J Econ Finan (2018) 42:174–190
DOI 10.1007/s12197-017-9392-z

The effects of income and population demographics on single-county bank performance

Ken B. Cyree[1] · **Brandon C. L. Morris**[2]

Published online: 4 May 2017

Abstract We investigate the role market demographics play in determining the accounting performance of banks. Specifically, we study the effect of wealth and population to determine if these market factors drive performance, or if performance is related only to bank-specific variables. Using a sample of single-county banks, we find that market demographics play an important role in determining bank performance. Our univariate findings show banks that operate in low population counties outperform those in high population counties. We also show the lowest performing group of banks operates in counties characterized by high population and high income. Our multivariate tests confirm that as county-level population decreases, bank performance increases. Moreover, we observe a significant low-income advantage after controlling for other determinants of profitability. We also find that low population levels significantly mitigate the negative effects of the 2008 financial crisis for small, single-county banks.

Keywords Bank performance · Bank growth · Profit efficiency · Macroeconomic · Local banks

JEL Classification G21

✉ Ken B. Cyree
kcyree@bus.olemiss.edu

Brandon C. L. Morris
brandon.morris@wright.edu

[1] School of Business Administration, University of Mississippi, 253 Holman Hall North, Oxford, MS 38677, USA

[2] Raj Soin School of Business, Wright State University, 284 Rike Hall, 3640 Colonel Glenn Highway, Dayton, OH 45434, USA

1 Introduction

Bank performance is widely studied, but the results are often inconclusive depending on the focus of the study. For example, the performance of banks is often related to merger activity as in Akhavien, Berger and Humphrey (1997) or ownership effects as in Bonin et al. (2005). DeLong and DeYoung (2007) study the effects of multiple bank acquisitions on changes in return on assets (ROA), changes in return on equity (ROE), and changes in other variables such as interest margin or core deposits. Demographic effects such as county-level income and population levels could be important determinants of bank performance, yet is typically missing in studies of bank performance. Therefore, in this study, we investigate whether county-level demographics of population and income are related to bank performance.

Berger and Mester (1997) list profit efficiency correlates that other studies have used including the type of regulator, market structure, geographic diversification, corporate control, concentration, asset size, holding company status, capital levels, and other exogenous variables. Berger and DeYoung (2001) focus on the distance of a banking operation from the home office as an explanation for efficiency differences in banking operations. Berger and Humphrey (1997) provide a survey of the usefulness of efficiency studies and conclude they largely inform government policy, address research issues and help to evaluate managerial performance. One theme discussed by Berger and Humphrey (1997) is the market-power versus efficient-structure debate about the determinants of profitability. In general, market power does seem to affect the prices of some types of local deposits and loans, but has little apparent effect on profit efficiency.

The common thread in most studies is to benchmark performance and then study the effects of a firm event, such as consolidation, or different business models on performance. Rarely is the focus of a banking performance study on the market in which they operate. Thus, in this study we explore whether and to what extent bank performance is affected by the demographics of a market. In particular, we ask if per-capita wealth and population size are correlated with accounting performance measures for single-county banks. We use single-county banks since it avoids issues with estimating the appropriate market demographic measures for banks that operate in many different counties. We investigate the effects of including these variables in traditional bank performance models to see if customer demographics in the market are related to performance, and whether the inclusion of these variables change other performance dynamics. To a certain extent, we are testing bank-specific choices made by bank managers versus the good fortune of being in the right market. We recognize that bank managers choose their market, but our focus is on the effects of the wealth and population in those markets on performance, if any.

Our results show that population and income characteristics are an important determinant of bank profitability. The initial univariate results show significant performance gains for firms competing in low-population counties as compared to firms competing in high-population counties. In our multivariate tests, we find that both low income and low population levels improve the performance of single-county banks. We also show that low-population mitigates some of the effects of the 2008 financial crisis, both at the onset of the crisis and in the years following. These findings are consistent with the notion that the market a bank chooses to operate in can lead to significant advantages. We conjecture the type of competition in the market is important in low-population areas that have fewer local banking options. In other words, the quality of competition is important if banks choose not to compete on pricing or fees, even though market concentration measures such as Herfindahl indices or the share of deposits by multi-county banks do not show abnormal concentration as is the case in our findings. Our results are consistent with Pilloff (1999) and Cyree and Paul Spurlin (2012) who find that rural areas have less competitive effects as banks do not appear to compete on loan or deposit pricing, even if the market has a reasonable number of competitors and therefore have lower measures of concentration. Therefore, one implication of our findings is that traditional measures of concentration do not necessarily measure the quality of the competition, and our results are consistent with low population areas have the lowest quality of competition.

2 Related literature

Prior bank performance studies are largely focused on market power, corporate events such as mergers and acquisitions, differences in performance across regulatory environments or around changes in regulation, and other factors such as the effects of small business lending. In most studies, very little attention is given to the markets in which the bank is operating, other than perhaps identifying them as rural or urban. Several methods for controlling for market conditions are used throughout the literature. We discuss the relevant methods and findings below.

The most common demographic variables are population, rural or urban market selection, and per-capita income. Hannan and Prager (2004) find that population, income, and rural markets negatively influence pricing behavior of retail banks. Hannan and Prager (2004) use the difference in deposit interest rates as the dependent variable in their study, while our study focuses on how income and population affect total firm performance – in particular the accounting measures of ROA and pre-tax ROA. Berger et al. (2007) report conflicting evidence that market demographics partly determine return on equity. In part of their sample from 1982 to 1990, higher populations reduce returns and higher income improves returns. However, Berger et al. (2007) find no relationship between population and returns, and the effect of income changes signs in the regression model for the period from 1991 to 2000. These studies are typical of the mixed evidence on market demographics and bank performance.

