research showing how the presence of or proximity to a school is positively related to higher levels of economic vitality of a community (Lyson, 2002; Sipple, Francis, & Fiduccia, 2019) motivated me to create a study that examines the two together. In a small rural community where the school is a staple community institution, does the quality of the school (as determined by the state) relate to greater economic vitality? And conversely, does having a strong economic base in a community predict the quality of its school?

With this in mind, I set out to combine school and community level data to explore the relationship between school quality and community economic well-being of small rural communities in Indiana. Indiana presents an interesting state for analysis due to its relatively long history of school reform, dating back to 1987, that has led to an exhaustive school accountability grading system (Hiller et al., 2012). In 2012, Indiana ranked number one in the nation for “standards, assessment, and accountability” according to Ed Week’s (2012) *Quality Counts* report card, qualifying it as a useful example for other states. The 2004-2005 school year saw the first school assessment grades under the most recent set of legislation, Public Law 221. Since then, Indiana’s school accountability grading system has undergone a series of changes to land on its current system of grading. This system takes into account individual student performance, individual student growth to proficiency, and multiple measures, which include graduation rates and college and career readiness, and rates schools and districts on an A-F scale (IDOE, n.d.). The A-F grades are not built on a rewards system, but rather, the state intervenes when schools extremely underperform. For traditional public schools, state intervention is determined based on consecutive F grades (IDOE, n.d.). The amount and extent of state intervention changes based on how many consecutive years a school has received an F, but the public is notified through a hearing and various types of interventions are deployed in order to work on the improvement of the school. Further and to reiterate, whether a school receives an A or a D does not warrant intervention from the state—only an F grade receives attention from the state level.

Key changes over the years have revolved around 1) understandability and accessibility of the results—focusing on jargon-free language that is easily understood by students, parents, and communities (Hiller et al., 2012) and 2) measuring multiple aspects of individual growth rather than purely performance and performance in relation to peers (IDOE, n.d.).

On the former point, because one of the goals of the A-F grades is to make the results easily understood by the community and the only consequences from the state are associated with receiving an F, one might infer that the incentive of receiving an A has more to do with being able to boast a high grade as opposed to dodge receiving an F. In an era of school choice, where Indiana is a leading figure in the school voucher movement (Turner, 2017; EdChoice, 2020), school leaders may also feel pressure to perform in order to display a high grade for their school.

On the latter point, the IDOE shaped the new A-F system with the goal of valuing performance in addition to valuing schools that help students "catch up," "keep up," and "move up" (IDOE, n.d.). This focus on growth measures can be inferred as an attempt at moving beyond school performance measures that quantify community wealth as a proxy for student achievement. All too often, states use only test scores as a way to hold schools and school leaders accountable. Illustrating the dangers associated with using only this type of measure, Tienken and colleagues (2017) draw attention to the flaws in this type of model as seen in New Jersey's education system. In their study, Tienken et al. (2017) were able to accurately predict the percentage of middle school students who scored proficient on New Jersey's standardized tests in math and language arts using three family and community demographic variables from the US Census. Those variables included "(a) percentage of families in a community with income over $200,000 a year, (b) percentage of people in a community in poverty, and (c) percentage of people in a community with bachelor's degrees" (Tienken et al., 2017, p. 1). Accordingly, these authors make the argument that state mandated standardized tests to evaluate schools and school administrators are not an objective measure of the school's effect on students, but rather, capture too much of the family and community factors that influence a student's performance in the classroom (Tienken et al., 2017). As such, the Indiana model poses an interesting case to study, as it appears to potentially value the school's ability to help a student progress in the right direction, outside of the external environment that influences their ability to perform. Because the Indiana accountability system measures growth as opposed to purely performance level, is it able to capture school effects outside of the community effects? Further, does having a high quality school in a community contribute to a more vital economy in a small, rural community? These are potential questions I set out to explore with ACS place data and Indiana's school accountability data.

Conceptual underpinning

In order to explore the nuances of the connection between school accountability and the local community, it is important to consider what external forces influence the school climate. According to Arum (2000), discussion around school-community relationships has shifted to not only include the local ecological community which physically surrounds the school but also the larger professional and institutional structure within which schools operate. This viewpoint acknowledges the institutional pressures schools face to perform according to state measures—largely occupying a school leader's bandwidth, especially if the results come with consequences for not performing according to par. Further, if a school leader's mind is occupied with the pressure to perform well according to standardized tests, how much time are they able to focus on the needs of the larger community in which their students reside?

Accordingly, in the era of standardized testing and accountability, providing social supports to the community may not be the focus of school leaders (Jennings et al., 2005; Budge, 2006; Schafft, 2016). However, Harmon and Schafft (2009) argue that "cultivating collaborative and meaningful school-community development will be a hallmark of good public schools that can meet the challenges facing rural communities and their students in the 21st Century" (p. 8). As such, one can see how school leaders, especially in rural areas where they are a major community institution, are potentially pulled in multiple directions to cater to both the needs of the community within which their students reside and the expectations of the state—and when education policy places the focus on performance, the latter often wins.

Although education policy has placed most of its attention on standardization and accountability, the recent COVID-19 pandemic brought many of a school's other roles to the public's attention in a way that they had not been recognized before. When education went virtual in the spring of 2020, and the school was no longer able to provide childcare or meals for its students, the school's role outside of the education it provides its students was more fully realized. Suddenly, the school's role as a community institution that provides social and economic aid for its students was brought to the forefront as students were unable to receive childcare and meals from the place they spend roughly eight hours a day.

Similarly, it would be remiss to discuss the impact of COVID-19 on schools and communities without also discussing the impact of the surge in the Black Lives Matter movement that was fueled by the death of George Floyd in late May 2020 and has brought attention to the pervasive inequity in America. Although rural communities are often made up of a majority white population (which will soon be evidenced in this project), trends across the US show demographic shifts in other populations. Specifically, the Hispanic population is the fastest growing population in rural America, with a growth rate over the past few years of just under 2% per year (Cromartie, 2018). American Indian populations have also seen an increase in number in rural areas, although at a smaller growth rate than Hispanic populations (roughly 0.5% increase in number each year) (Cromartie, 2018). In comparison, Black and non-Hispanic white populations have seen a loss in population size over the past few years (Cromartie, 2018). Accordingly, rural school leaders must acknowledge the changing demographic makeup of their schools and communities and be prepared to address and adapt to these changes so as to create an inclusive environment in their schools and communities and meet the needs of all their students. As such, schools and communities must be examined together, now more than ever, in order to determine how the two can work together to support rural communities and their students in the wake of globalization, out-migration, changing demographics, and economic downturn.

