JOURNAL OF INTERNATIONAL BUSINESS DISCIPLINES

Volume 9, Number 1 May 2014

Published By:
International Academy of Business Disciplines and Frostburg State University

ISSN 1934-1822 WWW.JIBD.ORG

JOURNAL OF INTERNATIONAL BUSINESS DISCIPLINES

Volume 9, Number 1 May 2014

Chief Editor:
Ahmad Tootoonchi
College of Business
Frostburg State University
101 Braddock Road
Frostburg, Maryland 21532
Tel: 301-687-4740
Tootoonchi@frostburg.edu

Editor:
Carol Gaumer
College of Business
Frostburg State University
101 Braddock Road
Frostburg, Maryland 21532
Tel: 301-687-4052
CGaumer@frostburg.edu

Processing Manager:
Jenna Porter
College of Business
Frostburg State University
101 Braddock Road
Frostburg, Maryland 21532
Tel: 301-687-4019
JDPorter0@frostburg.edu

Published By:
International Academy of Business Disciplines and Frostburg State University

ISSN 1934-1822 WWW.JIBD.ORG

JOURNAL OF INTERNATIONAL BUSINESS DISCIPLINES

EDITORIAL BOARD

External Reviewers

Ali Ashraf, Frostburg State University

Shakil Rahman, Frostburg State University

Jo Anna Shore, Frostburg State University

JOURNAL OF INTERNATIONAL BUSINESS DISCIPLINES

Volume 9, Number 1 May 2014

Editorial Note

The May 2014 issue of the *Journal of International Business Disciplines (JIBD)* has been the result of a rigorous process in two stages:

- Stage 1: all papers that were submitted to the 2014 IABD conference went through blind reviews, and high quality papers were recommended for publication in the *Business Research Yearbook (BRY)*.
- Stage 2: approximately ten percent of the articles published in the *BRY* and one invited manuscript were selected for possible publication in *JIBD*, and the respective authors were contacted and asked to resubmit their papers for a second round of reviews. These manuscripts went through a rigorous review process by the editorial board members and external reviewers. In the end, four articles were recommended by the editorial board for publication in the May issue of *JIBD*.

JIBD is committed to maintaining high standards of quality in all of its publications.

Ahmad Tootoonchi, Chief Editor
Journal of International Business Disciplines

BANK CONSOLIDATION AND SMALL BUSINESS LENDING IN THE AFTERMATH OF THE US FINANCIAL CRISIS: EVIDENCE FROM CALIFORNIA'S BANKING MARKETS

Burak Dolar, Western Washington University
burak.dolar@wwu.edu

ABSTRACT

In the last two decades, the extensive consolidation of the banking industry, and its likely consequences of increased presence of large banking institution as well as concentration in local banking markets, have raised concerns about the ability of small businesses to obtain funds. The effects of consolidation on small business lending may be more pronounced during a credit squeeze when the supply of loans becomes increasingly scarce. Our findings suggest that small business lending growth declined in California's moderately concentrated and highly concentrated urban banking markets relative to unconcentrated ones after the financial crisis of 2008. We also find that the effect of market share structure on small business lending growth is moderated by market concentration in the post-crisis era.

INTRODUCTION

The financial crisis of 2008 is likely to leave an important legacy on the global economy for years to come. The crisis, which reached its climax in September 2008 with the downfall of major U.S. financial institutions, revealed the structural deficiencies of the global financial system and highlighted the urgency of comprehensive reforms to promote stability and soundness of financial markets. The triggering event of the crisis is deemed to be the meltdown of the market for securitized financial products, particularly subprime mortgage-backed securities. The underlying factors of the crisis are, however, varied and had taken several years to evolve into the most severe economic downturn since the Great Depression. Even though the 18-month recession had ended in June of 2009, the economic recovery has remained weak compared to previous post-recession periods.

The U.S. Small Business Administration defines small businesses as independent firms having fewer than 500 employees. However, the vast majority of small businesses are small-scale operations; as of 2009, 89.9% of small firms employed fewer than 20 employees (U.S. Census Bureau). Representing more than 99% of all employer firms, small businesses have a significant impact on the U.S. economy. They generate more than half of the non-farm gross domestic product (GDP) and employ about half of all private sector employees. Small businesses also function as the engine of job growth in the nation's economy, generating 65% of all new jobs between 1993 and 2009 (U.S. Small Business Administration, n. d.).

Having access to reliable sources of credit is critical for the viability and success of small firms. Research findings indicate that small businesses using external financing, including loans from commercial banks and other financial institutions, are more likely to achieve higher profitability, liquidity, and growth (Haynes & Brown, 2009; Cole, 2010). Unlike their larger counterparts with direct access to money and capital markets, small businesses have limited sources of external funding which may become more-scarce during times of economic and financial distress. As banking institutions remain to be the most important source of credit for small firms, the negative effects of tightening of bank credit are particularly pronounced for small businesses (According to Ou and Williams, 2009, commercial banks and thrifts held 63.8% of the total small business debt followed by finance companies with 15.9% as of 2003).

Some of the decline in bank lending in the aftermath of the 2008 crisis may be explained by decreasing demand for credit; however, the supply side factors are also likely to have played a major role. The volume of small business loans (defined as business loans under $1 million) as well as their importance for banking institutions declined significantly during and after the financial crisis. Small business lending by depository institutions declined by 8.3% from 2008 to 2010. Small business loans constituted 15.3% of total bank assets in 2010, down from 16.8% in 2005. Likewise, the small business share of total business loans declined from 81.7% to 68.9% between 2005 and 2010 (Haynes and Williams, 2011). According to the survey conducted by the National Federation of Independent Business, only half of small businesses were able to obtain all or most of the loans they needed in 2009, down from 90% in the mid-2000s (Maltby, 2010).

The extensive bank consolidation in the last few decades raised concerns over the availability of banking services to small firms. Many argue that the consolidation of the industry along with increased concentration in local banking markets resulted in a lessening of competition and intensified the market power of remaining larger banking institutions. The declining position of small banks, in consequence, may have the greatest negative impact on small businesses relying on these institutions for banking services, particularly for sources of credit. According to Avery and Samolyk (2000), small business lending is more likely to be affected by consolidation than some other banking services because it is mostly local in nature. Large banks tend to be less focused on providing services to small customers since they may not find these transactions operationally efficient (Berger et al., 1999). Small banks are better suited than their larger counterparts to lend to small firms which are less capable of providing quantifiable financial data required for standardized business loans. In this regard, relationship lending offers an effective solution to the information asymmetry problem faced by small businesses. Moreover, the general decline in the supply of bank lending during a credit crunch is likely to exacerbate the vulnerability of small businesses to limited external financing options.

The goal of this paper is to investigate the impact of bank consolidation as well as its likely consequences of increased presence of large banks and concentration in local banking markets on small business credit in the tightened credit market of the post-crisis period. We use a dataset derived from banking markets in the state of California which has the largest population and the economy, and is home to the largest number of small businesses in the United States. California

has been one of the hardest-hit states by the recent financial crisis thus, supplies the basis for a natural experiment to study the effects of the crisis. Based on the Federal Housing Finance Agency (FHFA) State House Price Index, California's housing market experienced the third-largest decline with 43.7% between 2007 and 2012 (Federal Housing Finance Agency, 2012). Comparing small business loan growth in California's banking markets before and after the financial crisis enables us to test whether bank consolidation has had an impact the availability of small business credit in the period after the crisis.

Holding other factors constant, we find that small business lending growth declined in moderately concentrated and highly concentrated urban markets relative to unconcentrated ones after the financial crisis. On the other hand, our findings do not suggest that market concentration had an impact on small business lending growth in rural banking markets in the post-crisis era. Overall, our empirical results do not indicate a direct association between small business lending growth (except for commercial real estate loans under $250,000) and banking market share structure in California's urban and rural markets in the period after the financial crisis. However, our findings do indicate that the effect of market share structure on small business lending growth is moderated by market concentration. The increased market share of the largest banking institutions had a significantly negative effect on small business lending growth in moderately concentrated and, particularly in highly concentrated urban markets in the aftermath of the financial crisis.

RELATED LITERATURE

This study draws on two different streams of literature. First, we review past studies looking at the impact of consolidation in banking on small business lending. Then, we examine the literature on the availability of small business credit during credit crunches. Most consolidation activity involves large banks which tend to have disproportionately smaller portfolios of small business loans than small banks. However, the existing literature on bank consolidation and small business lending does not provide clear evidence of the negative or positive impacts of consolidation on the supply of small business loans. A variety of factors, including type of banking markets (urban or rural); market size; level of market concentration; type of consolidation (involving small or large institutions); acquiring institutions' organizational complexity and post-consolidation lending policy; responses of new entrants and other lenders in the same market to consolidation; and the study period are likely to influence the relationship between consolidation and the availability of small business lending (see, among others, Keeton, 1996; Berger et al., 1998; Peek and Rosengren, 1998; Strahan and Weston, 1998; Berger et al., 1999; Avery and Samolyk, 2000; Samolyk and Richardson, 2003; Avery and Samolyk, 2004; and Berger et al., 2004).

Bank consolidation and small business lending

Throughout much of its history, the U.S. banking industry consisted of a large number of small institutions operating mostly in local markets. However, factors including banking deregulation, advances in financial technologies, and financial globalization have dramatically transformed the industry since the 1980s. The disappearance of many community banks (which have traditionally played a significant role in small business lending) and the dominance of the industry by large banking organizations raised concerns about the availability of credit to small businesses in the new banking landscape, given the systematic differences in small business lending policies between the two groups.

There are important differences between small and large banks regarding their small business lending loan portfolios. The research findings indicate that large banks have relatively smaller small business loan portfolios than community banks. Strahan and Weston (1996) report that the smallest banks (institutions with less than $100 million in total assets) devote 97% of their total commercial and industrial loans to small businesses whereas small business loans constitute only 17% of the total commercial and industrial loans of the largest banking institutions (those with more than $5 billion in total assets). According to Strahan and Weston (1998), the overall volume of small business lending increases monotonically with banking company size; however, as institutions grow larger, small business lending growth slows and lending to large businesses increases rapidly. In addition, the ratio of small business loans to total assets rises until banking companies reach about $300 million in total assets and declines thereafter, while the share of lending to large businesses in total assets increases (also see Berger et al., 1995; Keeton, 1995; Levonian and Soller, 1996; and Hancock et al., 2005).

Relationship-based lending plays a critical role in small business finance since most small firms find it difficult to produce formal financial records required for transaction-based loans. Small banks are better-suited to process "soft" information and make relationship loans than larger banks which are more likely to give standardized loans to businesses able to provide "hard" financial data. Berger et al. (1999) argue that large banking institutions may choose to provide fewer services (including business loans) to small customers since doing business with large and small customers at the same time requires a different set of policies and procedures and, may be scope inefficient in the eyes of large banks. It may be particularly inefficient for large, complex banks to provide relationship-based services to informationally opaque small businesses along with transaction-based services to large, informationally transparent customers. Berger et al., however, point out that small firms which are able to provide quantifiable financial data may also have access to transaction-based loans from large banks.

A number of empirical studies found that small and large banks differ significantly in the way they process and issue small business loans. The findings of Cole et al. (2004) indicate that large banks use a "cookie-cutter" approach (e.g., utilizing credit-scoring models) where the lending decision to small businesses is often based on standard, quantitative information gathered from

the borrower's financial statements. Large banking institutions are also more likely to supply small business loans when the borrower is larger, maintains formal financial records, and has a longer track record. On the other hand, in small banks, lending decisions are more likely to be made using a "character" approach. Small banks put less emphasis on financial variables but rather utilize private information about the borrower obtained through pre-existing relationships (the findings of Berger et al., 2005 suggest similar conclusions. Also see Berger and Udell, 2006 for the review of the research literature regarding the comparative advantages of small and large banking institutions in small business lending).

Small business lending during a credit crunch

By definition, a credit crunch is a general decline in the quantity of credit available to borrowers as financial institutions tighten their lending standards and/or raise the cost of credit. According to Bernanke and Lown (1991, p. 207), "a bank credit crunch [is] a significant leftward shift in the supply curve for bank loans, holding constant both the safe real interest rate and the quality of potential borrowers." Credit crunches are typically related to supply-side factors, including "more stringent examination standards, introduction of new capital regulations, the impact of loan losses on bank equity capital, and increased banker risk aversion" (Shrieves & Dahl, 1995, p. 5) (Syron, 1991; Peek & Rosengren, 1993 and 1995; Hancock & Wilcox, 1994; and Brinkmann & Horvitz, 1995, among others, look at the effects of capital regulations and regulatory enforcement on the supply of bank loans). Nevertheless, an overall contraction in lending provided by financial institutions may also be associated with decreasing demand for loans by borrowers due to declining aggregate demand and economic activity.

Small businesses are particularly vulnerable to tightening of bank credit as they have fewer options of external credit (than their larger competitors) and are more likely to rely on relationship loans from local financial institutions. In this respect, permanent (e.g., failure of the lender) or temporary (e.g., declined lending capacity of the lender) loss of a long-term relationship with a lender during an economic or financial crisis may dramatically decrease the availability of funds to a small business since it may not be easy to establish a new relationship with (or transfer the existing relationship to) a new lender. In fact, Bernanke (1983) observes that the tightening of credit during the Great Depression had the most significant impact on households, farmers, unincorporated businesses, and small corporations as these small borrowers had the highest direct or indirect reliance on bank credit. Bernanke argues that the financial crisis of the early 1930s reduced the effectiveness of intermediation between borrowers (particularly, of those aforementioned) and lenders as well as increased its cost. Thus, borrowers either were not able to obtain funds or for a given safe interest rate, they faced higher effective cost of credit when they were able to borrow.

Hancock and Wilcox (1998) analyze the effects of declines in bank capital on lending and aggregate economic activity during the credit crunch of the early 1990s. The results of Hancock and Wilcox indicate that small banks cut their lending (particularly, commercial and industrial

loans) significantly more than large banks in response to capital pressures. On the other hand, other studies looking at the early 1990s credit crunch reached different conclusions. The findings of Berger and Udell (1994) suggest that small banks reduced their lending proportionately less than the banking industry as a whole which contradicts the common perception that customers of small banks were the most adversely affected borrowers from tightening of bank credit. Likewise, based on a sample of banks operating in the state of New Jersey, Bernanke and Lown (1991) showed that large banks reduced commercial and industrial lending considerably more than small banks during the period.

Supply-side factors (along with reduced demand for loans) may have had a significant negative impact on bank credit in the aftermath of the financial crisis of 2008 given that the crisis originated and evolved in the financial sector. Ivashina and Scharfstein (2010) examine the availability of bank credit to large corporate borrowers in the period after the crisis. Their findings suggest that even though some of the decline in new lending can be attributed to decreasing demand for credit, the supply of funds may also have declined during the crisis. Ivashina and Scharfstein show that banks with more deposit financing (which is a more stable source of funding) cut their lending less than the ones relying more heavily on volatile funding sources (e.g., short-term debt). It may be possible to infer from their findings that large banks decreased lending (including small business loans) disproportionately more than small banks during and after the financial crisis since the former, on average, have a lower deposits-to-assets ratio than the latter.

EMPIRICAL MODEL AND RESULTS

Description of the dataset

Our study relies on the Reports of Condition and Income (Call Reports) and Thrift Financial Reports (TFRs) which are filed quarterly by all commercial banks and thrifts insured by the Federal Deposit Insurance Corporation (FDIC). The June 30 Call Reports and TFRs include data on the number and outstanding amount of loans to small businesses. Small business lending data are reported by original amount of loans rather than the size of the borrower. Loans to small businesses are defined as loans under $1 million. Two types of small business loans are reported by banking institutions: loans secured by nonfarm nonresidential properties or commercial real estate loans, and commercial and industrial loans. Furthermore, loans to small businesses are divided into three categories with respect to their outstanding balance: $100,000 or less, more than $100,000 through $250,000, and $250,000 through $1,000,000.