In an investigation on the impact of market structure on bank deposit interest rates, Rosen (2007) reports that per capita income has a strong negative effect on deposit interest rates. To control for market size, Rosen uses population and total market deposits and finds conflicting results – population has a negative effect on deposit rates while total deposits has a positive relation with population.[1]

In a study similar to our own, Athanasoglou et al. (2008) investigate bank-specific and macroeconomic determinants of bank profitability in a panel study of Greek banks.

[1] Population density was also used, although it largely had no influence on deposit rates. The authors explained that this was likely due to the market fixed-effect regressors causing the variable to measure the change in population density, which is normally small. Therefore, we do not include population density as a variable in our study.

The macroeconomic determinants are limited to inflation and deviations in real GDP, with both contributing positively to profitability. Local market conditions such as per capita income and population controls are ignored, or are insignificant in their estimates.

Berger et al. (2000) consider two macro variables, gross state product and gross regional product, while examining the profit persistence of US banks. Results suggest that regional/macroeconomic shocks remain strong determinants of performance persistence. Firms in the upper/lower end of the performance distribution tended to stay in those distributions when the local economies experienced positive or negative macro shocks. Hannan and Prager (2009) examine how the presence of multi-market banks affects single-market bank performance. Hannan and Prager use population as their sole demographic control after separating their sample into rural and urban markets (MSA or non-MSA). They find limited support that population negatively affects performance in rural markets, but find that population negatively affects performance for banks in urban markets.

Many other studies ignore market demographics. For example, Kwan (2003), DeYoung and Hasan (1998) and Vennet (2002) control for local market conditions by simply using a Herfindahl index. The Herfindahl index nicely adjusts for competition levels, but is not a direct measure of local economic and population conditions. Thus, we investigate if local market characteristics play a significant role in determining bank profits in addition to traditional concentration measures.

Although the scope of our study does not include profit efficiency, it is important to understand the market variables that research has shown to impact profit efficiency. Bos and Kool (2006) find that banks in large rural markets are on average more cost efficient and that population has a negative impact on cost efficiency. GonzÁLez (2009) finds that both GDP and inflation are positive and significant determinants of profit efficiency. Albertazzi and Gambacorta (2009) also find a positive link between net interest income and GDP. These studies indicate conflicting results on macroeconomic and demographic variables and the relation to bank profit or profit efficiency.

Cyree and Paul Spurlin (2012) show that when large multi-market banks are present in a rural market, banks have higher ROA but lower efficiency. Their results indicate the importance of controlling for the presence of larger competitors in the market due to the effect on bank performance of smaller and rural banks. In this study, we use the proportion of deposits by multi-county banks as a measure of the presence of competitors from outside the local market consistent with findings by Hannan and Prager (2009) and Cyree and Paul Spurlin (2012).

3 Data

The data we use in this study are obtained from several sources. Accounting performance data are obtained from the annual reports of condition and income (commonly called the Call Reports) filed by each institution, and branch-level deposits were gathered from the Federal Deposit Insurance Corporation's Summary of Deposits (SOD). County characteristics are obtained from both the Inter-University Consortium of Political and Social Research (ICPSR)[2] and the Bureau of Economic Analysis. Our research question is whether bank performance is influenced by the market economic environment or if firm-level determinants are related to performance? Ideally we would compare performance measures for every branch in every county, but unfortunately performance data at the branch level are not available. While the SOD data provides branch-level inputs such as total deposits, these data lack cost and profitability measures necessary to determine the county-level performance of multi-county institutions. Thus, for our empirical tests, we eliminate all multi-county banks and focus on banks with offices in only one county so that we can accurately capture how each county's population and income characteristics correlate with overall bank performance without the confounding effects of operating in multiple counties and markets.

Financial economists often use two primary measures of bank accounting performance, return on assets (ROA) and return on equity (ROE). Athanasoglou et al. (2008) explain that ROE disregards the risk associated with high leverage and that ultimately leverage ratios are many times the result of regulation, whereas ROA reflects the ability of a bank's management team to generate profits from their own assets. Berger (1995) indicates that "the profitability measures after-tax ROA and ROE are standards in banking research," and that he obtains similar results with pre-tax measures. Thus, based on prior research and industry and regulatory standards, we also use ROA and pre-tax ROA as our single-county bank (SCB) performance measures. We define pre-tax ROA as earnings before taxes divided by assets (EBTROA) since using this measure mitigates the possible effects of S-corporation election that is likely for small single-county banks in our sample. To prevent extreme outliers from skewing our results, we trim the data at the 1st and 99th percentiles.

Our final data set consists of 22,758 county-year observations on 2063 US counties and 49,839 SCB-year observations on 5920 unique SCBs during the years 2001 through 2014. This annual data set is an unbalanced panel due to acquisitions, bank failures, and new entry across our sample period.

It is customary in the bank-performance literature to classify banks as either community banks or large single-market or multi-market banks, and in many cases the analysis is based on these distinctions (e.g., DeYoung and Hasan (1998); Deyoung et al. (2004); Rosen (2007); Berger et al. (2007); and Hannan and Prager (2009)). The focus of our study is to identify the effects market demographics have on performance, therefore we include all single-county banks in each market and do not eliminate any particular class of bank based

[2] The Inter-University Consortium for Political and Social Research (ICPSR) provides a rich data set that contains an array of county characteristics by which researchers can investigate contextual influences at the county level. Data describing age, sex, and race demographics, crime rates, government distributions, employment statistics, and similar variables from several government agencies have been compiled into one useable data set covering the sample period 2000–2007. We use only county land area from this database, which does not change throughout our study period, and therefore we use the land area data across our entire sample period.

on size. We do note that a single-county bank is likely to be a community bank given small size and local focus and ownership.