The theoretical underpinning of this argument to conceptualize school and community development as one and to use multiple levels of data to study rural communities is founded in Casto and colleagues' (2016) conception of "community aware" policy. This approach to policy works to move past the silos of educational research and policy that only address the "thin" needs of individuals and institutions to examine the "thick" needs of families and communities (Dean, 2010; Casto et al., 2016). Take changing demographics as an example. If a school notes that the number of students on free/reduced lunch is increasing, they adjust and make sure that all of those students receive food at school. This approach addresses a thin concept of need, because the school recognizes an increase in free/reduced priced students and feeds them. A thicker approach to this level of need would be to take the increase in free/reduced price students and consider how that relates to the families within the community. Does ACS data show that the community is also experiencing an increase in poverty levels? Higher levels of unemployment? If so, how might the school establish other programs that support the whole family rather than just focusing on the individual? Further, how might student performance benefit from addressing community level issues and improving the external environment that students inhabit? In this case, school level data and community level data can be used together in order to give local leaders and researchers a clearer picture of the community as a whole to inform "community aware" education policy.

This was the motivation behind the School-Community Framework created by LeBeau, Casto, and Sipple (in preparation) (Figure 3). The framework was created as a tool for school and local leaders to examine the social, economic, and educational roles a school plays and, specifically, how those play out in an individual community. As such, the framework aims to be an adaptable tool for schools and communities to use to better understand their relationship in order to make well-informed decisions about their unique environment. It is broken into spheres that overlap and represent the level of analysis (school/community) and role of the school (economic/social). These spheres are further divided into four quadrants, representing the roles schools take on: economic force, social force, preparer of citizens, and preparer of workers.[1] The current study examines the school as an economic force, asking what role the quality of a school plays in the economic vitality of a community.

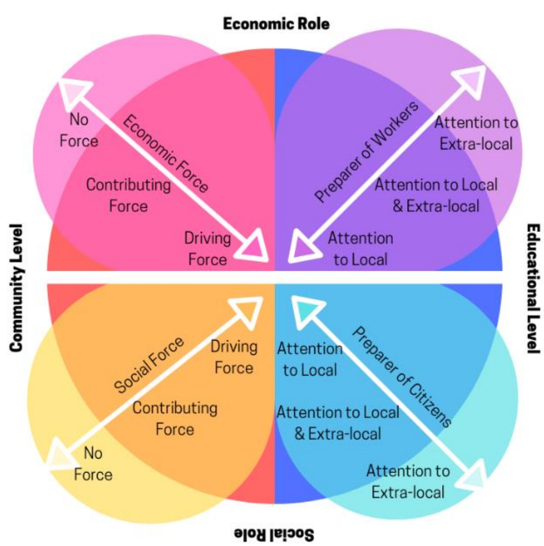

Figure 3. School-Community Framework. Source: LeBeau et al. (in preparation)

This all goes to say that the focal point of education policy is crucial to how school leaders work and operate as leaders in both their school and the community in which their students reside. If the education policy governing a state is built on an accountability system that

[1] For more information about the School-Community Framework, reference LeBeau et al. (in preparation).

measures community effects, but the school leaders are focusing on improving school effects, there is a mismatch. However, in the case of Indiana, if the accountability system truly measures school effects, then school leaders can be more cognizant of where they put their time and energy. This study is interested in further exploring Indiana's quality measure in relation to the economic well-being of a community, according to ACS measures, in order to contribute to the conversation surrounding the establishment of more "community aware" policy.

Project example: Methods and obstacles

Data

I utilized and combined multiple sources of data for this study. As Sipple, Fiduccia, and LeBeau (forthcoming) argue, this type of multi-faceted approach offers important insights into the complexities of the systems that work together in rural communities, but they also acknowledge that this data integration is not without its challenges. This rang true in this project, as I combined data from four different sources to acquire my final dataset. ACS data provided population sizes along with demographic and economic indicators for the places. The Indiana Department of Education provided quality scores in the form of letter grades for each individual school. The National Center for Education Statistics' (NCES) Elementary and Secondary Information System (ElSi) provided the latitude and longitude locations of each school along with the official description of the school district type.[2] Lastly, the US Census Bureau provided TIGER shapefiles to serve as geographic polygons for the places so that I could match schools within their places. I utilized Stata, Excel, and ArcMap to carry out these merges. With each merge, I established guidelines for which observations to keep and drop. The main criteria included places in Indiana with populations of 2,500 or less (using Lyson, 2002 as a guideline)

[2] I utilized ElSi's Agency Type variable in order to determine if the school was part of a local school district as opposed to a charter school or other type of state agency.

and the traditional public schools that fell within the boundaries of those places. The most recent form of Indiana's grading system, focusing on growth, was released in the 2015-2016 school year (IDOE, 2018), motivating my decision to use three years of data: 2016, 2017, and—the most recent year available—2018.[3] This provided me with three letter grades for each school in my analysis, and I limited my analysis to only schools that had all three years of data. This led me to a total of 485 places with populations less than or equal to 2,500—338 places without schools and 147 places with schools, with a total count of 213 schools.

Before delving into the details of working with ACS data and applying Jurjevich and colleagues (2018) guidelines, I will provide an overview of the variables used in this analysis. Independent and dependent variables were chosen building off two previous studies examining similar questions (Lyson, 2002; Sipple, Francis, & Fiduccia, 2019), with the goal of running a regression model that captures the relationship between school quality (or lack of) and the income level of a community, above and beyond the effects of other factors. Below is a brief description of my variables, and a more detailed description of how the ACS defines each variable can be found in Table 1. All dependent and independent variables were pulled from the ACS with the exception of the school presence/quality variable which was pulled from the IDOE. As previously discussed, all ACS estimates represent the averages over 5 years of data.

[3] 2018-2019 data was not used because it was labeled as a "hold harmless" grade due to changes in state testing.