Deposit balances of individual institutions as well as county totals are gathered from the FDIC's 'Summary of Deposit Surveys,' which collect deposit balances for banking institutions as of June 30 of each year. The balanced panel dataset we exploit in the empirical work to follow consists of 368 observations from 26 urban and 20 rural banking markets in the state of California and

covers the period from 2004 through 2011. Urban and rural markets are analyzed separately because of their significantly different characteristics. We follow the convention in the literature and define urban and rural banking markets as Metropolitan Statistical Areas (MSAs) and non-MSA counties, respectively. As of 2011, 97.7% of the total population in California lives in urban markets, hence the research findings for urban markets are particularly significant. Since Call Reports and TFRs do not include data on the geographical origination of small business loans, we extracted the market-level data from a total of 9,415 observations gathered from 936 different commercial banks and thrifts which had operated branches in the state between 2004 and 2011. Due to limited data sources, our paper is one of the few studies looking at the effect of bank consolidation on small business lending at the market-level.

Variables and a priori expectations

The variables used in the regressions are described in Table 1, along with sources where data were obtained. The data are annual and all dollar amounts are converted to year 2011 constant dollars using the Consumer Price Index (CPI). Small business loans are defined as commercial real estate loans (CRE), and commercial and industrial loans (CI) under $1 million. Following the taxonomy of the Call Reports and TFRs, we stratify small business lending into two groups based on loan amounts: loans under $250,000 and loans more than $250,000 through $1 million.

In December 2008, the National Bureau of Economic Research (NBER) announced that the US economy had entered into a recession in December of the previous year. In the US financial services industry, the initial signs of the financial crisis started to emerge in 2007 with the collapse of the market for securitized financial products, particularly subprime mortgages. The crisis culminated in the second half of 2008 with the freeze of global financial markets and the downfall of a number of major US financial institutions. In our sample, the post-crisis era encompasses banking data beyond the June 30, 2008 Call Reports and TFRs. Table 2 supplies the descriptive statistics for all variables for the pre- and post-crisis periods classified by urban and rural banking markets. Tables 3 and 4 show the correlation matrices for independent variables from the urban and rural banking market samples, respectively. The results indicate that the growth rate of all sizes of small business loans showed a significant decline between the periods of 2004–2008 and 2009–2011 both in urban and rural banking markets. This decline was even more pronounced for larger loans and in urban markets. The deposit market share of the largest banking institutions is almost 20 percentage points higher in urban markets (about 64%) than rural markets (about 46%). Urban and rural banking markets in California are, on average, considered to be moderately concentrated (in which the Herfindahl-Hirschman Index [HHI] of deposit market concentration is between 0.1 and 0.18) and highly concentrated (in which the HHI of deposit market concentration is over 0.18), respectively. Also, over 70% of rural markets and only about 10% of urban markets are categorized as highly concentrated.

Following the approach by Avery and Samolyk (2004), we construct annual estimates (as of June 30) of small business lending for every institution in a given banking market and sum up each

institution's loan volume to calculate the total small business lending in the market. For example, if Bank A (A) and Bank B (B) are the only two banks operating branches (and collecting deposits) in Market K (K) then K's total small business lending (SBL) is calculated as the following: SBLK = (Deposit AK / Deposit ATotal) * SBL ATotal + (Deposit BK / Deposit BTotal) * SBL BTotal. This method assumes a reasonable correspondence between the geographical distribution of an institution's deposits and small business lending (in other words, the former is a good proxy for the latter). Avery and Samolyk conduct different tests to assess the validity of this key assumption and find that there is a reasonably high correlation (0.87 and 0.8 for urban and rural markets, respectively) between the geographical distribution of deposits and small business loans. The change in small business lending is measured as the annual percentage growth rate from the previous year. In order to limit the impact of extreme outliers, we winsorized three observations in each tail for the small business lending growth variables.

TABLE 1. DESCRIPTION OF VARIABLES

Variable Name	Description
Small business lending growth	
CRE>250k	Growth rate of CRE more than $250,000 through $1 million (Source: Call Reports).
CRE<250k	Growth rate of CRE under $250,000 (Source: Call Reports).
CI>250k	Growth rate of CI more than $250,000 through $1 million (Source: Call Reports).
CI<250k	Growth rate of CI under $250,000 (Source: Call Reports).
Event parameter	
POST	Equals 0 for observations from the years 2004 through 2008 and 1 for observations from the years 2009 through 2011.
Market share structure	
MSM	Deposit market share of medium-size institutions (total assets between $1 billion and $25 billion) for the previous year (Source: FDIC Summary of Deposits).
MSL	Deposit market share of the largest institutions (more than $25 billion in total assets) for the previous year (Source: FDIC Summary of Deposits).
Market concentration	
HHIM	Dummy variable for moderately concentrated banking markets in which the HHI of deposit market concentration from the previous year is between 0.1 and 0.18 (Source: FDIC Summary of Deposits).
HHIL	Dummy variable for highly concentrated banking markets in which the HHI of deposit market concentration from the previous year is over 0.18 (Source: FDIC Summary of Deposits).
HHINDEX	The HHI of deposit market concentration for the previous year (Source: FDIC Summary of Deposits).
CHHINDEX	Lagged growth rate of the HHI of deposit market concentration (Source: FDIC Summary of Deposits).
Market characteristics	
DEPO	Amount of deposits in a banking market (in 1000s) for the previous year (Source: FDIC Summary of Deposits).
CHDEPO	Lagged growth rate of deposits in a banking market (Source: FDIC Summary of Deposits).
INS	Lagged change in the number of institutions (Source: FDIC Summary of Deposits).
EFFRAT	Lagged change in the market average of the efficiency ratio. The ratio is defined as noninterest expense less amortization of intangible assets as a percent of net interest income plus noninterest income (Source: Call Reports).
NONCURAT	Lagged change in the market average of the noncurrent loans to loans ratio. The ratio is defined as total noncurrent loans and leases (loans and leases 90 days or more past due plus loans in nonaccrual status) as a percent of gross loans and leases (Source: Call Reports).
CAPRAT	Lagged change in the market average of the equity capital to assets ratio. The ratio is defined as total equity capital as a percent of total assets (Source: Call Reports).
POPU	Lagged growth rate of population (Source: US Census Bureau).
PCINCOME	Lagged growth rate of per capita income (Source: US Bureau of Economic Analysis).
Economic indicators	
HOUINDEX	Lagged growth rate of the House Price Index (Source: FHFA).
INVPLAN	Lagged change in the inventory plans component (percent increase minus percent decrease in the next three to six months as of June) of the Small Business Optimism Index (Source: National Federation of Independent Business Small Business Economic Trends Survey).
CAPEXPLAN	Lagged change in the capital expenditure plans component (percent planning a capital expenditure during next three to six months as of June) of the Small Business Optimism Index (Source: National Federation of Independent Business Small Business Economic Trends Survey).
MSAGDP	Lagged growth rate of the GDP in urban markets (Source: US Bureau of Economic Analysis).
UNEMP	Lagged change in the unemployment rate in rural markets (Source: California Employment Development Department).

TABLE 2. DESCRIPTIVE STATISTICS [a]

	2004–2008				2009–2011			
	Urban		Rural		Urban		Rural	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
CRE>250k	0.0600	0.2024	0.0364	0.1724	-0.0700	0.0752	-0.0501	0.1078
CRE<250k *	0.0110	0.1432	-0.0469	0.1610	-0.1203	0.1187	-0.0832	0.1566
CI>250k	0.0450	0.1811	0.0515	0.2089	-0.0711	0.1117	-0.1060	0.2189
CI<250k *	-0.0086	0.1687	-0.0230	0.1185	-0.0675	0.1041	-0.0879	0.1300
POST	0	0	0	0	1	0	1	0
MSM *	0.2449	0.1120	0.2563	0.1887	0.2472	0.1078	0.2584	0.1896
MSL *	0.6482	0.1187	0.4551	0.2468	0.6396	0.1316	0.4741	0.2182
HHIM	0.7231	0.4492			0.8077	0.3967		
MSLxHHIM	0.4809	0.3148			0.5164	0.2788		
HHIL *	0.1000	0.3012	0.7300	0.4462	0.1026	0.3054	0.7167	0.4544
MSLxHHIL *	0.0602	0.1885	0.3248	0.3020	0.0666	0.2065	0.3368	0.2967
HHINDEX *	0.1418	0.0889	0.2798	0.1373	0.1362	0.0344	0.2816	0.1489
CHHINDEX	0.0009	0.1266	-0.0074	0.0641	0.0448	0.1461	0.0180	0.1467
DEPO *	$3.07x10^7$	$6.76x10^7$	$5.77x10^5$	$5.25x10^5$	$3.21x10^7$	$7.32x10^7$	$5.43x10^5$	$4.93x10^5$
CHDEPO *	0.0467	0.1770	0.0201	0.0467	0.0021	0.0990	-0.0245	0.1029
INS	0.5231	2.0316	0.0300	0.3881	-1.1923	2.9588	-0.1000	0.3025
EFFRAT	0.0167	0.2093	0.0028	0.3036	-0.0096	0.2689	0.0073	0.0704
NONCURAT	-0.0002	0.0035	-0.0009	0.0033	0.0152	0.0079	0.0168	0.0144
CAPRAT	0.0044	0.0172	0.0038	0.0109	-0.0012	0.0136	-0.0010	0.0176
POPU *	0.0108	0.0110	0.0048	0.0109	0.0131	0.0110	0.0053	0.0212
PCINCOME *	0.0483	0.0234	0.0499	0.0340	-0.0028	0.0376	0.0226	0.0488
HOUINDEX	0.1283	0.1123	0.1298	0.0852	-0.1402	0.0796	-0.1122	0.0104
INVPLAN	-1.6	3.8931	-1.6	3.8977	0.0	2.1742	0.0	2.1785
CAPEXPLAN	-0.6	3.7346	-0.6	3.7390	-3.0	4.5755	-3.0	4.5844
MSAGDP	0.0650	0.0341			0.0025	0.0340		
UNEMP			-0.0820	0.5727			2.3450	1.3051
n	130		100		78		60	

[a] We test the following null hypothesis using a two-tailed test for all the variables: $H_0: \mu_{urban} = \mu_{rural}$. * denotes a significance level of 1less than 10%.

TABLE 3. CORRELATION MATRIX FOR THE URBAN MARKETS SAMPLE

	POST	MSM	MSL	HHIM	MSLxHHIM	HHIL	MSLxHHIL	HHINDEX	CHHINDEX	DEPO	CHDEPO
POST	1										
MSM	0.0103	1									
MSL	-0.0337	-0.7437	1								
HHIM	0.0952	-0.4061	0.1393	1							
MSLxHHIM	0.0570	-0.5521	0.4318	0.9363	1						
HHIL	0.0041	0.3424	-0.0684	-0.5880	-0.5505	1					
MSLxHHIL	0.0160	0.1830	0.0580	-0.5645	-0.5285	0.9600	1				
HHINDEX	-0.0373	0.0010	0.1543	-0.2616	-0.2362	0.6944	0.7667	1			
CHHINDEX	0.1571	-0.1934	0.2441	-0.0074	0.0237	0.1598	0.2464	0.3626	1		
DEPO	0.0098	-0.1242	0.2849	-0.3248	-0.2330	-0.0074	0.0486	-0.0756	0.0959	1	
CHDEPO	-0.1407	-0.1620	0.1933	-0.0615	-0.0548	0.2297	0.3316	0.5831	0.7568	0.0767	1
INS	-0.3260	0.1098	-0.1363	-0.0444	-0.0608	-0.0468	-0.0759	-0.0699	-0.1195	-0.1016	-0.0220
EFFRAT	-0.0548	0.0212	-0.0228	-0.0347	-0.0395	0.0199	0.0226	-0.0103	0.0205	0.0051	0.0198
NONCURAT	0.8023	0.0683	-0.1134	0.0774	0.0177	-0.0282	-0.0216	-0.0655	0.1871	-0.0384	-0.1120
CAPRAT	-0.1680	0.0600	-0.0018	0.0079	0.0094	0.0143	0.0112	0.0431	-0.0975	-0.0337	0.0026
POPU	0.1007	0.0807	-0.0447	-0.0672	-0.0786	0.2153	0.2051	0.2044	0.0056	-0.2031	0.1266
PCINCOME	-0.6444	0.0448	-0.0263	0.0134	0.0288	-0.1047	-0.1387	-0.0643	-0.4287	-0.0013	-0.0987
HOUINDEX	-0.7901	-0.0807	0.0855	-0.0104	0.0133	-0.0078	0.0140	0.0930	-0.0935	0.0415	0.2332
INVPLAN	0.2260	-0.0316	0.0091	0.0424	0.0281	0.0000	0.0136	0.0034	-0.0105	-0.0027	0.0460
CAPEXPLAN	-0.2758	-0.0164	0.0452	0.0199	0.0367	-0.0322	-0.0286	0.0470	-0.2295	0.0041	0.0208
MSAGDP	-0.6658	-0.0147	0.0439	-0.0150	0.0192	0.0094	-0.0139	0.0544	-0.2044	-0.0169	0.0773

	INS	EFFRAT	NONCURAT	CAPRAT	POPU	PCINCOME	HOUINDEX	INVPLAN	CAPEXPLAN	MSAGDP
INS	1									
EFFRAT	0.1379	1								
NONCURAT	-0.1222	0.0239	1							
CAPRAT	0.1612	0.4687	-0.1666	1						
POPU	-0.0342	-0.0199	0.0465	0.0515	1					
PCINCOME	0.1812	-0.0728	-0.5622	0.1568	-0.0735	1				
HOUINDEX	0.1459	-0.0361	-0.7748	0.1730	0.0488	0.5006	1			
INVPLAN	-0.2874	-0.0855	0.0443	-0.1047	0.1606	0.0825	-0.1183	1		
CAPEXPLAN	-0.0755	-0.1048	-0.2817	0.1541	0.0813	0.4349	0.4123	-0.1385	1	
MSAGDP	0.1548	-0.1060	-0.5949	0.1162	0.2030	0.7415	0.6475	0.0086	0.3789	1

TABLE 4. CORRELATION MATRIX FOR THE RURAL MARKETS SAMPLE

	POST	MSM	MSL	HHIL	MSLxHHIL	HHINDEX	CHHINDEX	DEPO	CHDEPO	INS
POST	1									
MSM	0.0055	1								
MSL	0.0392	-0.5149	1							
HHIL	-0.0145	-0.2947	-0.0556	1						
MSLxHHIL	0.0194	-0.5116	0.6438	0.6800	1					
HHINDEX	0.0062	-0.0592	-0.3622	0.5442	0.1138	1				
CHHINDEX	0.1196	-0.1570	0.1776	0.1350	0.2427	0.3076	1			
DEPO	-0.0329	-0.0697	0.2295	-0.5018	-0.2991	-0.5759	-0.0192	1		
CHDEPO	-0.2857	0.0390	-0.1103	0.0022	-0.0674	-0.2244	-0.4029	0.0457	1	
INS	-0.1740	0.1316	-0.1312	-0.0706	-0.1256	-0.0629	-0.3949	-0.0766	0.1690	1
EFFRAT	0.0090	0.0132	-0.0049	0.1168	0.1007	-0.0290	-0.0607	0.0156	0.0244	0.1826
NONCURAT	0.6834	-0.0077	0.0381	-0.0213	0.0178	-0.0158	0.0467	-0.0184	-0.1043	-0.1078
CAPRAT	-0.1679	0.1257	-0.0780	0.0110	-0.0456	0.0119	0.0922	0.0054	0.0779	0.1828
POPU	0.0139	0.0366	0.0519	-0.0915	0.0005	-0.0966	0.0988	-0.0079	0.0644	0.0460
PCINCOME	-0.3146	-0.0145	0.0716	0.0889	0.1014	0.0192	-0.0981	-0.0559	-0.0480	0.0177
HOUINDEX	-0.8671	-0.0412	-0.0325	-0.0047	-0.0291	-0.0018	-0.1228	0.0300	0.3002	0.1750
INVPLAN	0.2260	-0.0037	-0.0014	0.0694	0.0634	0.0309	0.0287	-0.0142	0.0337	-0.1260
CAPEXPLAN	-0.2758	-0.0386	0.0063	-0.0233	-0.0261	0.0118	-0.0111	0.0098	-0.0989	0.1456
UNEMP	0.7901	0.0496	-0.0297	-0.0232	-0.0211	-0.0249	0.0616	-0.0409	0.0053	-0.1489

	EFFRAT	NONCURAT	CAPRAT	POPU	PCINCOME	HOUINDEX	INVPLAN	CAPEXPLAN	UNEMP
EFFRAT	1								
NONCURAT	0.0486	1							
CAPRAT	0.1831	-0.1236	1						
POPU	-0.0029	-0.0564	0.1101	1					
PCINCOME	0.0705	-0.3494	-0.0471	-0.1065	1				
HOUINDEX	-0.0596	-0.6546	0.0924	0.0203	0.2049	1			
INVPLAN	0.2155	0.0465	-0.0520	0.1485	0.1613	-0.2021	1		
CAPEXPLAN	-0.0830	-0.3636	0.1765	0.1158	0.3020	0.3665	-0.1385	1	
UNEMP	0.0414	0.7116	-0.1771	-0.1004	-0.4577	-0.7924	0.1328	-0.5415	1

In order to analyze the impact of the presence of large banks (which tends to increase as a result of consolidation since larger institutions are more heavily involved in mergers and acquisitions) on small business lending, we employ the market share structure variables (i.e., MSM and MSL) which control for, respectively, the share of market deposits held by medium-size and the largest banking institutions. Our primary focus is to study the relationship between bank consolidation and small business lending growth in the aftermath of the recent financial crisis. In order to do so, we include interaction terms between the market share structure variables and the event dummy variable POST which takes a value of 0 for observations from the years 2004–2008 and a value of 1 for observations from 2009–2011.