Table 1 presents summary statistics for variables used in the analysis. Panel A reports county-level summary statistics, Panel B reports SCB statistics, and Panel C reports per capita income, population, number of SCBs, ROA and pre-tax ROA for each year in the sample. The Herfindahl index (HHI) we employ is a county-level index that measures the concentration of total deposits in each county for all banks in county (not just SCBs) and is our proxy for market competition.[3] A few patterns in these data are worth noting. First, the number of unique banks decreased throughout the sample, consistent with increased numbers of mergers and acquisitions, and failures due to the financial crisis starting in 2007 Q3. Figure 1 charts ROA and EBTROA across our sample period. From 2001 to 2006, ROA and EBTROA remains relatively consistent, but from 2007 to 2009 SCB profitability sharply declined due to the financial crisis. In 2010, SCB profits begins to improve, and this pattern continues to the end of our sample period.

4 Empirical results

In this section, we discuss the empirical tests of whether or not single-county bank performance is significantly related to the local market factors of income and population.

4.1 Univariate statistics

We investigate whether bank performance is due strictly to managerial skill or if the bank's market population and income levels drive profitability. We begin our analysis by examining profitability of SCBs that reside in low to high population centers. Figure 2 graphs average ROA by population quartile over time. Between 2001 and 2014, banks in low population counties generally outperform banks in high-population counties. What is striking about Fig. 2 is the dramatic decline in profitability of high-population SCBs during the great recession followed by an impressive rebound. Figure 3 shows ROA across county-income quartiles. We see a similar pattern in high-income markets with SCBs underperforming through the financial crisis (2009), but eventually outperforming banks in lower income quartiles by 2011 and continuing that

[3] Our primary sample included only single-county banks, although the deposit Herfindahl is computed using both single-county banks and multi-county banks to provide an accurate description of competition in each county. We use total deposits for all banks in the county, consistent with most banking studies. We do not use multi-county banks in our sample because we cannot decompose ROA to the county level with SOD and call report data. Our Herfindahl is scaled so that its value ranges from 0 to 1, rather than from 0 to 10,000. It is computed for each county (i) for each year (t), using all unique bank deposits (j):

$$HHI_{it} = \sum_{j=1}^{N} \text{Deposits}_{jit}^{2}$$

Table 1 Summary statistics

Panel A: County characteristics

	Mean	Median	S.D.	Minimum	Maximum
Deposit HHI	0.285	0.238	0.166	0.044	1.000
MC Bank Share	0.662	0.723	0.268	0.00	1.000
Land area	934.2	641.7	1211.6	6.3	20,052.5
Number of single-county banks per county	2.239	1.000	2.861	1.000	91.000

Panel B: Single-county bank characteristics

	Mean	Median	S.D.	Minimum	Maximum
ROA (%)	0.770	0.894	1.001	−4.734	3.457
EBTROA (%)	0.957	1.128	1.179	−7.914	8.601
Age	64.077	75.000	43.076	0.000	187.000
Assets, $ millions	344.091	78.102	3945.881	1.442	215,919.762
Ag loan ratio	0.053	0.011	0.086	0.000	0.656
Commercial loan ratio	0.008	0.000	0.040	0.000	0.825
Big CD ratio	0.124	0.112	0.099	0.000	0.926
Allowance ratio	0.009	0.008	0.006	0.000	0.273
Credit loss ratio	0.003	0.001	0.006	−0.026	0.225
Demand deposit ratio	0.138	0.124	0.082	0.000	0.941
Urban	0.521	1.000	0.500	0.000	1.000

Panel C: Time characteristics

Year	Per Capita Income	Population	SCB count	ROA%	EBTROA%
2001	25,538.05	120,392.33	4838	0.926%	1.250%
2002	25,666.03	122,661.25	4624	1.033	1.361
2003	26,987.07	125,053.52	4394	1.011	1.340
2004	28,338.27	128,516.81	4233	1.068	1.385
2005	29,498.55	132,503.28	4028	1.063	1.367
2006	30,724.94	136,193.87	3889	0.991	1.264
2007	32,709.64	142,687.98	3754	0.823	1.063
2008	34,835.09	146,929.54	3581	0.285	0.401
2009	33,874.67	151,163.35	3392	−0.063	0.043
2010	34,894.65	155,090.82	3158	0.345	0.454
2011	37,675.04	157,298.94	2977	0.599	0.746
2012	39,051.22	160,595.32	2825	0.775	0.945
2013	40,022.78	163,170.99	2663	0.845	0.998
2014	41,101.59	166,233.39	2490	0.896	1.065

Summary statistics for the sample of 22,758 county-year observations and 49,839 single-county bank year observations. The sample covers years 2001–2014. MC Bank Share is the percentage of total county-level deposits made by multi-county banks. Land area is county size in square miles. ROA is return on assets. EBTROA is pre-tax ROA. All bank-level ratios are scaled by assets. Ag loan represents agricultural loans. Big CD totals all time deposits over $100,000. Allowances are for loan losses and lease losses including charge-offs and recoveries