Table 1. Description of dependent and independent variables

	Variable Name	Label	Description	2016 Obs	2017	2018
Dependent Variable	**Economic Characteristic**					
	PCInc	Per Capita Income	Per Capita Income of families (inflation adjusted dollars)	483	480	480
Independent Variables	**Population Characteristics**					
	Pop	Population Size	Total population estimate	485	485	485
	PctU18	% Under 18	Percentage of total population under 18 years of age	485	485	484
	PctPov	% Poverty	Percentage of families and all people whose income in the past 12 months is below the poverty level	485	485	484
	PctWhite	% White	Percentage of total population white alone	485	485	484
	Workforce Characteristics					
	PctMgmt	% Management Workers	Percentage of civilian employed population 16 years and over who work in professional scientific and management and administrative and waste management services	484	483	483
	PctSelf	% Self Employed	Percentage of civilian employed population 16 years and over who are self-employed in own not incorporated business workers	484	483	483
	Educational Attainment					
	PctHS	% High School Graduate	Percentage of population 25 years and over that are a high school graduate or higher	485	485	484
	PctBch	% Bachelors Degree	Percentage of population 25 years and over that have a bachelor's degree or higher	485	485	484
	School Presence/Quality					
	HighQual	Presence of high quality school	Indicates school/no school and quality of school	485	485	485

Dependent variable

Economic characteristic

In choosing a dependent variable for analysis, I began with three economic variables: housing value, per capita income, and household income. Upon a few initial analyses, I found that housing value and household income had more missing values than per capita income (roughly between 20-30 missing observations for each variable). Similarly, further in the analysis, I found that both income variables yielded similar results, convincing me to choose just one for the regression—per capita income (*PCInc*).

Independent variables

Population characteristics

In order to capture the characteristics of the population, I included a range of variables. The raw population estimate (*Pop*) served as a way to make sure places fit the less than or equal to 2,500 criteria and also served as a community size variable. I then included percent of the population under 18 years (*PctU18*) in order to gauge the proportion of children in the community. My last two population variables were percent of the population in poverty (*PctPov*) and percent of the population that identifies as white (*PctWhite*). The poverty measure served as a measure of community wealth and the percent white variable illustrated diversity (or lack thereof) in a community. As descriptive statistics will show, a majority of the communities sat roughly between 90-100% white, depicting a largely homogenous population. For this reason, creating a binary variable for the percentage of white residents compared to all other races was the logical conclusion.

Work characteristics

Because the jobs that people hold undoubtedly contribute to the income they receive, I included a couple variables to depict certain types of workers in the community. The management workers variable (*PctMgmt*) captures the percentage of individuals working in professional and managerial positions, and the self-employed workers variable (*PctSelf*) shows the percentage of individuals who receive a self-employment income.

Educational attainment

The amount of education an individual receives often determines what type of occupation they go into and what level of income they receive, leading me to include two variables surrounding education attainment. The two variables include the percent of people over 25 who are high school graduates (*PctHS*) and the percent of people over 25 who have a bachelor's degree (*PctBch*).

School presence and quality

The data I collected on the location of schools and school quality underwent multiple steps to create the final school quality variable. To begin, I turned to the IDOE letter grade data and turned this variable into numerical values, replacing the A-F values with a scale from 1-5 respectively. This took care of the letter grade data, so I set it aside and began making indicator variables for the places with schools. I pulled up the data merged in ArcMap to determine which places had schools within their boundaries and which did not. With this information, I created a binary variable indicating whether or not a place had a school (*NoSch*). Next, because there were a total of 50 places with more than one school, I created a variable that indicated how many times a community's ACS values were duplicated. Places with more than one school were represented multiple times in my data because their data was replicated each time a new school's information was attached to that place. For this reason, I needed to collapse all of the school information for one place into one observation so that I only had one result for each community. I did this by taking the mean letter grade of all the schools in a community. After examining the standard deviations from the mean of these scores, I found that the range of grades in a community did not deviate much from their mean, confirming my choice to display the mean grade of a community.

With each community either receiving one grade ranging from 1-5 or being flagged as not having a school, I then created an ordinal variable that ranged from no school to low quality school to high quality school. I deemed a school grade "low quality" if it received a C, D, or F and "high quality" if it received an A or a B. However, after running a few tests on these variables, I found that there was not much difference between either having a low quality school or not having a school, compared to having a high quality school. This resulted in creating a binary variable (*HighQual*) that compared communities with a high quality school to communities with a low quality school or no school at all.

ACS obstacle: Margin of error

With my variables set and data cleaned, I moved on to run descriptive statistics on my variables. Descriptive statistics on the ACS variables and the school presence/quality variable can be found in Table 2 and Table 3, respectively. Because the summary statistics of ACS estimates are similar across the years, Table 2 only includes 2018 results for ease of presentation. However, this next step in the research process led me to question the validity of my data and brings me to a major obstacle that rural researchers must prepare for and acknowledge when working with ACS data. While exploring the population size estimate, I found there to be large fluctuations within communities across the three years of data. To exemplify, Raglesville CDP had a noticeable fluctuation in population size that went from 30 in 2016, 24 in 2017, and all the way down to 0 in 2018, appearing as though it had no inhabitants. Similarly, another CDP larger in size, Landess CDP, went from a population estimate of 287 in 2016, dropped down to 102 in 2017, and dropped even further to 53 in 2018, indicating that this CDP dramatically lost over 200 inhabitants in two years. Both of these examples happen to be CDPs; however, as previously stated, CDPs are statistical entities defined by the US Census Bureau in order to provide data for places that do not have active governments but are seen as having a relatively high population density and local identity. And out of the 147 places with schools, 18 of those are defined as CDPs, making their ACS results important sources of information for those school leaders.