We also use deposit market concentration variables (i.e., HHIM, HHIL, HHINDEX, and CHHINDEX) to examine the effect of higher market concentration (a likely consequence of bank consolidation since, unlike a large institution, a small bank cannot capture a large share of a local banking market due to size limitations) on the availability of small business loans in local banking markets. HHIM and HHIL are entered as dummy variables for moderately concentrated (the HHI from the previous year is between 0.1 and 0.18) and highly concentrated (the HHI from the previous year is over 0.18) banking markets, respectively. The excluded group includes unconcentrated banking markets where the HHI from the previous year is below 0.1. The variable HHIM is not included in regression tests conducted for rural banking markets where the vast majority of markets are classified as highly concentrated. In this regard, the base group consists of moderately concentrated and unconcentrated banking markets. HHINDEX denotes the HHI for the previous year and CHHINDEX is defined as the lagged growth rate of the HHI. Again, our main focus is to study any possible effects of market concentration on the supply of small business credit for the period after the financial crisis therefore, we introduce interaction terms between the variables HHIM and HHIL, and the event dummy variable POST.

As mentioned previously, the existing literature does not provide conclusive evidence on whether bank consolidation has a positive or negative impact on small business credit and a number of factors are likely to affect the relationship between them. Nevertheless, in a recent study, Berger et al. (2007) look at the effects of "market size structure" (defined as the distribution of market shares of different size classes of banks in a local baking market), an important yet overlooked aspect of market structure in the previous literature, on the supply of small business loans. The findings of Berger et al. are not consistent with finding of previous studies that large banks may have a disadvantage in lending to small firms, particularly in case of providing relationship loans to informationally opaque businesses. Their results suggest that the probability that a small firm obtains credit from a bank of a certain size stratum is roughly proportional to the local banking market share of that size stratum. Also, in general, the association between market concentration and small business lending is not studied explicitly in the literature. On a first approximation, one may suggest that increased market concentration (hence, less competition) is likely to have a negative impact on the availability of small business loans since lack of competition may reduce banks' incentives to establish relationships with and provide services to small firm. However, DeYoung et al. (1999) (among others) argue that the effect of concentration on small business lending is complex and thus indeterminate.

Small and large banks tend to behave differently in terms of their lending practices during periods of credit squeeze. Berger et al. (1999) argue that large, organizationally complex, diversified institutions may fare better in times of financial crises and, hence, continue to supply credit to small businesses. On the other hand, small, undiversified banks may have to reduce lending to small borrowers during times of financial stress. Lending by large banks may particularly be resilient during local and regional economic downturns as these institutions have more geographically diverse loan and deposit portfolios. In contrast, however, anecdotal evidence suggests that small banks may have fared better than their larger competitors in the latest financial crisis as most small institutions entered the crisis with stronger capital positions and less exposure to subprime mortgage loans.

We include a number of market characteristics variables in our empirical specifications which are likely to influence small business loan growth. DEPO denotes the amount of deposits in a banking market and serves as a proxy for market size. We do not have an *a priori* expectation for DEPO. INS is defined as the change in the number of institutions. An increase in the number of banks operating in a banking market is likely to have a positive effect on small business lending growth, *ceteris paribus*. CHDEPO, POPU, and PCINCOME are included to control for the growth rate of deposits, population, and per capita income in a banking market, respectively. Holding other factors constant, the signs on the coefficients of CHDEPO, POPU, and PCINCOME are likely to be positive. We control for the financial condition and performance of banks in a given market by entering NONCURAT, EFFRAT, and CAPRAT as explanatory variables. We expect a negative association between NONCURAT, which controls for the change in the market average of the ratio of noncurrent loans to loans, and small business lending growth, *ceteris paribus*. EFFRAT and CAPRAT denote the change in the market average of the efficiency and equity-to-asset ratios, respectively. We have no *a priori* expectations for the signs on their coefficients.

We also employ economic indicators which could affect the small business loan growth rate. HOUINDEX denotes growth rate of the House Price Index published by the FHFA. We use the Metropolitan Statistical Areas and State Nonmetropolitan Areas Indexes as of the second quarter for urban and rural markets, respectively. This variable is hypothesized to have a positive effect on small business credit growth, *ceteris paribus*, as loans secured by real-estate (including personal residences) are a major source of financing for small business owners. The variables INVPLAN (change in the inventory plans) and CAPEXPLAN (change in the capital expenditure plans), which are two of the components of the Small Businesses Optimism Index conducted by the National Federation of Independent Business, allow us to examine possible effects of credit demand by small businesses on the response variable. We hypothesize that small business loan growth is likely to accelerate due to anticipated increases in inventory and capital expenditures. The variable MSAGDP accounts for the growth rate of the GDP in urban markets. The GDP data is not available for rural markets therefore, we enter UNEMP (change in the unemployment rate) to control for the macroeconomic conditions in rural banking markets. We expect that, holding other factors constant, the effects of MSAGDP and UNEMP on small business loan growth is positive and negative, respectively.

Model specifications

We use three different specifications of a fixed effects model with robust standard errors to study the possible effects of bank consolidation on small business lending in the period after the financial crisis of 2008. We have chosen to study the effects of market share structure and market concentration on small business lending growth in two different specifications because of the high multicollinearity between the market structure interaction terms (i.e., POSTxMSM and POSTxMSL) and market concentration interaction terms (i.e., POSTxHHIM and POSTxHHIL).

In the first specification, we analyze the relationship between banking market share structure (defined in terms of deposit market share of medium-size and the largest banks) and small business lending growth, with principal focus being the post-crisis period. The specification has the following general form:

$$\text{(Small Business Lending Growth)} = f(\text{Event Dummy, Market Share Structure, Event Dummy x Market Share Structure, Market Concentration, Market Characteristics, Economic Indicators}) + \alpha + \varepsilon \quad (1)$$

where α is the market fixed effect which contains all time invariant factors and ε is the idiosyncratic error term.

The next specification tests whether or not market concentration has had a negative impact on small business lending in the aftermath of the financial crisis. Interacting the event dummy variable POST with HHIM and HHIL (dummy variables for moderately concentrated and highly concentrated banking markets, respectively) enables us to estimate the marginal change in the latter variables associated with the financial crisis. The control variables are the same as the ones in Equation (1). The model is of the following general form:

$$\text{(Small Business Lending Growth)} = f(\text{Event Dummy, Market Concentration, Event Dummy x Market Concentration, Market Share Structure, Market Characteristics, Economic Indicators}) + \alpha + \varepsilon \quad (2)$$

The last specification introduces two new interaction variables. We interact MSL (which denotes the share of market deposits held by the largest banks) first with HHIM and then with HHIL. Since our focus is the time period after the crisis, we further interact these interaction terms with POST, creating three-way interactions. The three-way interactions allow us to investigate whether the effect of large bank presence in local banking markets on small business lending is moderated by market concentration. The coefficient on POSTxMSLxHHIL, for example, indicates the change in the marginal impact of large bank presence on small business lending growth in highly concentrated banking markets after the financial crisis. We use the same control variables as the previous specifications. The specification takes the following form:

(Small Business Lending Growth) = f (Event Dummy, Market Share Structure, Market Concentration, Market Share Structure x Market Concentration, Event Dummy x Market Share Structure x Market Concentration, Market Characteristics, Economic Indicators) + α + ε (3)

Results

The estimated fixed effects specifications generate the regression results reported in Tables 5–7. They explain between 19.6% and 46.4% of the variation in the small business lending growth in California's urban and rural banking markets. Table 5 includes results based on the specification in Equation (1) where we test whether banking market structure has an impact on small business loan growth. Table 6 presents the findings on the possible effects of market concentration on small business lending as specified in Equation (2). Finally, the results from estimating Equation (3) are presented in Table 7. For the sake of brevity, we do not discuss the details of the estimated coefficients on the control variables.

TABLE 5. EFFECTS OF MARKET SHARE STRUCTURE ON SMALL BUSINESS LENDING GROWTH [b]

	Urban Markets				Rural Markets			
	CRE>250k	CRE<250k	CI>250k	CI<250k	CRE>250k	CRE<250k	CI>250k	CI<250k
Intercept	-0.599	-0.844**	-0.471	-1.325***	0.490**	-0.008	0.331	-0.338
	0.553	*0.326*	*0.329*	*0.364*	*0.222*	*0.325*	*0.284*	*0.305*
POST	0.018	0.621***	0.008	0.266**	0.147	0.265**	-0.083	0.020
	0.142	*0.165*	*0.138*	*0.101*	*0.091*	*0.120*	*0.111*	*0.066*
MSM	0.837	0.520	0.303	1.548***	-0.203	0.020	-0.493	0.247
	0.506	*0.524*	*0.362*	*0.419*	*0.255*	*0.448*	*0.305*	*0.225*
POSTxMSM	-0.265	-0.381	-0.019	-0.239	-0.059	-0.163	-0.148	-0.241*
	0.263	*0.254*	*0.224*	*0.210*	*0.186*	*0.207*	*0.245*	*0.118*
MSL	0.919	1.232***	0.931*	1.571***	-0.112	-0.495	-0.805**	-0.126
	0.850	*0.327*	*0.504*	*0.530*	*0.291*	*0.364*	*0.348*	*0.296*
POSTxMSL	-0.127	-0.661***	-0.017	-0.109	-0.147	-0.392***	0.071	0.006
	0.174	*0.179*	*0.140*	*0.114*	*0.100*	*0.120*	*0.151*	*0.130*
HHM	0.019	-0.105**	-0.054	-0.060**				
	0.034	*0.039*	*0.040*	*0.024*				
HHL	0.111	-0.039	0.207*	0.139	-0.078**	0.079	-0.026	0.003
	0.114	*0.066*	*0.114*	*0.110*	*0.033*	*0.092*	*0.065*	*0.066*
HHINDEX	-0.435	-0.346	-0.877***	-0.503	0.183	0.684	1.379***	0.384
	0.539	*0.261*	*0.298*	*0.420*	*0.316*	*0.551*	*0.393*	*0.281*
CHHINDEX	-0.092	0.142	-0.169	0.118	0.075	0.201	-0.085	-0.031
	0.151	*0.163*	*0.123*	*0.126*	*0.220*	*0.268*	*0.209*	*0.119*
DEPO	-4.50×10^{-9}	-4.34×10^{-9}**	-4.54×10^{-9}*	-2.96×10^{-9}	-8.08×10^{-7}***	-3.17×10^{-7}	-4.81×10^{-7}	2.45×10^{-7}
	2.93×10^{-9}	*1.68×10^{-9}*	*2.45×10^{-9}*	*2.68×10^{-9}*	*2.66×10^{-7}*	*3.63×10^{-7}*	*3.97×10^{-7}*	*3.79×10^{-7}*
CHDEPO	0.112	-0.095	0.134	-0.060	-0.049	-0.648**	-0.326	-0.488**
	0.169	*0.151*	*0.099*	*0.194*	*0.289*	*0.252*	*0.347*	*0.216*
INS	-0.003	0.004	0.002	0.012***	0.027	-0.007	0.038	-0.018
	0.005	*0.004*	*0.002*	*0.004*	*0.029*	*0.023*	*0.043*	*0.027*
EFFRAT	0.004	-0.012	0.001	0.018	0.024	-0.003	-0.055	-0.014
	0.043	*0.052*	*0.055*	*0.027*	*0.020*	*0.026*	*0.047*	*0.020*
NONCURAT	0.495	-1.416	-0.467	-1.842	-0.038	-0.316	4.663*	1.025
	1.257	*1.568*	*2.554*	*1.578*	*1.040*	*1.440*	*2.664*	*1.952*
CAPRAT	0.209	-0.568	-0.008	-0.244	-2.052***	-3.139***	-1.594	-0.056
	0.675	*0.551*	*0.767*	*0.416*	*0.531*	*0.792*	*1.035*	*0.696*
POPU	1.730	1.550	1.300	-0.699	0.630	0.154	-1.403	-0.994
	1.129	*1.057*	*1.302*	*1.162*	*0.726*	*0.863*	*0.928*	*0.809*
PCINCOME	-0.417	2.478***	1.650**	2.356***	0.814***	0.882***	0.993**	0.495
	0.539	*0.553*	*0.764*	*0.523*	*0.284*	*0.286*	*0.458*	*0.290*
HOUINDEX	-0.038	0.358**	0.054	0.259	0.551**	0.712**	0.677*	0.385**
	0.114	*0.132*	*0.243*	*0.156*	*0.231*	*0.251*	*0.381*	*0.165*
INVPLAN	-0.008	-0.011***	-0.007*	-0.011***	-0.003	-0.001	0.002	0.001
	0.006	*0.003*	*0.004*	*0.003*	*0.003*	*0.004*	*0.006*	*0.003*
CAPEXPLAN	-0.001	-0.012***	-0.012**	-0.007**	-0.015***	-0.008**	-0.018***	-0.005*
	0.005	*0.002*	*0.005*	*0.003*	*0.005*	*0.003*	*0.006*	*0.003*
MSAGDP	0.341	-0.175	-0.092	-0.826**				
	0.605	*0.379*	*0.399*	*0.303*				
UNEMP					-0.025	0.025	-0.010	0.016
					0.024	*0.022*	*0.037*	*0.017*
R^2 (with-in)	0.196	0.446	0.267	0.340	0.366	0.361	0.338	0.238
n; # of groups	208; 26	208; 26	208; 26	208; 26	160; 20	160; 20	160; 20	160; 20

[b] Standard errors are reported in italics beneath the estimated coefficients.
***, **, and * denote 1%, 5%, and 10% significance, respectively.