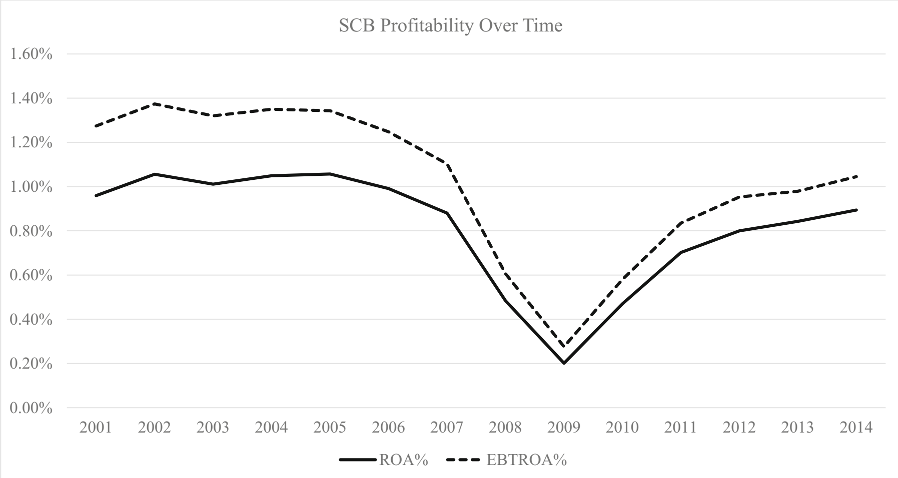

Fig. 1 Charts single-county bank ROA and EBTROA between years 2001–2014

trend to the end of our sample. These results give us the first indication of the important role market demographics play in bank performance.

Our next task is to compare profitability in the demographic extremes. If population and income play an important role in banking profits, then we should more easily see these performance differences between the banks that operate with contradistinct demographics (rich county banks versus poor county banks). Many counties overlap in both extreme income and extreme population, therefore we take into account both income and population when labeling SCBs in Table 2. SCBs that operate in counties that are in both the highest population quartile and highest income quartile are tagged as

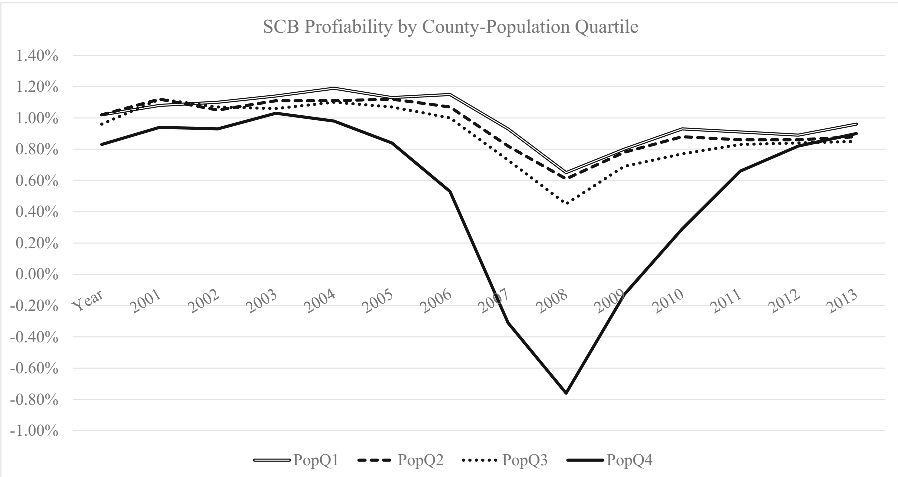

Fig. 2 Charts single-county bank profitability (ROA) between years 2001–2014. PopQ represents county-level population quartile, where PopQ1 represents low population counties and PopQ4 represents high population counties

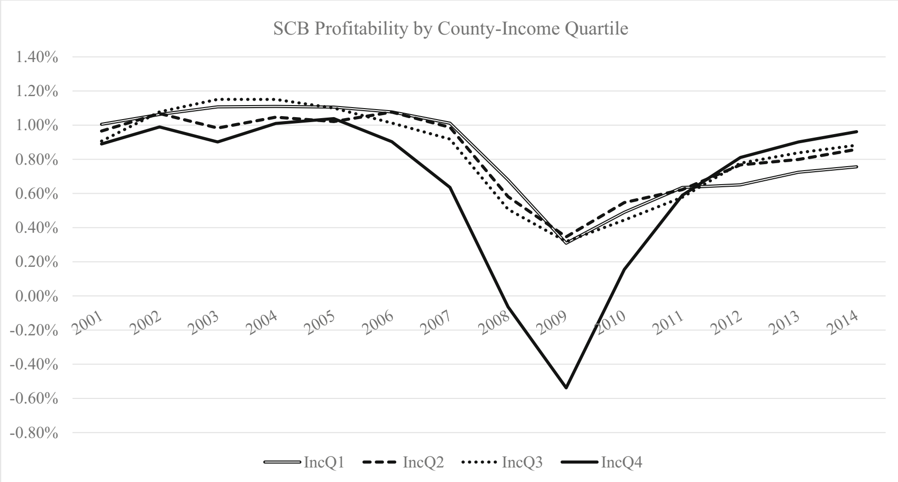

Fig. 3 Charts SCB profitability (ROA) between years 2001–2014. IncQ represents county-level income quartile, where IncQ1 represents the lowest income counties and IncQ4 represents the highest income counties

HP-HI, while SCBs operating in counties that fall into the lowest population quartile and lowest income quartile are tagged *LP-LI*. Similar groupings for high-population, low-income (*HP-LI*), and low-population, high-income (*LP-HI*) are also used in our analysis. Panel A reports the average ROA of SCBs that operate in each of the four demographic quadrants. We find that firm performance is highly dependent on market demographics. The market with the highest performing group of firms, on average, have low populations and high incomes (*LP-HI*). The market with the lowest performing group of firms are characterized by high populations and high incomes (*HP-HI*). We also see a distinct low-population advantage, as ROA is significantly higher in the low-population counties compared to their high-population counterparts. The income effect is not as clear. For high population counties, higher income produces worse returns, whereas in low population counties, higher income produces higher returns. In Panel B we test the quadrant's ROA against the full sample ROA of 0.786% as reported in Table 1, and find further evidence that the returns of SCBs in low-population areas are significantly higher than the average bank's returns. In panel C we show that the returns in each demographic quadrant is significantly different from one another. Overall, Table 2 shows that distinct performance differences exist between banks based on market demographics. Not only are these performance differences statistically significant, but they are economically meaningful – banks that operate in counties with favorable demographics (*LP-HI*) more than double the performance of firms operating with the least favorable demographics (*HP-HI*) as measured by ROA.