Table 2. 2018 Descriptive statistics for ACS variables

Variable	**Mean**	**Std. Dev.**	**Min**	**Max**
PCInc18	25653.54	12587.76	5469.00	128650.00
Pop18	794.32	615.71	0.00	2472.00
PctU1818	22.72	8.43	0.00	66.70
PctMgmt18	6.06	6.85	0.00	97.20
PctSelf18	4.55	4.90	0.00	63.40
PctPov18	14.61	11.27	0.00	100.00
PctHS18	87.17	8.79	21.10	100.00
PctBch18	14.65	13.59	0.00	90.00
PctWhite18	96.56	4.53	62.28	100.00

Table 3. Descriptive statistics for school presence and quality variables

Grade	**Obs**		
	2016	2017	2018
A	31	34	37
B	64	66	66
C	44	38	33
D	8	8	9
F	0	1	2
Total Schools	147	147	147
HighQual Variable			
No School/Low Quality	397	389	384
High Quality	88	96	101
Total Places	485	485	485

This fluctuation led me to consult a couple of colleagues at Cornell University who work with the US Census Bureau. Upon initial consultation, they both acknowledged that the US Census place-based data is prone to large amounts of error for small rural populations; however, they could not offer a different unit of analysis from the Census that fit my research question. The school district as a unit often encompasses multiple places, creating a larger unit of analysis, but this did not allow me to parse through individual communities. My colleagues brought the details of the ACS MOE to my attention, explaining why my data—taken as is—was lacking reliability. From here, I explored the MOE of my estimates to get a better understanding of the level of error associated with the results.

There are a couple of ways a researcher can move forward with this information: 1) alter the research question so that it fits a different level of analysis or pulls from a different source of data; or 2) move forward with the data, working with the margin of error and acknowledging its presence in the analysis and results. Although the first option is a viable option for researchers who are able to tap into other data sources or ask different questions, it is not a realistic option for all—especially rural researchers interested in individual communities. Accordingly, a nuanced conceptual understanding and a guide to navigate this data is necessary for a rural researcher's research tool kit. This is where the guidelines set out by Jurjevich and colleagues (2018) can and should be utilized. To reiterate from before, these five recommendations include: "Report the MOE of ACS estimates, indicate when they are not reporting the MOE, provide context for the (un)reliability of ACS data, consider alternatives for reducing statistical uncertainty, and conduct a test of statistical significance when comparing ACS estimates over time" (Jurjevich et al., 2018, p. 113). Using this set of guidelines, I will outline how I applied these techniques to my own research project, while also adjusting the suggestions to fit the intricacies of my data.

1) Reporting the MOE of ACS Estimates & 2) Indicating When They Are Not Reported

For the sake of this project, I will display 2018 MOE statistics because the statistics for all three years of data proved to be similar across the board. Because most of my variables were represented as percentages, I chose to present the margins of error as summary statistics for the variables this applied to. It is easy to compare across estimates, as they all range from a scale of 0-100. This does not include the percent white (*PctWhite*) and percent under 18 (*PctU18*) variables because they were combinations of other variables and therefore do not have MOE statistics. In comparison, the population count variable (*Pop*) was a raw number, making it difficult to compare the MOE for small populations and larger populations. For this variable, I chose to calculate the coefficient of variation (CV), as demonstrated earlier, for a group of places chosen based on their population estimate to display a range of sizes, to illustrate how it impacts the smallest to largest populations.

MOE summary stats for percentage variables. Table 4 displays the MOE summary statistics for four of the percentage variables. To clarify, these MOE summary statistics are different from the coefficient of variation, so their reliability cannot be determined according to the CV reliability guidelines. Rather, the MOE summary statistics provide a useful range of values within which the actual value is likely to fall. The mean MOE sits between 8.40% and 10.91% for the four variables. Likewise, the four variables have similar minimums and maximums; however, it is worth noting the large gap between the minimum and maximum. For *PctPov*, the minimum is 0.90 and the maximum is 89.2, displaying the variation in MOE across places—some places experience much larger levels of MOE than others.

Table 4. Margin of error for percentage estimates					
	n	**Mean**	**Std. Dev.**	**Min**	**Max**
PctMgmt18	483	9.75	13.19	0.90	85.90
PctPov18	484	10.91	10.21	0.90	89.20
PctHS18	484	8.69	8.99	0.60	79.00
PctBch18	484	8.40	8.69	1.60	79.00

In order to break this down further and determine which places have estimates with larger amounts of error, I separated the observations based on population size. Table 5 displays the MOE statistics for places with populations less than or equal to 500, and Table 6 displays these statistics for populations between 500 and 2,500. Notably different between the two are their means. The smaller populations display a higher mean MOE ranging from 12.31% to 17.45% compared to the larger populations which have mean MOE ranging from 4.61% to 7.43%. And although their minimum values sit closely to each other, the difference in the maximum MOE for both sizes of populations also sees a notable difference ranging from 79% to 89.2% for smaller populations compared to 28.3% to 60.7% for the larger population sizes.

Table 5. Margin of error for percentage estimates. Places with populations less than or equal to 500					
	n	**Mean**	**Std. Dev.**	**Min**	**Max**
PctMgmt18	198	17.15	17.87	1.60	85.90
PctPov18	199	15.90	13.22	1.00	89.20
PctHS18	199	13.28	11.85	1.20	79.00
PctBch18	199	12.31	12.03	1.60	79.00

Table 6. Margin of error for percentage estimates. Places with populations between 500 and 2,500					
	n	**Mean**	**Std. Dev.**	**Min**	**Max**
PctMgmt18	285	4.61	3.01	0.90	28.30
PctPov18	285	7.43	5.10	0.90	60.70
PctHS18	285	5.48	3.79	0.60	54.80
PctBch18	285	5.66	3.06	1.80	33.30

Accordingly, these summary statistics reflect the variation of MOE across the places, and it is crucial for researchers to acknowledge the larger MOE in smaller population sizes in order to better understand their data and to accurately portray their results to their stakeholders.

MOE CV statistics for population variable. When examining the population size, which is a raw number variable, I calculated the CV for five of the places. I chose five locations that captured a wide range of sizes, from a population of 127 to 2,000. As seen in Table 7, as population size increases, there is a roughly upward trend in MOE; however, this is coupled with a downward trend in CV. Further, according to the Esri (2014) guidelines outlined earlier[4], three out of the five places fall into the "medium reliability" range and should be used with caution. Therefore, although this is not an exhaustive list of CVs for all 485 observations, this supports the results of the MOE summary statistics of the percent variables which also found that the smaller population sizes have higher amounts of error.

Table 7. Coefficient of variation for select places

Place	**Population Size**	**MOE**	**CV**
Dubois CDP	127	67	30.63
Switz City town	250	101	24.56
Reynolds town	501	94	11.41
Wolcottville town	1021	207	12.32
Ferdinand town	2000	276	8.39

High Reliability	Medium Reliability

3) Provide context for the (un)reliability of ACS data

These statistics display large amounts of error in the estimates of the smallest populations. And as seen in Table 5, there are just under 200 places with population sizes of 500 or below. This

[4] A CV of 12 or less is considered "high reliability," between 12 and 40 is considered "medium reliability," and anything above 40 is "low reliability" (Esri, 2014).

is a little less than half of the total observations. With that being said, it must be acknowledged that much of the data has high levels of error associated with its estimates.