TABLE 6. EFFECTS OF MARKET CONCENTRATION ON SMALL BUSINESS LENDING GROWTH [c]

	Urban Markets				Rural Markets			
	CRE>250k	CRE<250k	CI>250k	CI<250k	CRE>250k	CRE<250k	CI>250k	CI<250k
Intercept	-0.483	-0.423	-0.394	-1.227***	0.519**	0.063	0.462*	-0.173
	0.555	*0.313*	*0.297*	*0.379*	*0.221*	*0.300*	*0.251*	*0.296*
POST	-0.022	0.224***	0.065	0.226***	0.041	-0.013	-0.151	-0.117
	0.066	*0.053*	*0.084*	*0.060*	*0.066*	*0.065*	*0.115*	*0.069*
HHIM	0.038	-0.062	-0.037	-0.045*				
	0.033	*0.050*	*0.032*	*0.022*				
POSTxHHIM	-0.097**	-0.124***	-0.065**	-0.082***				
	0.037	*0.040*	*0.024*	*0.028*				
HHIL	0.290*	0.096	0.392***	0.290**	-0.090*	0.050	-0.056	-0.035
	0.154	*0.083*	*0.136*	*0.139*	*0.046*	*0.089*	*0.064*	*0.066*
POSTxHHIL	-0.255**	-0.248***	-0.258***	-0.215**	0.030	0.075	0.068	0.088*
	0.121	*0.086*	*0.092*	*0.091*	*0.052*	*0.065*	*0.057*	*0.044*
HHINDEX	-0.677	-0.330	-1.218***	-0.708*	0.124	0.533	1.458***	0.467*
	0.443	*0.241*	*0.276*	*0.385*	*0.298*	*0.496*	*0.332*	*0.261*
CHHINDEX	-0.176	-0.007	-0.268**	0.048	0.072	0.190	-0.103	-0.057
	0.152	*0.148*	*0.126*	*0.127*	*0.223*	*0.245*	*0.202*	*0.117*
MSM	0.725	0.281	0.217	1.453***	-0.165	0.113	-0.571**	0.147
	0.491	*0.481*	*0.343*	*0.410*	*0.222*	*0.406*	*0.233*	*0.198*
MSL	0.774	0.638*	0.841*	1.451**	-0.077	-0.403	-0.832**	-0.149
	0.844	*0.343*	*0.447*	*0.546*	*0.277*	*0.345*	*0.344*	*0.286*
DEPO	-4.23×10^{-9}*	-4.94×10^{-9}***	-3.75×10^{-9}**	-2.74×10^{-9}	-8.50×10^{-7}***	-4.18×10^{-7}	-6.44×10^{-7}	3.85×10^{-8}
	2.08×10^{-9}	*1.57×10^{-9}*	*1.76×10^{-9}*	*1.94×10^{-9}*	*2.70×10^{-7}*	*2.78×10^{-7}*	*3.84×10^{-7}*	*3.80×10^{-7}*
CHDEPO	0.176	0.023	0.208	-0.008	0.035	-0.427*	-0.370	-0.497**
	0.170	*0.144*	*0.122*	*0.194*	*0.275*	*0.213*	*0.312*	*0.191*
INS	-0.002	0.006	0.003	0.013***	0.028	-0.005	0.033	-0.022
	0.005	*0.004*	*0.002*	*0.004*	*0.027*	*0.026*	*0.044*	*0.028*
EFFRAT	-0.011	-0.015	-0.014	0.005	0.031	0.014	-0.045	-0.001
	0.045	*0.043*	*0.052*	*0.026*	*0.025*	*0.032*	*0.048*	*0.023*
NONCURAT	0.277	-0.752	-0.739	-2.028	-0.235	-0.828	4.916*	1.262
	1.026	*1.679*	*2.520*	*1.554*	*1.014*	*1.459*	*2.713*	*1.872*
CAPRAT	0.500	-0.296	0.300	0.001	-2.295***	-3.766***	-1.546	-0.080
	0.669	*0.538*	*0.811*	*0.429*	*0.551*	*0.813*	*1.032*	*0.722*
POPU	1.720	1.058	1.452	-0.717	0.447	-0.312	-1.179	-0.757
	1.075	*1.001*	*1.229*	*1.127*	*0.664*	*0.798*	*0.887*	*0.775*
PCINCOME	-0.470	2.342***	1.544*	2.316***	0.768**	0.768**	0.953*	0.439
	0.513	*0.612*	*0.753*	*0.530*	*0.306*	*0.325*	*0.479*	*0.315*
HOUINDEX	-0.042	0.377***	0.019	0.257	0.524**	0.639**	0.664*	0.353**
	0.106	*0.120*	*0.228*	*0.157*	*0.225*	*0.232*	*0.353*	*0.163*
INVPLAN	-0.007	-0.010***	-0.007	-0.011***	-0.003	-0.001	0.002	0.001
	0.007	*0.004*	*0.004*	*0.003*	*0.003*	*0.004*	*0.005*	*0.003*
CAPEXPLAN	-0.001	-0.012***	-0.012**	-0.008**	-0.015***	-0.007**	-0.018***	-0.005
	0.005	*0.002*	*0.005*	*0.003*	*0.005*	*0.003*	*0.006*	*0.003*
MSAGDP	0.311	-0.222	-0.095	-0.852***				
	0.606	*0.403*	*0.408*	*0.307*				
UNEMP					-0.027	0.020	-0.011	0.013
					0.023	*0.020*	*0.033*	*0.016*
R^2 (with-in)	0.213	0.443	0.290	0.356	0.361	0.321	0.337	0.233
n; # of groups	208; 26	208; 26	208; 26	208; 26	160; 20	160; 20	160; 20	160; 20

[c] Standard errors are reported in italics beneath the estimated coefficients. ***, **, and * denote 1%, 5%, and 10% significance, respectively.

The coefficients on MSM and MSL in Table 5 indicate the marginal external effect of replacing the market share of banks in the base group (i.e., small banks with less than $1 billion in total assets) with market shares of medium-size and large banks (see Rosen, 2007 for a similar interpretation of banking market structure variables). In all four regressions estimated for urban markets, the coefficients on MSM (significant in one regression) and MSL (significant in three regressions) are positive providing some evidence that an increase in the market shares of the largest banks and medium-size banks, with an equal decline in the market share of small banks,

had a positive impact on small business lending growth in the period before the financial crisis, *ceteris paribus*. The findings for rural markets are, however, inconclusive as the estimated coefficients on MSM and MSL are mostly negative but not significant.

TABLE 7. EFFECTS OF MARKET SHARE STRUCTURE ON SMALL BUSINESS LENDING GROWTH MODERATED BY MARKET CONCENTRATION [d]

	Urban Markets				Rural Markets			
	CRE>250k	CRE<250k	CI>250k	CI<250k	CRE>250k	CRE<250k	CI>250k	CI<250k
Intercept	-0.669	-0.552	-0.035	-1.299***	0.612**	-0.332	0.313	-0.378
	0.541	*0.397*	*0.323*	*0.415*	*0.234*	*0.257*	*0.304*	*0.290*
POST	-0.055	0.253***	0.063	0.202***	0.087	0.073	-0.147	-0.102
	0.077	*0.052*	*0.088*	*0.056*	*0.056*	*0.070*	*0.097*	*0.059*
MSM	0.739	0.327	0.197	1.455***	-0.157	-0.052	-0.587***	0.112
	0.500	*0.468*	*0.354*	*0.414*	*0.234*	*0.408*	*0.206*	*0.201*
MSL	1.055	0.858	0.278	1.575**	-0.411	0.031	-0.513	0.238
	0.835	*0.542*	*0.485*	*0.601*	*0.280*	*0.348*	*0.446*	*0.287*
HHIM	0.057	0.078	-0.342***	-0.019				
	0.143	*0.224*	*0.111*	*0.150*				
HHIL	0.556	0.066	-0.032	0.357	-0.275*	0.425*	0.145	0.222
	0.419	*0.271*	*0.230*	*0.263*	*0.159*	*0.220*	*0.186*	*0.144*
HHINDEX	-0.889*	-0.553**	-1.473***	-0.923**	0.129	0.665	1.446***	0.468*
	0.441	*0.207*	*0.256*	*0.413*	*0.320*	*0.526*	*0.296*	*0.258*
CHHINDEX	-0.151	0.038	-0.232*	0.072	0.041	0.249	-0.075	-0.022
	0.151	*0.149*	*0.128*	*0.128*	*0.236*	*0.262*	*0.212*	*0.111*
MSLxHHIM	-0.065	-0.231	0.488**	-0.064				
	0.227	*0.360*	*0.184*	*0.237*				
POSTxMSLxHHIM	-0.087	-0.224***	-0.089**	-0.077				
	0.061	*0.056*	*0.042*	*0.045*				
MSLxHHIL	-0.337	0.151	0.832***	-0.018	0.417	-0.640	-0.403	-0.499
	0.441	*0.443*	*0.297*	*0.322*	*0.334*	*0.419*	*0.388*	*0.340*
POSTxMSLxHHIL	-0.355**	-0.512***	-0.535***	-0.356***	-0.035	-0.092	0.135	0.141
	0.162	*0.114*	*0.143*	*0.125*	*0.084*	*0.128*	*0.098*	*0.106*
DEPO	$-2.98x10^{-9}$	$-4.81x10^{-9}$***	$-3.06x10^{-9}$*	$-1.94x10^{-9}$	$-7.68x10^{-7}$**	$-1.99x10^{-7}$	$-6.45x10^{-7}$	$6.61x10^{-8}$
	$2.01x10^{-9}$	*1.70x*	*$1.69x10^{-9}$*	*$1.72x10^{-9}$*	*$2.90x10^{-7}$*	*$2.78x10^{-7}$*	*$3.78x10^{-7}$*	*$3.70x10^{-7}$*
CHDEPO	0.168	-0.011	0.185	-0.020	0.047	-0.543**	-0.328	-0.462**
	0.171	*0.137*	*0.126*	*0.196*	*0.289*	*0.223*	*0.327*	*0.211*
INS	-0.002	0.006	0.001	0.012***	0.036	0.004	0.033	-0.021
	0.005	*0.004*	*0.002*	*0.004*	*0.026*	*0.032*	*0.042*	*0.028*
EFFRAT	-0.011	-0.020	-0.018	0.005	0.012	0.006	-0.033	0.010
	0.044	*0.047*	*0.051*	*0.028*	*0.027*	*0.027*	*0.050*	*0.021*
NONCURAT	0.312	-1.321	-1.045	-2.122	-0.182	-0.568	4.714*	1.064
	1.119	*1.595*	*2.496*	*1.526*	*1.105*	*1.474*	*2.717*	*1.847*
CAPRAT	0.502	-0.271	0.390	0.009	-2.075***	-3.458***	-1.768	-0.286
	0.688	*0.553*	*0.822*	*0.434*	*0.583*	*0.810*	*1.140*	*0.689*
POPU	1.922	1.279	1.779	-0.589	0.532	0.096	-1.312	-0.859
	1.204	*1.035*	*1.302*	*1.223*	*0.647*	*0.757*	*0.937*	*0.906*
PCINCOME	-0.446	2.419***	1.586**	2.346***	0.858**	0.858**	0.908*	0.403
	0.510	*0.582*	*0.743*	*0.524*	*0.316*	*0.341*	*0.486*	*0.338*
HOUINDEX	-0.031	0.379***	0.008	0.258	0.524**	0.680***	0.650*	0.341*
	0.106	*0.119*	*0.226*	*0.156*	*0.236*	*0.237*	*0.353*	*0.169*
INVPLAN	-0.007	-0.011***	-0.007	-0.011***	-0.003	-0.001	0.003	0.001
	0.007	*0.004*	*0.004*	*0.003*	*0.003*	*0.004*	*0.006*	*0.003*
CAPEXPLAN	-0.001	-0.012***	-0.013**	-0.008**	-0.015***	-0.008***	-0.018***	-0.005*
	0.005	*0.002*	*0.005*	*0.003*	*0.005*	*0.003*	*0.006*	*0.003*
MSAGDP	0.310	-0.212	-0.094	-0.840**				
	0.626	*0.375*	*0.405*	*0.308*				
UNEMP					-0.029	0.025	-0.011	0.014
					0.023	*0.020*	*0.034*	*0.017*
R^2 (with-in)	0.212	0.464	0.301	0.357	0.365	0.333	0.341	0.244
n; # of groups	208; 26	208; 26	208; 26	208; 26	160; 20	160; 20	160; 20	160; 20

[d] Standard errors are reported in italics beneath the estimated coefficients. ***, **, and * denote 1%, 5%, and 10% significance, respectively.

In regressions, we estimated for urban markets, the coefficients on POSTxMSM and POSTxMSL are all negative yet insignificant in all but one of the cases. In the second column of Table 5, the statistically significant coefficient on POSTxMSL implies that a 1% increase in the market share of the largest banks (accompanied by a 1% decline in the market share of small banks) results in, on average, 0.67 percentage points decline in the growth rate of CRE under $250,000 in the post-crisis era, *ceteris paribus*. The coefficients on POSTxMSM and POSTxMSL in regressions estimated for rural markets are mostly negative but again insignificant in most cases (except in regressions where CRE<250k and CI<250k are entered as dependent variables). Overall, the findings do not provide enough evidence to suggest an association between banking market structure and small business loan growth in the aftermath of the financial crisis in both urban and rural markets in California.

In regressions, we estimated using California's urban banking market data (see Table 6), the coefficients on the concentration variable HHIL are all positive and statistically significant in most cases suggesting that small business lending growth in highly concentrated markets was higher relative to the excluded group, namely unconcentrated banking markets before the financial crisis, *ceteris paribus*. The findings for moderately concentrated markets are not conclusive as the coefficients on HHIM have mixed signs and are not statistically significant. Finally, the coefficients on HHIL are mostly negative but insignificant in the regression tests conducted for rural markets.

The coefficients on POSTxHHIM and POSTxHHIL are all negative and statistically significant at conventional levels in estimated regressions. These findings indicate that, holding other factors constant, the small business lending growth rate (for all loan categories) in moderately concentrated, and particularly in highly concentrated banking markets declined compared to unconcentrated banking markets in the aftermath of the financial crisis. In first column of Table 6, the coefficients on POSTxHHIM and POSTxHHIL indicate that the growth rate of CRE more than $250,000 through $1 million in moderately concentrated and highly concentrated markets declined by, on average, 9.7 and 25.5 percentage points, respectively, relative to unconcentrated markets in the period after the crisis, *ceteris paribus*. The results for rural banking markets, however, do not suggest any association between small business lending growth and market concentration since the coefficients on POSTxHHIL are only marginally significant in one of the regressions.

In regressions, we estimated both for urban and rural banking markets (see Table 7), the coefficients on MSLxHHIM and MSLxHHIL have mixed signs and they are mostly statistically insignificant. These results do not provide any evidence suggesting that market concentration has a moderating effect on the association between market share structure and small business lending in the pre-crisis-era. On the other hand, in regressions estimated for urban markets, the coefficients on POSTxMSLxHHIM (significant in two regressions) and POSTxMSLxHHIL (significant in all regressions) are all negative indicating that, holding other factors constant, the relationship between market share structure of local banking markets and small business lending growth is moderated by market concentration in the period after the financial crisis. In other

words, in moderately concentrated and, particularly in highly concentrated markets, the increased market share of the largest banking institutions has a significantly negative effect on small business lending growth after the crisis. For example, in the first column of Table 7, the coefficient on POSTxMSLxHHIL implies that in highly concentrated markets, the growth rate of CRE more than $250,000 through $1 million declines by 0.36 percentage points when the market share of the largest banks increases by 1% (accompanied by a 1% decline in the market share of small banks) in the aftermath of the financial crisis of 2008, *ceteris paribus*. It is important to note that the findings from Table 5 (where MSL is not interacted with the market concentration variables) do not imply a significant negative relationship between banking market structure and small business lending in the post-crisis era. Thus, the negative effect of the increased market share of the largest banks on small business loan growth becomes more evident in California's more concentrated banking markets. In regressions estimated using observations from rural markets, the coefficients on POSTxMSLxHHIL do not yield similar results since they have mixed signs and are all insignificant.