Firms competing in markets with larger populations or in markets with more income likely face increased competition levels, which might explain the low-population advantage found in Table 2. Another explanation of our results might be that low-income markets have riskier customers and therefore the banks in these markets are earning returns commensurate with the risk of their product mix. Thus, the question

Table 2 Univariate analysis of ROA by income and population quartiles

Panel A: Mean ROA for double sorted quartiles			
	Highest population (HP)	Lowest population (LP)	
Highest Income (HI)	$ROA_{HP,HI} = 0.459\%$ $n = 14{,}672$	$ROA_{LP,HI} = 1.027\%$ $n = 2{,}413$	
Lowest Income (LI)	$ROA_{HP,LI} = 0.708\%$ $n = 563$	$ROA_{LP,LI} = 0.952\%$ $n = 1{,}327$	
Panel B: Single sample t-tests The mean ROA from each quadrant in Panel A is tested against the mean ROA among all SCBs. Sample: n= 49,839 with mean ROA = H_0 = 0.768%.			
Single-County Banking Quadrant	Difference $ROA_{quartile} - H_0$	t-stat	Pr > |t|
HP-HI	-0.309%	29.83	<.0001
HP-LI	-0.060%	1.21	0.2255
LP-HI	0.259%	18.55	<.0001
LP-LI	0.184%	8.01	<.0001
Panel C: Two sample t-tests Pairwise comparison of mean ROA from each quadrant in Panel A			
	Difference	t-stat	Pr > |t|
Income T-Tests			
HP-HI vs. *HP-LI*	-0.249%	4.63	<.0001
LP-HI vs. *LP-LI*	0.001%	2.95	0.0032
Population T-Tests			
HP-HI vs. *LP-HI*	-0.568%	21.74	<.0001
HP-LI vs. *LP-LI*	-0.245%	5.11	<.0001
Cross-quadrant T-Tests			
HP-HI vs. *LP-LI*	-0.493%	14.05	<0.0001
HP-LI vs. *LP-HI*	-0.320%	8.51	<0.0001

Cross sectional single-county bank (SCB) return on assets performance (ROA) by high and low income and population county quartiles. Counties are ranked in quartiles across the United States from high to low per-capita income and population. ROA is expressed in percent. The sample covers years 2001–2014

remains – if we control for competition, product risk, past due loans, and other determinants of bank performance, does the low-population advantage disappear? The next section explores this research question in more detail.

4.2 Modeling SCB performance

To model ROA and EBTROA using regression analysis, we use a similar structure to that of Berger et al. (2007), Athanasoglou et al. (2008), and Hannan and Prager (2009). Our approach is to estimate:

$$\begin{aligned} \text{Profit} = f(&\text{population, income, bank characteristics, product } mix \\ &\text{variables, county characteristics, year controls}), \end{aligned} \tag{1}$$

where profit is either ROA or EBTROA, county population is measured in tens of thousands, and county income is per capita income measured in thousands.[4] We include control variables for SCB characteristics, product mix, county characteristics, and yearly fixed effects. The bank characteristics we control for are the natural log of total assets and firm age. Generally, the effect of growing assets on profitability has been positive until banks become extremely large and experience diseconomies of scale. Thus, we expect a positive coefficient on Ln(Assets) since our banks are small because they operate in a single county. We also follow Berger et al. (2007) and control for firm age. Berger et al. (2007) find that more experienced firms outperform de novo banks, and we suspect that this relationship will hold in our sample as well. We control for the mix of banking products and product risk by including proportions of agricultural loans, commercial loans, large time deposits, allowance for loan and lease losses, credit losses, and demand deposits (checking accounts) all scaled by total assets.[5] As discussed earlier, one explanation of our results is that SCBs in low income areas could be serving high-risk customers, thus the reported high returns are a result of riskier loans and securities.

The competition controls we employ are a Herfindahl index of county-level deposits (*Deposit HHI*) and multi-county bank deposit share (*MC Bank Share*). Hannan and Prager (2009) find that the presence of large multi-market banks decreases profitability for rural single-market banks, but has no influence on the profitability of single-market banks in urban markets. Therefore we control for competition that SCBs face from multi-county banks using the portion of deposits that multi-county banks have in each area. We expect *MC Bank Share* to be negatively related to SCB profitability. Another facet of market structure that may affect profitability is the physical size of the market, measured in square miles. Rural counties have been found to be more profitable than urban counties (Hannan and Prager 2009), although, in contrast, Rosen (2007) shows that population density positively affects performance. Therefore, we cannot determine a priori the influence the size of the county will have on bank performance. We scale the coefficients on *Land area* by 1000 mile^2 for ease of interpreting estimates. In addition to the physical size of the market, we also use the dummy variable *Urban* to control for whether the market is classified as urban or rural.[6]

Finally, we include yearly fixed effects to control for differences in profitability through time not captured by the other controls in place. We exclude year 2001 as the base case. We tested our model with firm fixed-effects and county fixed-effects, but

[4] In addition, for robustness we also estimated models using Net Interest Margin (NIM) and Net Non-interest margin (NNIM) as performance measures and found little to no relation with our population or income variables. NIM was correlated with ROA with a 0.261 Pearson Correlation Coefficient, and NNIM was correlated at 0.109. Given their weak correlation with the traditional measure of ROA, and their lack of explanatory power, we do not present these results. We thank an anonymous referee for suggesting this specification to investigate the relation between the sources of bank income and population and income.