4) Consider alternatives for reducing statistical uncertainty

Because the variables in this analysis have a considerable margin of error, it is our role as researchers to decrease as much of the statistical error as possible. Jurvevich et al. (2018) offer a few adjustments that can be made in order to increase reliability such as collapsing data when possible and using multi-year estimates. However, in this case, I also suggest that researchers control the controllable and eliminate as much statistical error as possible within the statistical analysis they are running. I did this by transforming the variables in order to create a more normal distribution.

I began by running an analysis on all of my variables in order to see if any of the possible transformations improved the normality of the distribution of the variable. For my dependent variable (*PCInc*), taking the log of the variable created a more normal distribution, and for the population variable (*Pop*), taking the square root of the variable also resulted in an improved distribution. However, after examining the other variables, I found that no transformations improved the normality of their distribution. Accordingly, I turned these variables into ordinal categorical variables so as to improve the analysis. I did this by separating the variables into four categories. Because the summary statistics were similar across all three years of data, I took the approximate mean and standard deviation across the years and applied this as my criteria for segmenting the variable. This resulted in two categories extending a standard deviation outside of the mean (one standard deviation above and one standard deviation below), and then my final two categories included anything that extended above or below one standard deviation. Some variables that had means close to 0% or 100% only have three categories, as one standard deviation from the mean reached beyond 0% or 100% (*PctMgmt* and *PctSelf*). Lastly, I created a binary variable for the *PctWhite* variable. As previously mentioned, this variable was very concentrated and most observations sat within an 8-10% range, so splitting the variable into multiple categories resulted in comparing populations that were not very different. Thus, in creating the binary variable, I kept with the standard deviation format and split the variable into one standard deviation above and below the mean compared to anything outside of that. One standard deviation above brought the variable to 100% and one standard deviation below was at 92% white. So 92% white was my cutoff point that split the binary variable in half.

After transforming and altering my variables, I conducted regression diagnostics on the initial model and the adjusted model. The diagnostic tests showed improvement in the model after the transformations. Therefore, by transforming and adjusting my variables, I was able to decrease the amount of statistical uncertainty in my regressions.

5) Conducting a test of statistical difference

This last guideline set out by Jurjevich and colleagues does not apply to the data in this example. Rather, this suggestion to conduct a test of statistical difference is best utilized when comparing two nonoverlapping time periods of data (ex. 2010-2014 and 2015-2019 5-year estimates). In the case of my data and research goals, I did not intend to compare my results across the years, so I do not need to test for the significance of the difference across years. And in addition to this, the 5-year estimate periods overlapped, as they were consecutive and pooled years, so it did not fit the nonoverlapping time periods criteria.

Project example: Data analysis and results

After following the guidelines and carrying out statistical tests to better understand the error associated with my data, I conducted an OLS regression to test how the quality of the school (or lack of) relates to the per capita income level of the community, above and beyond other demographic factors. In this case, the log of the per capita income variable (*LogPCInc*) served as my dependent variable. Although this variable had some missing observations, I chose to continue my analysis without these observations, as they were few in numbers. I ran the regressions with 483 observations for 2016 and 480 for 2017 and 2018. Because the initial number of places was 485, this slight decrease was not significant enough to warrant reason to adjust my analysis. Further, the results of each regression are outlined below.

OLS regression results

Across the three years, the OLS regressions produced similar results in terms of direction of relationship and significance. Six of the control variables are significantly related to the log of the per capita income variable. Table 8 displays the results in more detail across the years. *PctMgmt_*, *PctSelf_*, *PctHS_*, and *PctBch_*[5] are positively related, illustrating an increase in per capita income as these variables increase. Conversely, *PctU18_* and *PctPov_* are negatively related, and so per capita income decreases as they increase. Three variables are not significantly related to the log of the per capita income. The square root of the raw population count (*SqPop*) is not significantly related to the per capita income level of the community. Similarly, the binary percent white variable (*HighWhite*) showed no relationship with the dependent variable.

[5] The _ after each of the variables indicates that they have been transformed or altered to improve statistical significance as laid out in the previous section.

Table 8. OLS regression results for all three years			
Dependent Variable: LogPcInc			
Variable	**Coefficient**		
	2016	2017	2018
SqPop	-0.00	0.00	-0.00
PctU18	-0.01**	-0.01**	-0.01**
PctMgmt_	0.04*	0.06**	0.04**
PctSelf_	0.05**	0.06**	0.06**
PctPov_	-0.10**	-0.10**	-0.11**
PctHS_	0.07**	0.06**	0.08**
PctBch_	0.18**	0.17**	0.14**
HighWhite	0.00	-0.02	0.00
HighQual	-0.05^	-0.04^	-0.04
R-squared	0.60	0.60	0.56
Note. Significant p-values are denoted by ^ 0.10; * 0.05; ** 0.01			

School presence/quality

The independent variable of interest, the school presence/quality variable (*HighQual*), was not significantly related to the log of the per capita income variable at the 0.05 level. Put differently, having a high quality school in comparison to a low quality school or no school in a community is not significantly related to the per capita income level of a community. Compared to the other two independent variables that were not related with high p-values (*SqPop* and *HighWhite*), the p-value of the *HighQual* variable is worth noting, as it was 0.12, 0.08, and 0.08 for 2018, 2017, and 2016 respectively. I mention this and draw attention to the 0.10 p-level in the tables because of the low number of observations in my study (N= 480 and 483 across the years). Although this variable was not significant according to the suggested 0.05 level, its level of significance is borderline. Also of note is the coefficient of this variable. For all three years, the coefficient is negative. Therefore, as you go from not having a school or having a low quality school to having a high quality school, the per capita income level of a community decreases.

Further tests

The aforementioned results assess the relationship between having a high quality school in a community and the per capita income level of a community, above and beyond a list of controls. In order to explore this variable further and determine if Indiana's school quality measure relates to community wealth, I ran a correlation test and found that the *HighQual* variable was not collinear with any of the community data.