In order to examine the sensitivity of our main findings to alternative specifications, we performed robustness tests (the results are not shown in tables due to space constraints). First, we estimated all three specifications without control variables (i.e., variables controlling for market characteristics and economic indicators). In regressions estimated for both urban and rural markets, the signs of the coefficients on the key variables (i.e., POSTxMSM and POSTxMSL from Table 5; POSTxHHIM and POSTxHHIL from Table 6; and POSTxMSLxHHIM and POSTxMSLxHHIL from Table 7) did not change. In regressions estimated for urban markets, the statistical significance of the coefficients on POSTxMSL, POSTxHHIL, and POSTxMSLxHHIL in general remained unchanged. On the other hand, the statistical significance of the coefficients on POSTxHHIM and POSTxMSLxHHIM declined, and POSTxMSM improved. The statistical significance of the coefficients on variables POSTxMSM, POSTxMSL, POSTxHHIL, and POSTxMSLxHHIL in regressions estimated for rural markets mostly did not change.

Next, we estimated the original specifications by using pooled ordinary least squares (OLS) (with and without robust standard errors as well as with standardized beta coefficients) and robust regression (to correct for influential outliers) analyses. In regressions estimated using observations from both urban and rural markets, the signs of the coefficients on the key variables remained the same except in a few cases. In regressions estimated for urban markets, the statistical significance of the coefficients on POSTxMSM generally improved and POSTxMSL remained the same. However, there was a general decline in the statistical significance of the coefficients on variables POSTxHHIM, POSTxHHIL, POSTxMSLxHHIM, and POSTxMSLxHHIL. There was no appreciable change in the statistical significance of the coefficients on the key variables in regressions estimated for rural markets.

We then estimated all the specifications by combining the samples for urban and rural banking markets. Since the variable HHIM is not included in regression tests conducted for rural banking markets, we had to exclude it from the tests using the combined samples. Overall, regression

results were inconclusive as the coefficients on the key variables had mixed signs and they were mostly statistically insignificant, presumably because urban and rural markets have significantly different characteristics, as frequently cited in the research literature on the banking industry. Finally, we estimated our specifications using six new response variables, namely growth rate of CRE and CI under $100,000, more than $100,000 through $250,000, and under $1 million. The results were generally similar to those presented in Tables 5–7. Overall, the results from the robustness tests support our main findings.

CONCLUDING REMARKS

The U.S. banking industry experienced a widespread consolidation in the last three decades which raised concerns over the availability of banking services to small businesses. The issue becomes particularly important during a credit crunch when bank credit tends to become less accessible for small firms which typically have limited financing alternatives.

Using market-level data derived from banking markets in the state of California, we study the possible effects of market consolidation on the availability of small business loans in the aftermath of the financial crisis of 2008. More specifically, we test whether or not increased concentration in local banking markets as well as greater market presence of large banks had a negative impact on the small business credit supply in the period after the crisis. This paper is one of the few studies looking at the relationship between bank consolidation and small business lending at the market level. It contributes to the literature on the effects of bank consolidation on small business lending in local banking markets as well as the literature looking at the availability of bank credit to small businesses during a credit crunch.

This study has significant policy implications. Our empirical results suggest that small business loan growth declined in highly concentrated and moderately concentrated urban markets relative to unconcentrated ones in the post-crisis era. The results for rural markets, however, do not indicate an association between market concentration and small business loan growth in the period after the crisis. Overall, the empirical evidence do not suggest a direct relationship between banking market share structure and small business lending growth in the aftermath of the financial crisis. However, we find evidence suggesting that market concentration moderates the effect of market share structure on small business lending growth. In this respect, some of the concerns over the adverse effects of market consolidation on small business lending seem to be valid, especially during a credit squeeze.

ACKNOWLEDGEMENT

I benefited from comments by Gokhan Karahan who added considerable value to the paper. As is customary, however, the author accepts full responsibility for any remaining errors.

REFERENCES

Avery, R. B. & Samolyk, K. A. (2004). Bank Consolidation and Small Business Lending: The Role of Community Banks, *Journal of Financial Services Research*, 25, 291–325.

Avery, R. B. & Samolyk, K. A. (2000). Bank Consolidation and the Provision of Banking Services: The Case of Small Commercial Loans, *Federal Deposit Insurance Corporation Working Paper* Series, 00-01.

Berger, A. N., Rosen, R. J. & Udell, G. F. (2007). Does Market Size Structure Affect Competition? The Case of Small Business Lending, *Journal of Banking and Finance*, 31, 11–33.

Berger, A. N. & Udell, G. F. (2006). A More Complete Conceptual Framework for SME Finance, *Journal of Banking and Finance*, 30, 2945–2966.

Berger, A. N., Miller, N. H., Petersen, M. A., Rajan, R. G. & Stein, J. C. (2005). Does Function Follow Organizational Form? Evidence from the Lending Practices of Large and Small Banks, *Journal of Financial Economics*, 76, 237–269.

Berger, A. N., Bonime, S. D., Goldberg, L. G. & White, L. J. (2004). The Dynamics of Market Entry: The Effects of Mergers and Acquisitions on Entry in the Banking Industry, *Journal of Business*, 77, 797–834.

Berger, A. N., Demsetz, R. S. & Strahan, P. E. (1999). The Consolidation of the Financial Services Industry: Causes, Consequences, and Implications for the Future, *Journal of Banking and Finance*, 23, 135–194.

Berger, A. N., Saunders, A., Scalise, J. M. & Udell, G. F. (1998). The Effects of Bank Mergers and Acquisitions on Small Business Lending, *Journal of Financial Economics*, 50, 187–229.

Berger, A. N., Kashyap, A. K. & Scalise, J. M. (1995). The Transformation of the U.S. Banking Industry: What a Long, Strange Trip It's Been, *Brookings Papers on Economic Activity*, 2, 155–218.

Berger, A. N. & Udell, G. F. (1994). Did Risk-Based Capital Allocate Bank Credit and Cause a 'Credit Crunch' in the U.S.?, *Journal of Money, Credit, and Banking*, 26, 585–628.

Bernanke, B. S. & Lown, C. S. (1991). The Credit Crunch, *Brookings Papers on Economic Activity*, 2, 205–247.

Bernanke, B. S. (1983). Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression, *The American Economic Review*, 73, 257–276.

Brinkmann, E. J. & Horvitz, P. M. (1995). Risk-Based Capital Standards and the Credit Crunch, *Journal of Money, Credit and Banking*, 27, 848–863.

Cole, R. (2010). Bank Credit, Trade Credit or No Credit: Evidence from the Surveys of Small Business Finances, *US Small Business Administration Office of Advocacy, Banking and Financial Economic Research*, June.

Cole, R., Goldberg, L. G. & White, L. J. (2004). Cookie-Cutter versus Character: The Micro Structure of Small Business Lending by Large and Small Banks, *Journal of Financial and Quantitative Analysis*, 39, 227–251.

DeYoung, R., Goldberg, L. G. & White, L. J. (1999). Youth, Adolescence and Maturity of Banks: Credit Availability to Small Business in an Era of Banking Consolidation, *Journal of Banking and Finance*, 23, 463–492.

Federal Housing Finance Agency. (2012). State HPI Summary, Available from <http://www.fhfa.gov/Default.aspx?Page=215&Type=summary>.

Hancock, D., Peek, J. & Wilcox, J. A. (2005). The Effects of Mergers and Acquisitions on Small Business Lending by Large Banks, *US Small Business Administration Office of Advocacy, Banking and Financial Economic Research*, March.

Hancock, D. & Wilcox, J. A. (1998). The 'Credit Crunch' and the Availability of Credit to Small Business, *Journal of Banking and Finance*, 22, 983–1014.

Hancock, D. & Wilcox, J. A. (1994). Bank Capital and the Credit Crunch: The Roles of Risk-Weighted and Unweighted Capital Regulations, *Journal of the American Real Estate and Urban Economics Association*, 22, 59–94.

Haynes, G. W. & Williams, V. (2011). Lending by Depository Lenders to Small Businesses, 2003 to 2010, *US Small Business Administration Office of Advocacy, Banking and Financial Economic Research*, March.

Haynes, G. W. & Brown, J. R. (2009). How Strong Is the Link Between Internal Finance and Small Firm Growth? Evidence from the Survey of Small Business Finances, *US Small Business Administration Office of Advocacy, Banking and Financial Economic Research*, July.

Ivashina, V. & Scharfstein, D. (2010). Bank Lending During the Financial Crisis of 2008, *Journal of Financial Economics*, 97, 319–338.

Keeton, W. R. (1996). Do Bank Mergers Reduce Lending to Businesses and Farmers? New Evidence from Tenth District States, *Federal Reserve Bank of Kansas City, Economic Review,* Q III, 63–75.

Keeton, W. R. (1995). Multi-Office Bank Lending to Small Businesses: Some New Evidence, *Federal Reserve Bank of Kansas City, Economic Review*, Q II, 46–57.

Levonian, M. & Soller, J. (1996). Small Banks, Small Loans, Small Businesses, *Federal Reserve Bank of Saint Francisco, Weekly Letter*, 96–02.

Maltby, E. The Credit Crunch That Won't Go Away: Forget the Improving Economy, Entrepreneurs Still Find It Hard to Get Loans; Here's Why We're in This Mess and How We May Get Out of It, *Wall Street Journal*, 21 June 2010.

Ou, C. & Williams, V. (2009). Lending to Small Businesses by Financial Institutions in the United States, *US Small Business Administration Office of Advocacy, Banking and Financial Economic Research,* July.

Peek, J. & Rosengren, E. S. (1998). Bank Consolidation and Small Business Lending: It's Not Just Size That Matters, *Journal of Banking and* Finance, 22, 799–819.

Peek, J. & Rosengren, E. S. (1995). The Capital Crunch: Neither a Borrower nor a Lender Be, *Journal of Money, Credit, and Banking*, 27, 625–638.

Peek, J. & Rosengren, E. S. (1993). Bank Regulation and the Credit Crunch, *Federal Reserve Bank of Boston Working Paper*, 93–2.

Rosen, R. (2007). Banking Market Conditions and Deposit Interest Rates, *Journal of Banking and Finance*, 31, 3862–3884.

Samolyk, K. A. & Richardson, C. A. (2003). Bank Consolidation and Small Business Lending Within Local Markets, *Federal Deposit Insurance Corporation Working Paper Series*, 03-02.

Shrieves, R. E. & Dahl, D. (1995). Regulation, Recession, and Bank Lending Behavior: The 1990 Credit Crunch, *Journal of Financial Services Research*, 9, 5–30.

Strahan, P. E. & Weston, J. P. (1998). Small Business Lending and the Changing Structure of the Banking Industry, *Journal of Banking and Finance*, 22, 821–845.

Strahan, P. E. & Weston, J. P. (1996). Small Business Lending and Bank Consolidation: Is There Cause for Concern?, *Federal Reserve Bank of New York, Current Issues in Economics and Finance*, 2, 1–6.

Syron, R. F. (1991). Are We Experiencing a Credit Crunch?, *Federal Reserve Bank of Boston New England Economic Review*, July/August, 3–10.

U.S. Small Business Administration. Research and Statistics, Available from <http://www.sba.gov/advocacy/7495>.

U.S. Census Bureau. Statistics of U.S. Businesses, Available from <http://www.census.gov/econ/susb/>.

IMPACT OF COUNTRY OF ORIGIN ON THE PERCEPTION OF HOST COUNTRY NATIONALS: QUALITY AND MANAGERIAL SKILLS OF EXPATS

Margaret A. Goralski, Quinnipiac University
Margaret.Goralski@Quinnipiac.edu

Omar Salgado, EGADE Business School, Tecnológico de Monterrey, Mexico
Salgado.Omar@itesm.mx

ABSTRACT

Volkswagen had been manufacturing cars in Mexico for over fifty years when it decided to replace its German managers in China with managers from Mexico and other Latin American countries. These men and women were well trained and had held managerial positions for quite some time, however, although the work in China could be managed, the mindset of the workers could not. Volkswagen had in place an internationalization strategy for its executives worldwide, which included four principles: designing a formal process for talent scouts worldwide; a pool of internationally trained executives with international career plans; a chain of mentors that would help to develop international managers; and a formal plan of successors. The dilemma of introducing "third-country nationals" (TCNs) as managers in China to replace "parent-country nationals" (PCNs) underscored the fact that "host-country nationals" (HCNs) may not have the same respect for TCNs as they did for PCNs even with Volkswagen's internationalization strategy in place. The aim of this paper is to establish the theoretical foundations to research the impact of country of origin (COO) in the perception of host country executives about the quality and managerial skills of TCNs within three levels of analysis: Country, Company and Individuals.

INTRODUCTION

In 2014, Volkswagen celebrates 50 years of car manufacturing in Mexico and throughout the years the German company has made significant achievements in production and adaptation of cars, initially to Mexico and now to world markets; however, the Mexican subsidiary has achieved other accomplishments that deviate from what is directly related to the manufacture of cars such as the development of management skills in Mexican executives, who are nowadays sent to other markets such as China and Russia to take care of various management activities. The above phenomenon attracts academic curiosity, as it may be possible to construct an analogy with research on the impact of country of origin (COO) in the consumers' perception of product quality; consequently, the expats' COO may have an impact on the perception of host-country nationals (HCNs) on his/her managerial skills, especially where there is a potential ethnocentric attitude within cultures and organizations. This paper is part of a larger study that aims to

determine whether such phenomenon exists in business practice and to what extent or if it is simply part of academic curiosity. Hence, the work begins with the description of the phenomenon along three different levels of analysis, makes a comparison of the COO phenomenon between products and expatriate managers, builds six research propositions, concludes by identifying the main differences between both areas of study, and gives directions for further research.

CONCEPTS WITHIN INTERNATIONAL BUSINESS LITERATURE

It is possible to identify three main areas that encompass the phenomenon described above: (1) International Coordination Mechanisms (ICM), through which multinational corporations (MNCs) establish processes, policies, and people to facilitate the exchange of tangibles and intangibles between organizational units; (2) Subsidiary development process (SDP), regarding how MNCs nurture and empower foreign business units to perform a predefined role within the enterprise network; and finally (3) International talent hunt (ITH), which refers to how MNCs design and develop mechanisms and policies for balancing and clarifying the way individuals have access to international career paths including education and growth opportunities within the corporate network.

With regard to ICM, MNCs penetrate international markets because of three main reasons: market seeking, resource seeking, and efficiency seeking (Narula & Dunning, 2000; Wadhwa & Reddy, 2011), creating an inward/outward exchange of tangibles and intangibles that become a "learning system" (Lall & Narula, 2004) through which MNCs create multi-hub networks between the Headquarters (HQ) located in the MNC "home-country" and Subsidiaries located in a "host-country" (Criscuolo & Narula, 2007). Depending on the internationalization extent, MNCs could have part of their sales, assets, and people located or coming from host-countries, which in the case of the work force, is known as "host-country nationals" (HCNs) or "parent-country nationals" (PCNs) for those coming from the MNCs' original country. The United Nations Conference on Trade and Development (UNCTAD) has developed a transnationality index (TNI), which ranks MNCs depending on the arithmetic mean of three ratios: foreign assets over total assets, foreign sales over total sales, and foreign people over total people.