[5] We use the Allowance for Loan and Lease Losses (ALLL) for a measure of future losses and the Provision for Loan and Lease Losses (PLLL) that we term "Credit Losses' to measure actual current losses. These two measures are positively correlated, but in our regression models have very low variance inflation factors below one, thus we keep each in the models to account for both current and expected risk of the lending portfolio.

[6] We use the classification scheme provided by the National Center for Health Statistics when labeling counties as urban or rural and is available at: http://www.cdc.gov/nchs/data_access/urban_rural.htm.

believe these controls were too stringent in determining market characteristics that persist over the time series such as land area, average income, and population levels.[7]

4.3 Regression results

We now discuss the results from regression models that investigate whether SCBs operating in low population markets continue to have higher accounting profits after controlling for competition levels and other determinants of bank performance found in other studies. The answer to this question has implications for the strategic choice of entry for expanding firms, market selection for new startups, and bank manager performance evaluation.

Table 3 contains the results from our full sample regression using Eq. (1). Columns one and three use *ROA%* as the dependent variable and columns two and four use *EBTROA%* as the dependent variable. We find a distinct low-income advantage as the continuous variable *County Income* is negative and significant in every specification, consistent with our univariate findings. We also find that a low-population advantage is prevalent, as the continuous variable *County Population* is negative and significant in every specification. This result is consistent with what we see in Fig. 2, that banks in the low population quartiles generally outperform banks in the upper population quartiles.

As expected, the coefficients on firm size and firm age are positive in each specification indicating larger and older SCBs have higher performance. Our county-level competition proxy, *Deposit HHI*, is not a statistically significant determinant of SCB profitability. Each product mix variable is a strong predictor of banking profits as shown by highly significant coefficients in the results. We find that higher proportions of agricultural loans, loan loss allowances, and demand deposits all improve firm performance, while larger proportions of commercial loans, large CDs, and credit losses predict lower SCB performance. *MC Bank Share* is negative and significant as the presence of big banks in each county lowers profitability of the SCB. The allowances are a measure of previous loan losses, whereas credit losses represent expected future losses due to past-due loans, thus the difference in the effect of these variables. *Land area* is positive in each model but shows limited significance, while the urban dummy variable indicates that SCBs are less profitable in urban markets than rural markets, consistent with the findings of Hannan and Prager (2009).

In columns 3 and 4 of Table 3, we follow the same county-classification system used earlier in our analysis and include the variables *HP-HI, HP-LI, LP-HI,* and *LP-LI.* Given that each of these demographically-extreme quadrants show substantial performance differences, we test whether the unique combination of extreme demographics is responsible for our earlier findings, or if the base-line effect of income and population as captured by the continuous income and population variables adequately explain these differences. We find that the SCBs in counties with high populations and high incomes (*HP-HI*) have significantly lower performance, over and above the baseline effect of the continuous population and income variables. After controlling for other determinants of profitability, firms in *HP-HI* counties underperform the average firm by

[7] Our primary results do not substantially change by using GLS fixed effects models. Both county population and county income estimates remain negative and statistically significant at the 1% level.

Table 3 Regression analysis of SCB profitability

Variables	(1) ROA%	(2) EBTROA%	(3) ROA%	(4) EBTROA%
County Income	−0.0093***	−0.0113***	−0.0092***	−0.0114***
	(−23.28)	(−24.36)	(−20.24)	(−21.60)
County Population	−0.0006***	−0.0006***	−0.0005***	−0.0006***
	(−16.46)	(−16.18)	(−15.25)	(−15.14)
HP-HI			-0.0386***	−0.0314**
			(−3.07)	(−2.16)
LP-HI			0.0842***	0.1000***
			(4.52)	(4.63)
HP-LI			-0.0612*	−0.0530
			(−1.78)	(−1.33)
LP-LI			0.0219	−0.0009
			(0.93)	(−0.03)
Ln(Assets)	0.3369***	0.4399***	0.3399***	0.4425***
	(77.89)	(87.72)	(77.67)	(87.18)
Age	0.0032***	0.0033***	0.0031***	0.0032***
	(31.59)	(28.03)	(30.88)	(27.46)
Agricultural Loans	1.2733***	1.4883***	1.2040***	1.4074***
	(25.65)	(25.85)	(23.38)	(23.56)
Commercial Loans	−0.4374***	−0.3550***	−0.4287***	−0.3416***
	(−4.48)	(−3.14)	(−4.39)	(−3.02)
Big CDs	−0.3614***	−0.5305***	−0.3591***	−0.5279***
	(−7.81)	(−9.88)	(−7.76)	(−9.83)
Allowances	16.3706***	24.4401***	16.4473***	24.5156***
	(21.95)	(28.26)	(22.05)	(28.34)
Credit Losses	−77.0860***	−94.9777***	−77.0057***	−94.9056***
	(−100.96)	(−107.26)	(−100.82)	(−107.15)
Demand Deposits	0.7720***	0.8061***	0.7833***	0.8148***
	(16.43)	(14.80)	(16.63)	(14.92)
Deposit HHI	−0.0277	0.0349	−0.0498*	0.0170
	(−0.96)	(1.05)	(−1.68)	(0.49)
MC Bank Share	−0.2543***	−0.2487***	−0.2393***	−0.2359***
	(−14.62)	(−12.34)	(−13.53)	(−11.50)
Urban	−0.1271***	−0.1304***	−0.1130***	−0.1176***
	(−13.85)	(−12.25)	(−11.62)	(−10.42)
Land Area	0.0052*	0.0045	0.0044	0.0036
	(1.87)	(1.40)	(1.56)	(1.12)
Constant	−2.5909***	−3.4390***	−2.6240***	−3.4615***
	(−48.12)	(−55.08)	(−47.84)	(−54.42)
Year Dummies	Yes	Yes	Yes	Yes
Observations	49,839	49,839	49,839	49,839
R-squared	0.3718	0.3908	0.3722	0.3912