To take this one step further and determine if any of the community variables could predict whether or not the school(s) in a community were high quality, I ran a logistic regression with the *HighQual* variable as my dependent variable (Table 9). In order to improve interpretability for this test, I used the original, untransformed variables for both per capita income (*PCInc*) and population size (*Pop*) and slightly adjusted the variables to be in units of 1,000 and 100 respectively. Therefore, each unit increase in per capita income is interpreted as an increase of $1,000 (*PCInc_1000*) and each unit increase in population is interpreted as an increase of 100 people (*Pop_100*). For all three years of data, these two variables (*PCInc_1000* and *Pop_100*) were significant predictors at the 0.05 and 0.01 level respectively. For every $1,000 increase in per capita income, a community is ~6-9% less likely to have a high quality school. On the other hand, for every 100 person increase in population, a community is ~13-15% more likely to have a high quality school. Lastly, for the 2017 model, *PctBch_* was a significant predictor of the *HighQual* variable and illustrated that for every unit increase in *PctBch_*, a community is 73% more likely to have a high quality school. In other words, above and beyond the education level in the community and community size, there is a negative relationship between per-capita income and school quality, at least as measured by the state of

Indiana. I remind the readers that this measure of quality relies heavily on year to year growth in achievement rather than absolute level of achievement in any given year.

Table 9. Logistic regression results for all three years

Dependent Variable: HighQual

Variable	**Odds Ratio**		
	2016	2017	2018
PCInc_1000	0.91*	0.93*	0.94*
Pop_100	1.13**	1.14**	1.15**
PctU18	1.00	0.99	0.99
PctMgmt_	0.95	0.98	0.80
PctSelf_	1.46^	1.33	1.24
PctPov_	0.79	1.01	0.81
PctHS_	1.15	1.07	1.16
PctBch_	1.32	1.73*	1.44
HighWhite	0.66	0.58	0.84

Note. Significant p-values are denoted by ^ 0.10; * 0.05; **0.01

Discussion

After having followed the guidelines outlined by Jurjevich et al. (2018) and presenting my data and results as clearly and transparently as possible, I now analyze the implications of these results and how they contribute to a school/community discussion as well as an ACS discussion. Analyzing the results in relation to other studies, policies, and prior knowledge allows me to come up with multiple explanations as to *why* I received the results I did and the multiple interpretations of what this means for rural schools and communities. This is perhaps the most critical outcome of this paper—recognizing the utility of ACS data to broadly examine and analyze geographic patterns and then using those results as a launch pad for further studies that can delve deeper into figuring out the *why* behind those results. I bring emphasis to the *whys* behind the results, because they are what will drive the motivation for researchers to dive deeper and ask more questions in order to make sure rural schools and communities have a voice in education policy. The following section is broken down into two parts outlining how these results contribute to the discussion surrounding school/community interactions and the discussion surrounding Indiana's accountability system. Each part outlines how these results have implications for both "community aware" education policy and rural researchers use of ACS data.

School/community interaction

Although the relationship between having a quality school in a community and the per capita income level of that community is considered borderline significant according to standard p-value guidelines, its coefficient is consistently negative all three years, making its results worth noting. These results indicate that communities that do not have a high quality school (or any school at all) have a higher income level on average than those with a high quality school. This differs from previous studies conducted in New York State that show that the presence or proximity to schools—not taking quality into account—relate to greater economic vitality of a rural community (Lyson, 2002; Sipple, Francis, & Fiduccia, 2019). This could be a result of the change in context (New York v. Indiana), or a result of the addition of the school quality attribute, as these previous studies did not include the quality of a school in their measurement. However, in order to test this latter point, I ran the same regressions without the school quality indicator. Using a binary school/no school variable, I found similar results. Although the p-values were not significant at the 0.05 level[6], the coefficient indicated that having a school in the community was negatively related to per capita income. Put differently, going from having a school to not having a school, the per capita income of a community increases.

[6] The p-value was 0.09 for 2018, 0.18 for 2017, and 0.31 for 2016.

Why?

This is where the *why* comes into question. Why are these results different from previous studies? Is it the difference in the years the studies were conducted? Is there something inherently different in the education and community makeup of Indiana and New York? Perhaps differences in state aid? And pushing the school quality question further, why is it that the school quality measure is not significantly related to community wealth? And thinking more critically about Indiana's accountability system: because the system is based on negative consequences and only intervenes with schools that receive consecutive F grades, a school that receives an A does not receive any different action from the state in comparison to a school that receives a D. Rather, schools that don't receive consecutive F grades, no matter if they are As or Ds, are treated the same by the state. Therefore, if the only benefit to getting an A is "bragging rights" on the part of the school, then perhaps the letter grade accountability system is more inherently tied to the community than initially thought. Because Indiana's accountability system was built partially with the intention of making it easily interpreted by families and the community (Hiller et al., 2012), it is easy to infer that this type of community-facing grading system also serves as a way for schools to attract families to their district. After all, Indiana is a national leader in the school choice movement and has one of the largest school voucher programs in the country (Turner, 2017; EdChoice, 2020). Accordingly, how much do school grades contribute to a school's ability to attract and maintain families? And in rural communities where school choice is not as widely available, how much does a school's letter grade contribute to the local pride and local identity of the community?

This relates to the multiple roles that a school plays in a rural community, as outlined by LeBeau et al. (in preparation) in their school-community framework. Initially, this study was created under the assumption of the school as an economic force that has the potential to contribute to the economic vitality of a community, as illustrated through its per capita income level. However, LeBeau and colleagues' (in preparation) description of an economic force also captures a school's potential to serve as an attraction to families and businesses to the community. Perhaps the results of this study indicate that the quality of a school could potentially have more of an impact on the migration patterns of a community—attracting or detracting families from the community. Similarly, as a social force, a school has the potential to be the central source of community pride and local identity. A school with a high letter grade has the ability to serve both of these functions by attracting families and serving as a source of community pride. Accordingly, when considering this in light of how school leaders approach Indiana's accountability system, is their focus on not getting an F so as to skirt state intervention? Or is their focus on achieving as high of a letter grade as possible to illustrate their success to their own families and community as well as the families and communities in their surrounding network?