Certainly, there are MNCs in which ratios account high or higher than they do in their home-country; this is the case of Nestlé, SA in which the foreign people over total people ratio accounts for 0.969 (UNCTAD, 2012), highlighting the importance of HCNs for the company. In consequence, MNCs would follow different ICM strategies depending on the reason to penetrate a particular market, the intensity of the exchange, and the peculiarities of the "host-country" (Rao, 2012), pursuing strategies such as: geographic and product market strategies, entry timing strategies, and organizational design strategies (Singh, 2012) via different approaches such as: formal direct control, socialization, and informal communication (Gusc, Bremmers, & Omta, 2005; Harzing, 2001). In addition to this, the international community produces the Global Competitiveness Report in which nations are ranked according to three main groups of factors:

Basic Requirements, Efficiency Enhancers, Innovations and Sophistication (World Economic Forum, 2013). According to the SDP, the traditional approach towards foreign markets is via corporate mandates through which subsidiaries replicate the corporate knowledge that has been proven successful in the home market (Birkinshaw J. , 1996). No adaptation is pursued unless local conditions endanger the feasibility of the venture. Even though foreign subsidiaries are mainly competence dependent from parent companies (Kuemmerle, 1999), they can persist in order to generate new technologies (Papanastassiou & Pearce, 1997), as has been shown for R&D determinants that subsidiary mandates have the purpose of either "competence-creating" or "competence-exploiting" (Cantwell & Mudambi, 2005; Filippov & Duysters, 2011; Giroud, Ha, Yamin, & Ghauri, 2012). Competence-exploiting mandates imply knowledge flow from HQ to the subsidiary, and it is expected that such knowledge covers the great majority of business situations, as well as commercially packaging knowledge to be easily adopted by the subsidiary. On the other hand, competence-creating would imply that the corporate knowledge does not cover the subsidiary's current business situations and therefore an adaptation or even a complete product and service re-invention is required.

In the case of competence-creating, a substantial change in corporate configuration (geographical coverage, scale and scalability of technology, distribution channels, suppliers profile, market niches, etc.) and coordination (corporate governance, advisory boards, etc.) is needed. Various authors have reported important cases, not only of high subsidiary influence on corporate initiatives, but also on the way the subsidiary feedbacks distinctive knowledge to Headquarters (HQ) and becomes a center of excellence (Ambos, Andersson, & Birkinshaw, 2010; Andersson, Forsgren, & Holm, 2002; Forsgren, Holm, & Johanson, 2005). Subsidiary strategic embeddedness is defined as the capacity of foreign subsidiaries to influence the strategy on MNCs (Garcia-Pont, Canales, & Noboa, 2009). It is currently acknowledged that because of the fact that MNCs are embedded in different business environments, each subsidiary's business network can be a source of competitive advantage (Forsgren, et al., 2005).

Authors divide subsidiary strategic embeddedness into internal and external embeddedness. Internal subsidiary strategic embeddedness is concerned with MNE mechanisms to diffuse firm-strategic advantages (Rugman & Verbeke, 2001), and that constitutes a source of power because of knowledge flows across the corporation (Mudambi & Navarra, 2004). Within internal strategic embeddedness, authors (Garcia-Pont, Canales, & Noboa, 2009) distinguish between "operational embeddedness" that reflects the set of relationships through which foreign subsidiaries carry out day-to-day activities and "strategic embeddedness" that reflects the capacity of the subsidiary to influence MNC's strategy. On the other hand, external embeddedness is defined as the capacity of foreign subsidiaries to collect commercially exploitable knowledge from their location that can be used as a source of power (Andersson, et al., 2002; Birkinshaw, Hood, & Jonsson, 1998; Meyer, Mudambi, & Narula, 2011). Authors argue that although external embeddedness can partially explain their level of knowledge creation (Andersson, Björkman, & Forsgren, 2005), it does impact positively on the subsidiary's expected performance and on the development of product and production processes in the MNC (Andersson, Forsgren, & Holm, 2002). Nevertheless, foreign subsidiaries face corporate and resource constraints to establish lateral relations (Birkinshaw & Morrison, 1995), because of the

probability that the entrepreneurial behavior of subsidiaries' managers may lead to a certain autonomy (Birkinshaw J. , 1996) and reduce subsidiaries' interest in contributing to the overall performance of the MNC (Andersson, Björkman, & Forsgren, 2007). Different modes and taxonomies to explain the way MNCs configure and develop international business have been recently developed allowing subsidiaries to participate in corporate knowledge creation (Noorderhaven & Harzing, 2009), (Salgado, Shi, & Banks, 2012), however, the ability of foreign subsidiaries to participate actively in this development is shaped by corporate and location barriers that foreign subsidiaries face (Salgado O. , 2011).

Regarding ITH, it was mentioned above that MNCs penetrate foreign markets usually with a package of tangibles and intangibles, that have been standardized enough to be exploited commercially; in consequence, most of the host-country nationals hired would be efficiency oriented. Furthermore during new ventures, either through new product introduction or extended geographic coverage, MNCs usually send expatriates to manage the venture and train the locals, who are expected to take over the responsibility within a medium or long term. The statements above project the so called "ethnocentrism staffing policy" (Banai, 1992; Banai & Sama, 2000; Van Merrewijk, 2011), which refers to two different groups of executives within MNCs: (1) the "in-group" which are usually PCNs with the "valid" corporate knowledge to be trusted and (2) the "out-group" which are usually HCNs who lack corporate knowledge and need to be trained. According to this, different managerial skills are needed within corporate networks (Bartlett & Ghoshal, 1992) in coordination; contribution, cross-pollination, and talent hunt duties.

One of the pitfalls of ethnocentrism staffing policies is the fact that expatriates receive more benefits and salary than host-country nationals do, leading to locals' inconformity on the amount and quality of input in comparison to compensation (Toh & DeNisi, 2005). This issue implies the perception and attitude of individuals that are heavily influenced by their cultural values. Hofstede (1980, 1991) identified four components of national culture: power distance, individualism vs. collectivism, masculinity vs. femininity, and uncertainty avoidance with the intention of claiming that work-related values, beliefs, norms, and self-descriptions vary from one culture to another. Definitions of the dimension are as follows (Minkov & Hofstede, 2011): Power distance is a social inequality, which includes a person's relationship with authority; Individualism vs. collectivism is the relationship between an individual and the group; Masculinity vs. femininity is the social implications of gender roles; Uncertainty avoidance is a way of dealing with uncertainty, relating to control of aggression and the expression of emotions. In a later work, Hofstede included a fifth dimension: long vs. short term orientation to address whether the focus of a person's efforts is on the future, present, or past.

The implication of Hofstede's research has brought a significant amount of related studies such as the GLOBE Project (Global Leadership and Organizational Behavior Effectiveness) which introduced other dimensions, as well as, reframe Hofstede's dimensions. The GLOBE project introduced the concept of practices and values to put emphasis on the difference between 'what people experience' and 'what they should desire to experience' (House, Hanges, Javidan, Dorfman, & Gupta, 2004). Therefore, various authors have stated that ITH is a multidimensional

phenomenon in which position, job, organization and acculturation are involved (Jun, Gentry, & Hyun, 2001; Shaffer, Harrison, & Gilley, 1999; Shin, Morgeson, & Campion, 2007). Position and job are two dimensions for which it should be relatively easy to plan and predict the outcome if HCNs or PCNs are assigned, because it depends on whether the executive will be responsible for technical tasks or management teams. However, organization and acculturation are two dimensions for which MNCs can plan and provide training, but success depends entirely on individuals to navigate the politics (organization) and bridging the cultural divide (acculturation).

In case corporate knowledge does not cover all business situations in the host-country, the input of local knowledge becomes imperative and therefore the use of PCNs is questioned even further, especially because of the cost of maintaining a cadre of PCNs. The alternative, to appoint third-country nationals (TCNs) is oftentimes viewed as less expensive than maintaining a cadre of PCNs and can create a cultural bridge between PCNs and HCNs (Tarique, Schuler, & Gong, 2006). Finally, the inclusion of an appropriate composition of HCNs, PCNs and TCNs has an impact on the technological and marketing performance of the subsidiary as well as being able to nurture and expand corporate knowledge (Fang, Jiang, Makino, & Beamish, 2008). The incorporation of the right staffing mix in foreign subsidiaries is claimed to be based on the adoption of polycentric attitudes, the consideration of subsidiary characteristics such as: its role within the corporate network, internationalization stage of the organization, and staff availability, as well as, an appropriate instrumentation of the staffing policy (Reiche, 2007; Shahrokh, 2012). Table 1 shows the summary of concepts reviewed in this section.

TABLE 1. SUMMARY OF MAIN CONCEPTS

Nr.	Area	Concept	Author
1	International Configuration mechanisms	Market seeking, resources seeking, efficiency seeking, learning system	(Lall & Narula, 2004; Narula & Dunning, 2000; Wadhwa & Reddy, 2011)
2		Product market strategies, entry timing strategies, organizational design strategies	(Rao, 2012; Singh, 2012)
3		Formal direct control, socialization and informal communication	(Gusc, et al., 2005; Harzing, 2001)
4	Subsidiary Development Process	Competence-creating, competence-exploiting	(Cantwell & Mudambi, 2005; Filippov & Duysters, 2011; Giroud, et al. 2012)
5		Operational & strategic embeddedness	(Andersson, et al., 2005; Garcia-Pont, et al., 2009)
6		Knowledge diffusion mechanisms & knowledge creation	(Noorderhaven & Harzing, 2009; Rugman & Verbeke, 2001; Salgado, et al., 2012)
7	International Talent Hunt	Ethnocentric & Polycentric attitudes	(Banai, 1992; Banai & Sama, 2000; Van Marrewijk, 2011)
8		position, job, organization and acculturation	(House, et al., 2004; Jun, et al., 2001; Shaffer, et al., 1999; Shin, et al., 2007)
9		Role within the corporate network, internationalization stage, staff availability, instrumentation of the staffing policy	(Reiche, 2007; Shahrokh, 2012)

THE COUNTRY-OF-ORIGIN PHENOMENON

The phenomenon of country-of-origin (COO) effect on a consumers' purchase intention, refers to differential attitudes that consumers may adopt depending on the geographical precedence of products. A consumer may lean his or her purchase decision according the mental map they have about the place that products are related with. Although the above statement is something that might happen during the time the purchase decision is made, the current international business complexity makes it hard to define a fair precedence of a product, because the COO phenomenon can imply many variations such as: country of design (COD), country of manufacture (COM), country of assembly (COA) among others (Ahmed & d'Astous, 1996) and a single product could involve several of the above mentioned options.

It is documented that the COO phenomenon dates back to post WWII period, when defeated countries were forced to put a 'made in' label as a form of punishment; so customers could decide whether or not to buy products from Germany or Japan for instance (Cai, 2002). Therefore, 'made in' labels not only denoted a derogatory remark towards the product and a warning system for the consumer but also created an association between the image of the country and the product. Country image refers to the economic, political, technological and social part of the country (Matin & Eroglu, 1993) but, it also includes other kinds of pejorative characteristics such as low quality and non-socially responsible practices e.g. child labor, slavery, and bribes; pre-setting in consumers' minds that products from least developed countries always underperform in comparison to those from developed countries (Lascu & Giese, 1995). Consequently, governments are investing in building country branding to improve the image of all related activities within the country such as manufacturing and services to which tourism is a remarkable example (Anholt, 2007; Dinnie, 2008).

The complexity of the phenomenon increases when a second association is created between companies and products. Companies build product-market strategies, configure and coordinate networks (Noorderhaven & Harzing, 2009) to penetrate international markets with their own brands; therefore, the concept of Brand of Origin (BO) was created and an evaluation of a product performance can turn into a brand evaluation (O'Cass & Lim, 2002). Furthermore, branding is not only taken as a cue for COO (Thakor, 1996) but also brands can act as de facto ambassadors of the image of a country, in the way that Japan is SONY and IKEA is Sweden (Anholt, 2003, 2007; Dinnie, 2008). Nevertheless, brands carry more features than purely product characteristics and they are classified as follows: 1) symbolic (status, luxury and wealth); 2) function; and 3) usage-based approaches (Thakor, 1996). These extra features are related to people's social identity and motivations (Jin, Chansarkar, & Kondap, 2006), therefore a third association between individuals and products is created. Consequently, even though quality & price are influential during the purchase decision, (Elliott, 1994) it is heavily influenced by the a-priori information consumers may have of the product which are classified as: 1) perceived knowledge; 2) objective knowledge; and 3) experience-based knowledge (Lin & Zhen, 2005). Finally, it is important to note that first consumers' impressions on products and services might be based on their current mental map (Argyris, 1991; Kara, Rojas-Méndez, Kucukemiroglu, & Harcar, 2009) and they are probably not conscious enough to take into consideration the whole

complexity of the COO phenomenon which includes Consumer Related Factors, Company Related factors (Gotsi, Lopez, & Andriopoulos, 2011), and Individual Related Factors (Jin, et al., 2006; Nooderhaven & Harzing, 2003). Table 2 shows a summary of the products and services associations and the moderating factors.

TABLE 2. PRODUCTS AND SERVICES ASSOCIATIONS AND MODERATING FACTORS

Associations	Moderating factors	Authors
Country	Economy Political Technological Social	(Ahmed & d'Astous, 1996; Anholt, 2007; Cai, 2002; Dinnie, 2008; Lascu & Giese, 1995; Matin & Eroglu, 1993)
Company	Brands Strategies Networks	(Gotsi, et al., 2011; O'Cass & Lim, 2002; Noorderhaven & Harzing, 2009; Thakor, 1996)
Individuals	Identity Motivations Perceptions	(Elliott, 1994; Gotsi, et al., 2011; Jin, et al., 2006; Kara, et al., 2009; Lin & Zhen, 2005; Nooderhaven & Harzing, 2003)

INTERPRETATION OF COO PHENOMENON WITHIN EXPATRIATE MANAGERS AREA

As mentioned previously, the aim of this paper is to establish the theoretical foundations of research about the impact of COO in the perception of host-country executives in regard to the quality and managerial skills of TCNs.

Components of COO phenomenon on Expatriate Managers

Based on the international business and country-of-origin literature discussed above, Figure 1 shows the interpretation of COO phenomenon within the area of expatriate managers. The figure contains several three-axis charts, one for each association or level of analysis. The levels of analysis are denoted by Y-axis and they are: Country (Y_a), Company (Y_b) and Individuals (Y_c).

In the Country (Y_a) level, it is possible to find the interrelation of three entities: 1) parent-country where the multinational was originally established; 2) host-country where the multinational has established foreign operations; and 3) a third country which is called in this paper "country-of-origin" as the place where TCN comes from.

In the Company (Y_b) level, the interrelated entities are: 1) Headquarters (HQs), which is the organizational unit that coordinates and regulates the multinational's worldwide activities and delineates strategy; 2) Subsidiary is the organizational unit, which executes the pre-defined strategy and controls the operations in the host-country; finally, 3) the Peer Subsidiary is another foreign subsidiary within the multinational's corporate network and from which the human talent

that is now responsible for the activities in the host-country has been obtained. In the Individual (Y_c) level, the interrelated entities are: 1) Parent Country Nationals (PCNs) who are executives hired by HQs; 2) Host Country Nationals (HCNs) who are executives hired by the Subsidiary; and 3) Third Country Nationals (TCNs) who are executives hired by a Peer Subsidiary. It is important to highlight the use of the word 'hired,' because it would be possible that HQs, Subsidiary, or Peer Subsidiary employ individuals from any other different nationality from the one where their operations are established.