Single-county bank (SCB) observations from the period 2001–2014. County Income is per capita income, scaled by 1000. County Population is scaled by 10,000. HP (LP) indicates that the SCB operates in a top (bottom) 25th percentile population county. HI (LI) indicates that the SC operates in a top (bottom) 25th percentile income county. HP-HI represents the intersection of high population and high income – the remaining variables follow this convention. Ln(Assets) is the natural log of total assets. Age is firm age in years. Agricultural Loans, Commercial loans, Allowances, Credit Losses, Big CDs, and Demand Deposits are scaled by assets. Big CDs represents all time deposits larger than \$100,000. Allowances are for loan losses and lease losses including charge-offs and recoveries. Deposit HHI is a concentration index of county-wide deposits. MC Bank Share represents the multi-county bank deposit share. Land area is the county square mileage scaled by 1000. T-statistics in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

3–5 percentage points.[8] We also find that firms operating in high-income counties, but with low populations (*LP-HI*) enjoy a 10 percentage point performance advantage over the average SCB in our sample. Banks in low-income counties do not show a statistically significant change from the average as shown by *HP-LI* and *LP-LI*. These results suggest that the unique combination of extreme demographic selection does explain some of the variance in firm performance. There appears to exist a more complex relationship between income and population. That is, we cannot say that banks should always avoid high-income counties, as high-incomes appear to be desirable in counties with smaller populations. Moreover, when high incomes are combined with high populations, the performance penalties appear to magnify. We suggest that further research is needed on this topic.

4.3.1 Robustness tests of the data and model

For robustness, we performed several modifications to the data and the model to help ensure that our results are not due to outliers or statistical problems. We first trimmed the two-sided tails of the data at 3% and 5% to see if outliers were creating issues. Our main results hold when trimming the data, thus we conclude that our results are indicative of the average banking market and not due to extreme observations. Additionally, we estimate the regressions using heteroskedastic-robust standard errors and our main results continue to be the same. We also remove some of the control variables and our results remain. Therefore, we believe our conclusion that as county-level population decreases, bank performance increases is not due to outliers in our data, multicollinearity, or heteroscedasticity in our model estimates for single-county banks.[9]

4.4 The 2008 financial crisis

The banking panic of 2007–2008 triggered the world-wide "great recession." The prices of many asset classes dropped substantially, equity market volatility sharply rose along with the cost of debt capital, and bank charge-offs of bad loans increased dramatically. One extension of our research is to ascertain whether market demographics had any role in mitigating or magnifying the negative effects of the financial crisis. To this end, we follow Ivashina and Scharfstein (2010) and include in our models two binary variables, *Crisis I*, which represents years 2007 and 2008 and *Crisis II*, which represents years 2009–2014. We then interact *Crisis I* and *Crisis II* with the firm's income and population demographics. These interaction variables will then show us how the marginal effect of income and population changed during the recession.

Table 4 presents the results of our regression models to study the marginal effects of income and population during the financial crisis. The coefficients on *Crisis I* and *Crisis II* are negative, as expected, since banks experienced severe reductions in accounting profits during the crisis. There is no additional marginal income effect during *Crisis I* since the interaction term *Income*Crisis I* is insignificant, but during the

[8] HP-HI ROA = −0.0386%, average firm ROA = 0.786%, therefore −0.0386/0.786 = −0.049. And 0.0314/1.025 = −0.306.

[9] We thank an anonymous referee for these suggestions. We do not present the tables here, but they are available upon request.

Table 4 The effect of population and income during the 2008 financial crisis

Variables	(1) ROA%	(2) EBTROA%
County Income	−0.0115***	−0.0141***
	(−18.16)	(−19.17)
County Population	−0.0002***	−0.0002***
	(−4.63)	(−3.47)
Crisis I	−0.0988***	−0.1857***
	(−2.71)	(−4.39)
Crisis II	−0.4325***	−0.5977***
	(−15.78)	(−18.79)
Income*Crisis I	−0.0016	−0.0018
	(−1.53)	(−1.50)
Income*Crisis II	0.0044***	0.0049***
	(5.65)	(5.34)
Population*Crisis I	−0.0007***	−0.0009***
	(−8.34)	(−8.86)
Population*Crisis II	−0.0006***	−0.0008***
	(−8.83)	(−10.15)
Ln(Assets)	0.3352***	0.4365***
	(77.63)	(87.11)
Age	0.0032***	0.0032***
	(31.46)	(27.77)
Agricultural Loans	1.2633***	1.4845***
	(25.49)	(25.81)
Commercial Loans	−0.4007***	−0.2956***
	(−4.11)	(−2.61)
Big CDs	−0.3090***	−0.4342***
	(−7.67)	(−9.29)
Allowances	16.7344***	24.7126***
	(22.55)	(28.70)
Credit Losses	−77.2893***	−94.8239***
	(−103.53)	(−109.47)
Demand Deposits	0.7402***	0.7536***
	(15.77)	(13.84)
Deposit HHI	−0.0358	0.0236
	(−1.25)	(0.71)
MC Bank Share	−0.2568***	−0.2562***
	(−14.75)	(−12.69)
Urban	−0.1222***	−0.1218***
	(−13.31)	(−11.44)
Land Area	0.0053*	0.0044
	(1.91)	(1.35)
Constant	−2.5053***	−3.3451***

Table 4 (continued)

Variables	(1) ROA%	(2) EBTROA%
	(−46.05)	(−53.00)
Observations	49,839	49,839
R-squared	0.3715	0.3912