Potential future studies

Although all of these potential *whys* might be a frustrating conclusion to the researcher, I argue that they highlight the numerous benefits of using ACS data. This is where I take Jurjevich and colleagues' (2018) five guidelines one step further and suggest using this data and its results as a launch pad for other studies. These types of questions illustrate how the results from this study can be used to propel other studies—qualitative and quantitative alike. This study was set up to examine the relationship between the quality of a school and the economic vitality of a community. However, after analyzing the results in conjunction with the details of the policy surrounding Indiana's accountability system, perhaps a future study might examine the relationship between school quality and migration patterns or population change as opposed to per capita income level. Rather than examining the school as an economic force that relates to the income level of a community, how does the school work as an economic force that relates to the in-migration or out-migration of families? In this case, population change might be utilized as the dependent variable. Further, this study also opens up the potential for a qualitative study that examines the way school leaders approach the state's accountability grading system. What are their motivations behind achieving a high letter grade? Is it to avoid state intervention? Or is it to illustrate the successes of the school to the wider community? If the answer lies in the latter response, then this has implications for future improvements and adjustments to Indiana's accountability system. If the accountability system is interested in school effects but school leaders are interested in the effect the grade has on the community, then the two should be discussed in conjunction so as to create a more "community aware" education policy.

Indiana's accountability system

Taking the results of this study and analyzing them in relation to what they tell us about Indiana's accountability system leads to many interesting conclusions. As previously outlined, the most recent version of Indiana's A-F grading system was shaped with the intention of measuring the individual growth of students in addition to raw performance scores (IDOE, n.d.). Accordingly, no correlations were found between the *HighQual* variable and other variables, leading the results of this initial test to support Indiana's accountability grade as a measure of school effects on student growth as opposed to a measure of community wealth. Similarly, the borderline significance of the *HighQual* variable's negative relationship to a community's per capita income level in the OLS regression supports this idea yet again, showing that the relationship of a community having a high quality school is *not* in fact positively related to an increased per capita income. Lastly, the results of the logistic regression, yet again, display a negative relationship between per capita income and the presence of a high quality school, as every $1,000 increase in per capita income results in a community being roughly 6-9% less likely to have a high quality school. And in addition, the percentage of people in poverty in a community is not a significant predictor of having a high quality school either.

Why?

With all of this in mind, it is logical to draw the conclusion that Indiana's accountability system is able to measure student growth above and beyond community factors. And this is one potential *why* explanation. However, there are multiple *whys* that can be drawn from these results. On the other hand, another explanation might be that Indiana's grades are meaningless, and the relationships are just spurious. Another potential conclusion is that there is something unique about rural schools and communities that allows them to meet the needs of their students above and beyond the impacts of family background. After all, rural researchers have argued that small school and district size have the ability to mitigate the impact of family socioeconomic status on student achievement (Howley, Strange, & Bickel, 2000; Johnson, Howley, & Howley, 2002).

A final and sobering conclusion might be that the ACS margin of error is so great in this study, that I am unable to produce accurate results. As was acknowledged earlier, much educational research has shown there to be a strong correlation between family socioeconomic status and the academic performance of a student (Coleman, 1966; Sirin, 2005; Tienken, 2017), and if these results do not coincide, perhaps the study exposes the level of error in the data.

Potential future studies

Because I cannot confidently conclude whether or not my results are impacted by the ACS MOE due to my small sample sizes, a future study might examine all places in Indiana in order to capture more places with higher populations and lower margins of error. Are the same results replicated across the state in communities of all sizes? Is the Indiana A-F grade able to capture school effects outside of community wealth in urban and suburban schools too? And yet another study: what would happen if raw standardized test outcomes were used as the dependent variable instead of Indiana's A-F grade that also captures growth and graduation rates? If using only standardized test scores as a measure is able to predict community wealth, then this would support the conclusion that Indiana's accountability system is able to measure school effects, above and beyond community wealth and family socioeconomic status after all.

Conclusion

The purpose of this paper was two-fold, exploring the conceptual and the technical of using a multifaceted approach to data in order to delve deep into the nuances of the relationship between education policy and rural communities. Examining Indiana's accountability system in relation to the economic vitality of communities brings up crucial questions about the ability of the accountability grades to measure beyond community effects and about how school leaders work at the intersection of institutional pressures and community pressures. It even calls into question the reliability of the ACS data used for analysis. As such, it's important to note that utilizing ACS-type data to study trends and patterns across rural geographies is still an incredibly useful tool, if navigated diligently, to provide the most reliable results to stakeholders—whether that be leaders at the state level or the school district level. Jurjevich and colleagues (2018) provide an incredibly useful set of guidelines to assist in the navigation of this type of data, and I encourage rural researchers to adhere to these guidelines whenever plausible. But additionally, I urge rural researchers to take these guidelines one step further and really struggle and chew on the *whys* and the implications of their results. This project did not provide a clear-cut conclusion followed by explicit recommendations for policymakers and school leaders. Rather, it brought up more questions than answers, and this, I argue, is one of the greatest benefits of using ACS-type data, despite its obstacles. If navigated with intention, rural researchers can use ACS-type data to propel future studies and delve deeper into the nuances of schools and communities so as to reach a better understanding and contribute to a more "community aware" education policy.

REFERENCES

Arum, R. (2000). Schools and communities: Ecological and institutional dimensions. *Annual Review of Sociology, 26*, 395-418.

Autti, O., & Hyry-Beihammer, E. K. (2014). School Closures in Rural Finnish Communities. *Journal of Research in Rural Education (Online)*, *29*(1), 1–17.

Budge, K. (2006). Rural leaders, rural places: Problem, privilege, and possibility. *Journal of Research in Rural Education, 21(*13). http://jrre.psu.edu/articles/2113.pdf

Carr, P.J. & Kefalas, M.J. (2009). *Hollowing out the middle: The rural brain drain and what it means for America*. Beacon Press.

Casto, H., McGrath, B., Sipple, John W., & Todd, L. (2016). "Community Aware" education policy: Enhancing individual and community vitality. *Education Policy Analysis Archives, 24*(50). http://dx.doi.org/10.14507/epaa.v24.2148

Corbett, M. (2007). *Learning to leave: The irony of schooling in a costal community.* Fernwook Publishing.