The figure also situates the interrelated entities in its appropriate context within the corporate network, apart from highlighting the entities that have more influence on the development of a strategy for the host-country. In consequence, the X-axis presents two levels of influence: Parent (X_a, X_b, X_c) and Subsidiary ($X_{a'}$, $X_{b'}$, $X_{c'}$), in which the Parent exerts more influence because of the established organizational hierarchy. According to this, the entities with higher influence are: Parent Country (X_a), HQs (X_b), and PCNs (X_c) in comparison to less influential entities such as: Host-Country ($X_{a'}$), Subsidiary ($X_{b'}$), and HCNs ($X_{c'}$). A third axis (Z-axis) is introduced to denote that there is another group of entities that might play a role between Parent – Subsidiary relationship and that exert a certain degree of influence. Entities along Z-axis are: COO (Z_a), Peer Subsidiary (Z_b), and TCNs (Z_c).

Moderating Factors along each different Level of Analysis

Figure 1 will show the translation of the moderating factors of the COO phenomenon from the area of products and services to the area of expatriate managers.

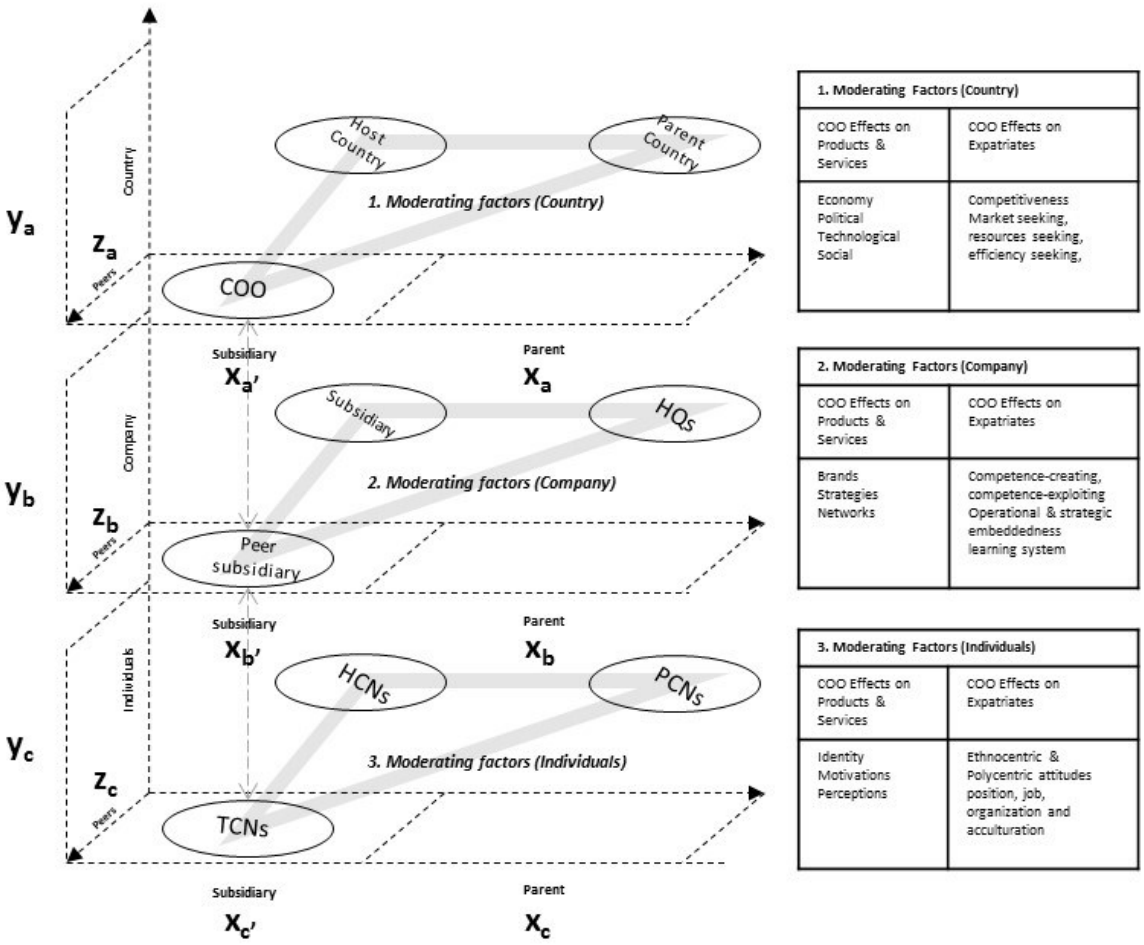


FIGURE 1. INTERPRETATION OF COO PHENOMENON WITHIN THE AREA OF EXPATRIATE MANAGERS

Moderating factors at Country level in the area of products and services refers to 'made in' labels that include the economic, political, technological, and social part of the country (Matin & Eroglu, 1993) even though, 'made in' labels also have certain influence on consumers' perception about product performance (Lascu & Giese, 1995). Consequently, proxies for such moderating factors in the area of expatriate managers are any a-priori information that HCNs have on COO competitiveness and economic strength (market-seeking, resource-seeking, efficiency-seeking) (Lall & Narula, 2004; Narula & Dunning, 2000; Wadhwa & Reddy, 2011). In concrete terms, the perception of HCNs about the quality and managerial skills of TCNs would be influenced by any a-priori information about the competitiveness and economic strength of the country-of-origin.

Moderating factors at Company level in the area of products and services refers to Brands, Strategies, and Networks that multinationals put forward with the aim to penetrate markets (Gotsi, et al., 2011; Noorderhaven & Harzing, 2009; O'Cass & Lim, 2002; Thakor, 1996). Therefore, proxies in the area of expatriate managers are any a-priori information that HCNs

have on competence creating and exploiting initiatives, operational and strategic embeddedness, as well as, the learning systems across the corporate network (Andersson, et al., 2005; Cantwell & Mudambi, 2005; Filippov & Duysters, 2011; Garcia-Pont, et al., 2009; Giroud, et al., 2012). The perception of HCNs about the quality and managerial skills of TCNs would be influenced by any a-priori information about the project type, the operational and strategic influence of the subsidiary, and any knowledge transfer mechanism across the corporate network.

Moderating factors at the individual level in the area of products and services refers to identity, motivations, and perceptions that consumer's project onto the products and services they choose (Elliott, 1994; Gotsi, et al., 2011; Jin, et al., 2006; Kara, et al., 2009; Lin & Zhen, 2005; Nooderhaven & Harzing, 2003). Thus, proxies in the area of expatriate managers are any a-priori information that HCNs have on the position, job, & organization type and the ethnocentric/polycentric attitude within the corporate network (Banai, 1992; Banai & Sama, 2000; House, et al, 2004; Shin, Morgeson, & Campion, 2007; Van Marrewijk, 2011). In concrete terms, the perception of HCNs about the quality and managerial skills of TCNs would be influenced by any a-priori information about on the job duties and attitudes towards staffing.

COUNTRY-OF-ORIGIN IMPACT ON HCNs PERCEPTION

In this section, research propositions are discussed for each level of analysis, defining positive and negative impacts of COO on the perception of HCNs on the quality and managerial skills of expatriate managers. It is important to underline the phrase 'impact on perception,' because research propositions would be defined based on HCNs' opinions and not on proven facts; such opinions might be based on current mental maps and not on informed or conscious considerations of the whole complexity of the COO phenomenon (Argyris, 1991; Kara, et al., 2009).

Research Propositions at Country Level

It was mentioned before that the interrelated entities at country level of analysis are Parent Country, Host Country and COO, for which a country's competitiveness and strength of the economy are moderating factors. Hence, HCNs may build their opinions about both Parent Country and COO based on the a-priori information they have, regardless of whether or not this information is based on data or a stereotype. Similarly to 'made-in' labels that tag products and services (Cai, 2002), as well as, enclosing certain country images and characteristics (Matin & Eroglu, 1993). Therefore, the main research proposition at country level is:

> ***P_1**: a-priori information, in the sense that the competitiveness of the country-of-origin is better than the one of the host-country, would have a positive effect on the perception of the HCNs about the quality and managerial skills of TCNs.*

The above proposition would have several variations depending of the combination of the availability of information and the particular COO competitiveness as shown in Table 3.

In concrete terms, HCNs would have a positive perception only if they have a-priori information in the sense that the competitiveness of the country-of-origin is better than the one of the host-country. Any other combination would produce a negative perception, since HCNs could adopt ethnocentric attitudes arguing that the host country is more likely to produce quality executives necessary for that job.

TABLE 3. VARIATIONS OF THE RESEARCH PROPOSITION AT COUNTRY LEVEL

Proposition	a-priori info available	Better COO competitiveness	HCNs perception on TCNs skills
P_1	Yes	Yes	Positive
P_{1a}	Yes	No	Negative
$P_{1,b}$	No	Yes	Negative
$P_{1,c}$	No	No	Negative

Research Propositions at Company Level

Similarly to Country Level, a HCN may build his or her opinion based on a company's data or a particular stereotype. The interrelated entities are HQs, Subsidiary, and Peer Subsidiary for which competence creating/exploiting initiatives, operational/strategic embeddedness, and learning systems are moderating factors. Company level is more complex than the previous one because the interrelation of entities occurs on a more tangible basis; therefore, it is more likely that HCNs have certain information about what is occurring in the Subsidiary such as: 1) the type of initiative (project) being developed in the Subsidiary; 2) the operational/strategic capability of the Subsidiary; and 3) the learning systems (knowledge sharing mechanisms) implemented across the organization, as a result of corporate communication.

The type of initiative being developed in the Subsidiary

For the case of the initiative that is being developed in the Subsidiary, if such initiative is competence-exploiting then the knowledge flow is mainly from HQ to Subsidiary, while if it is competence-creating, local input is required (Cantwell & Mudambi, 2005; Filippov & Duysters, 2011; Giroud, et al., 2012). Because competence exploiting initiatives are knowledge that has proven successful elsewhere, HCNs would have a duty to adopt such knowledge. Therefore, the following research proposition is given for the type of initiative being developed in the Subsidiary:

***P_2**: a-priori information, in the sense that the project being developed in the Subsidiary is the type of competence-exploiting initiative, would have a positive effect on the perception of the HCNs about the quality and managerial skills of TCNs.*

As stated with proposition P_1, P_2 has variations depending on the combination of the availability of information and whether or not the initiative is competence-exploiting as shown in Table 4. HCNs would have a positive perception only if they have a-priori information in the sense that the project being developed in the Subsidiary is the type of competence-exploiting. Any other combination would produce a negative perception, since HCNs could adopt ethnocentric attitudes arguing that local input is needed and TCNs are not able to provide such information.

TABLE 4. VARIATIONS OF THE RESEARCH PROPOSITION AT COMPANY LEVEL (TYPE OF INITIATIVE)

Proposition	a-priori info available	Project is competence-exploiting	HCNs perception on TCNs skills
P2	Yes	Yes	Positive
P2a	Yes	No	Negative
P2b	No	Yes	Negative
P2c	No	No	Negative

Operational/strategic capability of the subsidiary

For the case of the operational/strategic capability (embeddedness) of the Subsidiary, it was discussed before that 'operational embeddedness' reflects the set of relationships through which foreign subsidiaries carry out day-to-day activities, while 'strategic embeddedness' reflects the capacity of the subsidiary to influence MNC's strategy (Garcia-Pont, Canales, & Noboa, 2009). A higher operational embeddedness would lead to better technical and market performance (Andersson, Björkman, & Forsgren, Managing subsidiary knowledge creation: The effect of control mechanisms on subsidiary local embeddedness, 2005) in comparison with peer subsidiaries within the corporate network; while, a higher strategic embeddedness would reflect the importance given by the multinational to a specific subsidiary for a set of different reasons such as: need for market penetration, need for standardization, etc. Therefore, the following research proposition is given for operational / strategic embeddedness of the Subsidiary:

***P_3**: a-priori information, in the sense that the HCNs' Subsidiary has a higher strategic embeddedness in comparison to TCNs' Peer Subsidiary, would have a positive effect on the perception of the HCNs about the quality and managerial skills of TCNs.*

As shown in previous research propositions, P_3 has variations depending on the combination of the availability of information and whether or not the HCNs' Subsidiary has a higher strategic embeddedness in comparison to the TCNs' Peer Subsidiary as shown in Table 5.

TABLE 5. VARIATIONS OF THE RESEARCH PROPOSITION AT COMPANY LEVEL (SUBSIDIARY CAPABILITY)

Proposition	a-priori info available	Higher strategic embeddedness	HCNs perception on TCNs skills
P3	Yes	Yes	Positive
P3a	Yes	No	Negative
P3b	No	Yes	Negative
P3c	No	No	Negative

Learning systems and knowledge sharing mechanisms

For the case of learning systems and knowledge sharing mechanisms, it was discussed that learning systems enable multinationals to diffuse firm-strategic advantages (Rugman & Verbeke, 2001) and it was also mentioned that the access and the exchange of knowledge constitutes a source of power (Mudambi & Navarra, 2004), throughout different modes and directions (Nooderhaven & Harzing, 2009; Salgado, et al., 2012). In consequence, the existence and frequency of use of knowledge sharing systems within the corporate network is an important moderating factor to which the following research proposition is given:

> ***P_4**: a-priori information, in the sense that learning systems and knowledge sharing mechanisms are established and frequently used, would have a positive effect on the perception of the HCNs about the quality and managerial skills of TCNs.*

Also, P_4 has variations depending on the combination of the availability of information and whether or not learning systems and knowledge sharing mechanisms are established and frequently used as shown in Table 6.

TABLE 6. VARIATIONS OF THE RESEARCH PROPOSITION AT COMPANY LEVEL (LEARNING SYSTEMS & KNOWLEDGE SHARING)

Proposition	a-priori info available	Learning systems & knowledge Sharing	HCNs perception on TCNs skills
P4	Yes	Yes	Positive
P4a	Yes	No	Negative
P4b	No	Yes	Negative
P4c	No	No	Negative

Research propositions at the Individual level

In the case of the Individual level of analysis, the interrelated entities are PCNs, HCNs and TCNs for which attitudes, jobs, position, organization, and acculturation are moderating factors. A HCN may build his or her opinion based on particular data or stereotypes about the type of needs for both: the position and the workplace.

HCNs' identity and motivations for the position

A HCNs' opinions denote his/her identity and motivations (Elliott, 1994; Gotsi, et al., 2011; Jin, et al., 2006; Kara, et al., 2009; Lin & Zhen, 2005) and based on them, they may have already not only a number of tasks and duties that TCNs may develop in order to successfully lead in the position, but also a set of priorities that need to be established. Once on duty, a TCN may or may not communicate his/her motivations, priorities, and work plan and also the alignment between HCNs and TCNs identities and motivations for the position would impact on the HCNs opinions. Consequently, the following research proposition is given:

> ***P_5****: a-priori information, in the sense that TCNs' identity and motivations for the position are similar to the HCNs, would have a positive effect on the perception of the HCNs about the quality and managerial skills of the TCNs.*

Similarly to other research propositions, P_5 has variations depending on the combination of the availability of information and whether or not TCNs identity, motivation, and work plan are similar to the one of HCNs as shown in Table 7.

TABLE 7. VARIATIONS OF THE RESEARCH PROPOSITION AT INDIVIDUALS LEVEL (IDENTITY & MOTIVATIONS FOR THE POSITION)

Proposition	a-priori info available	Similar identity and motivations	HCNs perception on TCNs skills
P5	Yes	Yes	Positive
P5a	Yes	No	Negative
P5b	No	Yes	Negative
P5c	No	No	Negative

Ethnocentric attitudes in the workplace

Attitudes taken in the workplace by PCNs, HCNs, and/or TCNs would certainly have an impact on the opinion of all (Banai, 1992; Banai & Sama, 2000; House, et al., 2004; Shin, et al., 2007; Van Marrewijk, 2011). The attitude individuals take is related to the openness not only to listen to the other party, but also to adopt new practices. Therefore the following research proposition is given:

> ***P_6**: a-priori information, in the sense that TCNs' are open not only to listen but also to adopt new practices, would have a positive effect on the perception of the HCNs about the quality and managerial skills of TCNs.*

Again, P_6 has variations depending on the combination of the availability of information and whether or not TCNs are flexible to listen and adopt new practices as shown in Table 8.