Single-county bank (SCB) observations from the period 2001–2014. The dependent variables are return on assets (ROA) and pre-tax return on assets (EBTROA) in percent. County Income is per capita income in thousands. County Population is in tens of thousands. Crisis I is a binary variable equal to one for years 2007 and 2008. Crisis II is a binary variable equal to one for years 2009+. Crisis I and Crisis II are interacted with County Income and County Population in each model. Ln(Assets) is the natural log of total assets. Age is firm age in years. Agricultural Loans, Commercial loans, Allowances, Credit Losses, Big CDs, and Demand Deposits are scaled by assets. Big CDs represents all time deposits larger than $100,000. Allowances are for loan losses and lease losses including charge-offs and recoveries. Deposit HHI is a concentration index of county-wide deposits. MC Bank Share represents the multi-county bank deposit share. Land area is the county square mileage scaled by 1000. T-statistics in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

recovery years, higher income counties fared better than low income counties as Income*Crisis II is positive and significant. This finding is consistent with Fig. 3 where ROA in SCBs in income-quartile 4 rose sharply after 2009. We also find that higher populations exacerbated the negative effects of the financial crisis as shown by negative and significant coefficients for *Population*Crisis I* and *Population*Crisis II* in each model. Control variables in the regression are similar to the results presented in Table 3.

Collectively, these results indicate that those single-county banks operating in locations with smaller populations had better ROA, all else equal, as did SCBs in higher-income counties after 2009. Therefore, population and income mitigated some of the effects of the crisis, at least for single-county banks in our sample.

5 Summary and conclusions

The relationship between macroeconomic variables and bank profitability is a widely studied topic in the banking literature. However, prior research has not indicated a consensus on the effects of market characteristics on bank performance. Our primary research question is to what extent are the market characteristics of per capita income and population related to bank profits? To capture these effects, we use a sample of single-county banks since we can directly link total bank performance to its area's population and income levels. Our sample covers the period from 2001 to 2014. Our univariate statistics show that market demographics can lead to dramatic performance differences between banks. We find that banks in low-population counties outperform banks in high-population counties. After controlling for bank characteristics, product mix, risk, yearly fixed-effects, and market competition, our regression analysis shows that a low-population and a low-income advantage exist in these data as related to bank accounting performance. As an extension to our main result, we test for additional effects income and population on firm performance during the financial crisis. We find that higher income helped improve bank performance in the crisis recovery years after

2009. We also find that operating in counties with low populations mitigated some of the decline in bank profitability during the financial crisis. Our findings indicate future research should consider the income and population characteristics of the market where the bank operates to appropriately analyze bank performance. Additional study on whether or not our results for single-county banks generalize to the larger population of banks is another important avenue for future research.

References

Akhavein JD, Berger AN, Humphrey DB (1997) The effects of megamergers on efficiency and prices: evidence from a bank profit function. Rev Ind Organ 12:95–139

Albertazzi U, Gambacorta L (2009) Bank profitability and the business cycle. J Financ Stab 5:393–409

Athanasoglou PP, Brissimis SN, Delis MD (2008) Bank-specific, industry-specific and macroeconomic determinants of bank profitability. J Int Financ Mark Inst Money 18:121–136

Berger AN (1995) The profit-structure relationship in banking—tests of market-power and efficient-structure hypotheses. J Money Credit Bank 27:404–431

Berger AN, DeYoung R (2001) The effects of geographic expansion on bank efficiency. J Financ Serv Res 19: 163–184

Berger AN, Humphrey DB (1997) Efficiency of financial institutions: international survey and directions for future research. Eur J Oper Res 98:175–212

Berger AN, Mester LJ (1997) Inside the black box: what explains differences in the efficiencies of financial institutions? J Bank Financ 21:895–947

Berger AN, Bonime SD, Covitz DM, Hancock D (2000) Why are bank profits so persistent? The roles of product market competition, informational opacity, and regional/macroeconomic shocks. *J Bank Financ* 24:1203–1235

Berger AN, Dick AA, Goldberg LG, White LJ (2007) Competition from large, multimarket firms and the performance of small, single-market firms: evidence from the banking industry. J Money Credit Bank 39: 331–368

Bonin JP, Hasan I, Wachtel P (2005) Bank performance, efficiency and ownership in transition countries. J Bank Financ 29:31–53

Bos JWB, Kool CJM (2006) Bank efficiency: the role of bank strategy and local market conditions. J Bank Financ 30:1953–1974

Cyree KB, Paul Spurlin W (2012) The effects of big-bank presence on the profit efficiency of small banks in rural markets. J Bank Financ 36:2593–2603

DeLong G, DeYoung R (2007) Learning by observing: information spillovers in the execution and valuation of commercial bank M&as. J Financ 62:181–216

DeYoung R, Hasan I (1998) The performance of de novo commercial banks: a profit efficiency approach. J Bank Financ 22:565–587

Deyoung R, Hunter WC, Udell GF (2004) The past, present, and probable future for community banks. J Financ Serv Res 25:85–133

GonzÁLez F (2009) Determinants of bank-market structure: efficiency and political economy variables. J Money Credit Bank 41:735–754

Hannan TH, Prager RA (2004) The competitive implications of multimarket bank branching. J Bank Financ 28:1889–1914

Hannan TH, Prager RA (2009) The profitability of small single-market banks in an era of multi-market banking. J Bank Financ 33:263–271

Ivashina V, Scharfstein D (2010) Bank lending during the financial crisis of 2008. J Financ Econ 97:319–338

Kwan SH (2003) Operating performance of banks among Asian economies: an international and time series comparison. J Bank Financ 27:471–489

Pilloff SJ (1999) Multimarket contact in banking. Rev Ind Organ 14:163–182

Rosen RJ (2007) Banking market conditions and deposit interest rates. J Bank Financ 31:3862–3884

Vennet RV (2002) Cost and profit efficiency of financial conglomerates and universal banks in Europe. J Money, Credit, Bank 34:254–282