Coleman, J.C. (1966). Equality of educational opportunity. *National Center for Educational Statistics.*

Cromartie, J. (2018). *Rural America at a glance: 2018 edition.* United States Department of Agriculture. https://www.ers.usda.gov/webdocs/publications/90556/eib-200.pdf?v=7228.8

Dean, M. (2010). *Understanding human need.* Bristol: Policy Press.

EdChoice. (2020). Indiana – Choice scholarship program. *School Choice.* https://www.edchoice.org/school-choice/programs/indiana-choice-scholarship-program/

Education Week. (2012, January 12). *Report awards grades for education performance, policy;*

Nation earns a C, Maryland ranks first for fourth straight year [Press release]. https://www.edweek.org/media/qualitycounts2012_release.pdf

Environmental Systems Research Institute (Esri). (2014). *The American Community Survey: An Esri white paper.* https://www.esri.com/library/whitepapers/pdfs/the-american community-survey.pdf

Harmon, H. L. & Schafft, K. (2009). Rural school leadership for collaborative community development. *The Rural Educator, 30*(3), 4-9.

Hill, H.C. (2016, July 13). 50 years ago, one report introduced Americans to the black-white achievement gap. Here's what we've learned since. *Chalkbeat.* https://www.chalkbeat.org/2016/7/13/21103280/50-years-ago-one-report-introduced-americans-to-the-black-white-achievement-gap-here-s-what-we-ve-le

Hiller, S.C., DiTommaso, A., & Plucker, J.A. (2012). The evolution of Indiana's school accountability system. *Center for Evaluation & Education Policy, 10*(5).

Howley, C., Strange, M., & Bickel, R. (2000). Research about school size and school performance in impoverished communities. *ERIC Digest.*

Indiana Department of Education (IDOE). (n.d.). Indiana student centered A-F accountability system [Power Point slides]. https://www.doe.in.gov/sites/default/files/accountability/f-accountability-presentation.pdf

IDOE. (2018, March 22). History of Indiana's accountability system. *Accountability.* https://www.doe.in.gov/accountability/history-indiana%E2%80%99s-accountability-system

Jennings, N., Swidler, S., & Koliba, C. (2005). Place-based education in the standards-based reform era--Conflict or compliment? *American Journal of Education, 112*(1), 44-65.

Johnson, J.D., Howley, C.B., & Howley, A.A. (2002). *Size, excellence, and equity: A report on Arkansas schools and districts.* https://files.eric.ed.gov/fulltext/ED459987.pdf

Jurjevich, J.R., Griffin, A.L., Spielman, S.E., Folch, D.C., Merrick, M., & Nagle, N.N. (2018). Navigating statistical uncertainty: How urban and regional planners understand and work with American Community Survey (ACS) data for guiding policy. *Journal of the American Planning Association, 84*(2).

LeBeau, K., Casto, H., & Sipple, J. (in preparation). Empowering school and community leaders: A tool to assess a school's role amid changing contexts.

Lyson, T.A. (2002). What does a school mean to a community? Assessing the social and economic benefits of schools to rural villages in New York. *Journal of Research in Rural Education*, *17*(3), 131-137.

National Research Council (NRC). (2007). *Using the American Community Survey: Benefits and challenges.* Washington, DC: The National Academies Press. https://doi.org/10.17226/11901.

Ratcliffe, M. (n.d.). Understanding "Place" in Census Bureau dad products [Power Point slides]. United States Census Bureau. https://www.census.gov/content/dam/Census/data/developers/understandingplace.pdf

Schafft, K. (2016). Rural education as rural development: Understanding the rural school community bell-being linkage in a 21st-century policy context. *Peabody Journal of Education*, *91*(2), 137-154.

Sipple, J.W., Fiduccia, P.C., & LeBeau, K. (forthcoming). The why and how of enhancing data use in rural education research and practice. In A.P. Azano, K. Eppley, & C. Biddle (Eds.), *The Bloomsbury Handbook of Rural Education in the USA*. Bloomsbury

Publishing.

Sipple, J.W., Francis, J.D., & Fiduccia, P.C. (2019). Exploring the gradient: The economic benefits of 'nearby' schools on rural communities. *Journal of Rural Studies, 68,* 251-263.

Sirin, S.R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*(3), 417-453.

Spielman, S.E., Folch, D., & Nagle, N. (2014). Patterns and causes of uncertainty in the American Community Survey. *Applied Geography, 46*, 147-157.

Spielman, S.E. & Singleton, A. (2015). Studying neighborhoods using uncertain data from the American Community Survey: A contextual approach. *Annals of the Association of American Geographers, 105*(5), 1003-1025.

Tieken, M.C. (2014). *Why rural schools matter.* The University of North Carolina Press.

Tienken, C.H., Colella, A., Angelillo, C., Fox, M., McCahill, K.R., & Wolfe, A. (2017). Predicting middle level state standardized test results using family and community demographic data. *Research in Middle Level Education, 40*(1), 1-14.

Turner, C. (2017, May 12). The promise and peril of school vouchers. *NPR.* https://www.npr.org/sections/ed/2017/05/12/520111511/the-promise-and-peril-of-school-vouchers#:~:text=Indiana's%20Program%20Offers%20Lessons%20%3A%20NPR%20Ed%20Indiana's%20private%20school%20voucher,depends%20on%20whom%20you%20ask.

United States Department of Commerce. (1994). Chapter 9: Places. *Geographic Areas Reference Manual (GARM)* (pp. 9-1 - 9-31).

United States Census Bureau (USCB). (2009). *A compass for understanding and using American Community Survey data: What state and local governments need to know*. Washington, D.C.

USCB. (2020a, February 27). The importance of the American Community Survey and the 2020 Census. *American Community Survey (ACS).* https://www.census.gov/programs-surveys/acs/about/acs-and-census.html

USCB. (2020b, April 21). About the American Community Survey. *American Community Survey (ACS).* https://www.census.gov/programs-surveys/acs/about.html

USCB. (2020c). Chapter 3: Sampling error in the ACS. *Understanding and Using American Community Survey Data: What Researchers Need to Know,* (pp. 12-14).

Wei, R., Tong, D., & Phillips, J.M. (2015). An integrated classification scheme for mapping estimates and errors of estimation from the American Community Survey. *Computers, Environment and Urban Systems, 63,* 95-103.

Zuckerman, S. J. (2020). The role of rural school leaders in a school-community partnership. *Theory & Practice in Rural Education (TPRE), 10*(1), 73-91.