TABLE 8. VARIATIONS OF THE RESEARCH PROPOSITION AT INDIVIDUALS LEVEL (ETHNOCENTRIC ATTITUDES IN THE WORKPLACE)

Proposition	a-priori info available	Listen and flexible to adopt	HCNs perception on TCNs skills
P6	Yes	Yes	Positive
P6a	Yes	No	Negative
P6b	No	Yes	Negative
P6c	No	No	Negative

DISCUSSION

It is important to contextualize each of the research propositions because in that way, readers can better understand the relationships described in the previous section for country, company, and individuals as levels of analysis.

For the first research proposition (P_1), it is possible that individuals might have in mind certain characteristics that circumscribe a country, which has been named in this article as country image. Some characteristics are easy to link with country names such as educational level, if developed and undeveloped countries are compared; however, educational level is not easy to scale fairly if two developed countries are compared. Another misplacement of a countries' image is when individuals tend to put emphasis on particular good or bad characteristics that have been dragged from a whole image or stereotype of a country. This is the case of some of the developed countries that have moved away from manufacturing activities leading their economies mainly to service activities; hence, an industrial comparison between manufacturing activities in developed and undeveloped countries may lead to wrong assumptions. An example of this misunderstanding would be a comparison of the industrial activities between South Korea

and the UK. Finally, scaled down country images to predict individual outcomes is an invalid argument or fallacy.

The second (P_2), third (P_3) and fourth (P_4) research propositions, fall within the ability of the multinational to convey the mission and vision of the company. This is not just about a communication exercise across the organization, but also to achieve that strategic objectives are translated into viable and transparent operational actions for the entire organization. However, the translation of strategic objectives into operational actions is difficult since on one hand, the company would like to promote an internal competition and on the other, maintain a balance between its business units. In consequence, assignation of competence exploiting and creating initiatives, as well as, allowing certain autonomy in subsidiaries and promoting embeddedness in the location and the corporate network, is more than merely an investment topic. Therefore, enabling learning systems and knowledge sharing in all directions (HQs - Subsidiaries; Subsidiaries - HQs; and Subsidiaries - Peers) is one of the ways to balance corporate networks and encourage competition.

Fifth (P_5) and sixth (P_6) research propositions have to do with the atmosphere created within the subsidiary, i.e., if the multinational has promoted the exchange of knowledge not only from HQs to subsidiaries but also among peers, the workplace would be less ethnocentric and more oriented to listen and adopt different ideas. Also, it has to do with enabling mechanisms to educate and promote local talent for achieving positions elsewhere.

CONCLUSION

It is possible to conclude that the CCO phenomenon can be translated into the area of expatriate managers with some limitations. The first limitation is consumers' freedom of choice when deciding to buy a product or service with an origin that is a different country from the one in which it is being consumed; in contrast, HCNs do not have this choice, under the consideration that TCNs are sent by HQs to replace PCNs. Another important difference is that the service provided by TCNs can be edited and aligned to local needs and claims along the expatriate experience, while products and services are normally attached to predefined contracts and legally restricted. The third limitation is the difficulty of replacing the product or service if it does not meet the specifications for which it was intended; in contrast, the description in detail arguing why TCNs should be replaced involves qualitative aspects, because it is necessary to consider the learning curve and adaptation to the position. Finally, considering that the TCNs come from the exchange between business units within the same company, there is a greater chance to predict their performance throughout training and reward systems, in contrast to products traded outside the company which imply unilateral mechanisms to verify such performance.

SUGGESTIONS FOR FURTHER RESEARCH

As explained in the introduction, the purpose of this article was to establish the theoretical foundations to research the impact of COO in the perception of HCNs about the quality and managerial skills of TCNs; therefore, the practical implication of this study would be to increase the likelihood of success of an expatriate assignment abroad. This can be achieved by designing a system with quantitative dimensions based on each research proposition that together, are useful for measuring the *'Corporate Infrastructural Distance'* (CID) within three levels of analysis: country, company and individuals.

The above mentioned system would be composed of six dimensions: one for the country level, three for the company level and two for the individual level. The weighted average of these dimensions would give us an *'Index of Corporate Infrastructure'* (ICI) which allows us not only to compare business units within the same corporate network but also, extrapolate results to markets where the company still has no presence.

As an example, Figure 2 compares cases at country level of analysis as an analogy for the VW corporate network. The company has had a presence in Germany, Mexico, and China for long time but no consistent and uninterrupted presence in the U.S. The top charts show the relation between competitiveness, exploration, and exploitation within those three countries. For the case of competitiveness (1), it is possible to see that China and Mexico are far away from Germany in such dimension but, the U.S. is closer. A similar situation occurs for exploration dimension (2), because both Germany and the U.S. have the proper business environment. However the situation changes for the exploitation of initiatives (3), for which the European automotive market seems to be quite different from the American one in terms of consumers taste, usages, styles, etc.

Finally, it is important to recall that an effective operationalization of the dimensions proposed in the system will award more sense to the results obtained in comparisons, for example, it would be important not to let the dimension of "competitiveness" in a broad definition, overshadow a narrower definition i.e. the competitiveness in the automotive industry as one example.

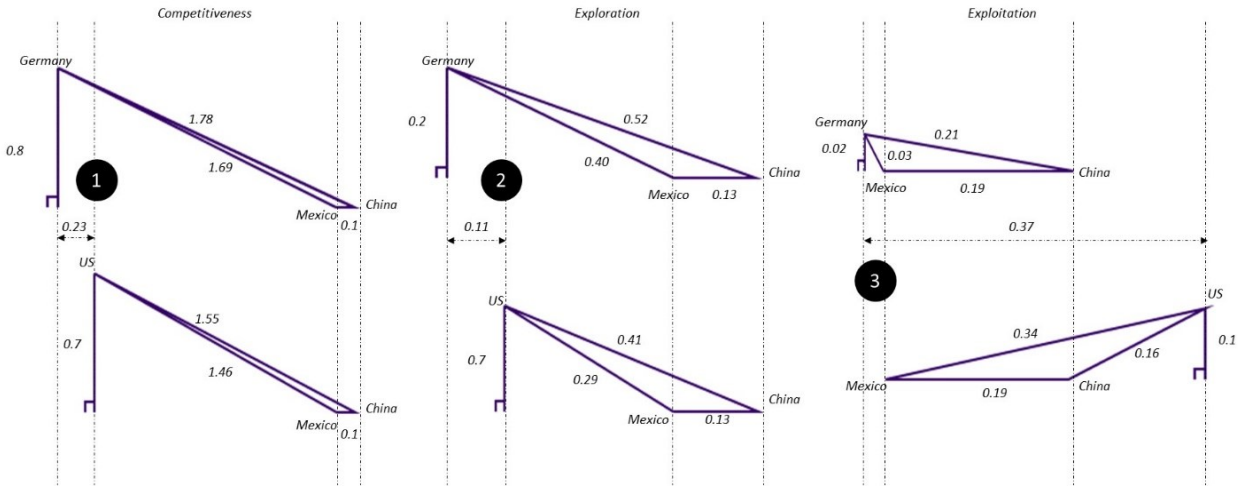


FIGURE 2. EXTRAPOLATION SYSTEM

REFERENCES

Ahmed, S., & d'Astous, A. (1996). Country of Orign and Brand Effects: A Multi Dimensional and Multi Atrribute Study. *Journal of International Consumer Marketing, 9*(2), 93-115.

Ambos, T., Andersson, U., & Birkinshaw, J. (2010). What are the consequences of initiative-taking in multinational subsidiaries? *Journal of International Business Studies, 41*(7), 1099-1118.

Andersson, U., Björkman, I., & Forsgren, M. (2005). Managing subsidiary knowledge creation: The effect of control mechanisms on subsidiary local embeddedness. *International Business Review, 14*(5), 521-538.

Andersson, U., Björkman, I., & Forsgren, M. (2007). Balancing subsidiary influence in the federative MNC: a business network view. *Journal of International Business Studies, 38*(5), 802-818.

Andersson, U., Forsgren, M., & Holm, U. (2002). The strategic impact of external networks: Subsidiary performance and competence development in the multinational corporation. *Strategic Management Journal, 23*(11), 979-996.

Anholt, S. (2003). *Brand new justice: The upside of global branding.* Oxford: Butterworth-Heinemann.

Anholt, S. (2007). *Competitive identity: The new brand management for nations, cities and regions.* Basingstoke: Palgrave Macmillan.

Argyris, C. (1991). Teaching Smart People How to Learn. *Harvard Business Review, 63*(9), 99–109.

Banai, M. (1992). The ethnocentric staffing policy in multinational corporations: a self-fulfilling prophecy. *The Internationai Journal of Human Resource Management, 3*(3), 451-472.

Banai, M., & Sama, L. (2000). Ethical dilemmas in MNCs' international staffing policies a conceptual framework. *Journal of Business Ethics, 25*(3), 221-235.

Bartlett, C., & Ghoshal, S. (1992). What is a global manager? *Harvard Business Review, 70*(5), 124–132.

Birkinshaw, J. (1996). How Multinational Subsidiary Mandates Are Gained And Lost. *Journal of International Business Studies, 27*(3), 467-495.

Birkinshaw, J., & Morrison, A. (1995). Configurations Of Strategy And Structure In Subsidiaries Of Multinational Corporations. *Journal of International Business Studies, 26*(4), 729-753.

Birkinshaw, J., Hood, N., & Jonsson, S. (1998). Building firm-specific advantages in multinational corporations: The role of subsidiary initiative. *Strategic Management Journal, 19*(3), 221-241.

Cai, Y. (2002). *University of Georgia, MSc Thesis.* Retrieved 2014, from http://www.fcs.uga.edu/ss/docs/cai_yi_200208_ms.pdf

Cantwell, J., & Mudambi, R. (2005). Mne Competence-Creating Subsidiary Mandates. *Strategic Management Journal, 26*(12), 1109-1128.

Criscuolo, P., & Narula, R. (2007). Using Multi-hub Structures for International R&D: Organisational Inertia and the Challenges of Implementation. *Management International Review (MIR), 47*(5), 639-660.

Dinnie, K. (2008). *Nation branding: Concepts, issues, practice.* Oxford: Butterworth-Heinemann.

Elliott, G. (1994). Consumer Perception of Product Quality and the Country-of Origin Effect. *Journal of International Marketing, 2*(2), 49-62.

Fang, Y., Jiang, G.-L., Makino, S., & Beamish, P. (2008). Multinational Firm Knowledge, Use Of Expatriates And Foreign Subsidiary Performance. *Academy of Management Proceedings*, 1-6.

Filippov, S., & Duysters, G. (2011). Competence-building in foreign subsidiaries: The case of new EU member states. *Journal for East European Management Studies, 16*(4), 286-314.

Forsgren, M., Holm, U., & Johanson, J. (2005). *Managing the Embedded Multinational* (Vol. 1). Northampton, MA: Edward Elgar.

Garcia-Pont, C., Canales, J., & Noboa, F. (2009). Subsidiary Strategy: The Embeddedness Component. *Journal of Management Studies, 46*(2), 182-214.

Giroud, A. p., Ha, Y. J., Yamin, M. p., & Ghauri, P. (2012). Innovation policy, competence creation and innovation performance of foreign subsidiaries: The case of South Korea. *Asian Business & Management, 11*(1), 56-78.

Gotsi, M., Lopez, C., & Andriopoulos, C. (2011). Building country image through corporate image: exploring the factors that influence the image transfer. *Journal of Strategic Marketing, 19*(3), 255-272.

Gusc, J., Bremmers, H., & Omta, S. (2005). Management Control Systems in Subsidiaries of Multinationals in the Emerging Market of Central Eastern Europe. *Management Futures, 8*(2), 57-78.

Harzing, A.-W. (2001). Of Bears, Bumble-Bees, and Spiders: The Role of Expatriates in Controlling Foreign Subsidiaries. *Journal of World Business, 36*(4), 366-379.

Hofstede, G. (1980). *Culture's Consequences: International Differences in Work-related Values.* Beverly Hills, CA: Sage.

Hofstede, G. (1991). *Cultures and Organizations - Software of the Mind.* New York, NY: McGraw-Hill.

House, R., Hanges, P., Javidan, M., Dorfman, P., & Gupta, V. (. (2004). *Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies.* Thousand Oaks, CA: Sage.

Jin, Z., Chansarkar, B., & Kondap, N. (2006). Brand Origin in an Emerging Market: Perceptions of Indian Consumers. *Asia Pacific Journal of Marketing and Logistics, 18*(4), 283-302.

Jun, S., Gentry, J., & Hyun, Y. J. (2001). Cultural Adaptation of Business Expatriates in the Host Marketplace. *Journal of International Business Studies, 32*(2), 369-377.

Kara, A., Rojas-Méndez, J. I., Kucukemiroglu, O., & Harcar, T. (2009). Consumer preferences of store brands: Role of prior experiences and value consciousness. *ournal of Targeting, Measurement & Analysis for Marketing, 17*(2), 127–137.

Kuemmerle, W. (1999). The drivers of foreign direct investment into research and development: An empirical investigation. *Journal of International Business Studies, 30*(1), 1-24.

Lall, S., & Narula, R. (2004). Direct Investment and its Role in Economic Development: Do We Need a New Agenda? *European Journal of Development Research, 16*(3), 447- 464.

Lascu, D., & Giese, T. (1995). Exploring country bias in a retailing environment: implications of retailer country of origin. *Journal of Global Marketing, 9*(1/2), 41-58.

Lin, L., & Zhen, J. (2005). Extrinsic product performance signaling, product knowledge and consumer satisfaction: an integrated analysis an example of notebook consumer behavior in Taipei city. *Fu Jen Management Review*, 65-91.

Matin, I., & Eroglu, S. (1993). Measuring a multi-dimensional construct: Country image. *Journal of Business Research, 28*(3), 191-210.

Meyer, K., Mudambi, R., & Narula, R. (2011). MNEs and Local contexts: location, control and value creation. *Journal of Management Studies, 48*(2), 253-277.

Minkov, M., & Hofstede, G. (2011). The evolution of hofstedes doctrine. *Cross Cultural Management, 18*(1), 10-20.

Mudambi, R., & Navarra, P. (2004). Is knowledge power? Knowledge flows, subsidiary power and rent-seeking within MNCs. *Journal of International Business Studies, 35*(5), 385-406.

Narula, R., & Dunning, J. (2000). Industrial Development, Globalization and Multinational Enterprises: New Realities for Developing Countries. *Oxford Development Studies, 28*(2), 141-167.

Nooderhaven, N., & Harzing, A. (2003). The "Country-of Origin effect" in Multinational Corporations: Sources, Mechanisms and Moderating Conditions. *Management International Review, 43*(2), 47-66.

Noorderhaven, N., & Harzing, A.-W. (2009). Knowledge-sharing and social interaction within MNEs. *Journal of International Business Studies, 40*(5), 719–741.

O'Cass, A., & Lim, K. (2002). Understanding the Younger Singaporean Consumer's View of Western and Eastern Brands. *Asia Pacific Journal of Marketing and Logistics*, 54-79.

Papanastassiou, M., & Pearce, R. (1997). Technology sourcing and the strategic roles of manufacturing subsidiaries in the U.K.: Local competences and global competitiveness. *Management International Review, 37*(1), 5-25.

Rao, M. T. (2012). Control and Coordination of Information Systems in Multinational Corporations: An Empirical Examination of Subsidiary and Host Country Factors. *Journal of Global Information Technology Management, 15*(4), 64-84.

Reiche, B. (2007). The effect of international staffing practices on subsidiary staff retention in multinational corporations. *International Journal of Human Resource Management, 18*(4), 523-536.

Rugman, A. M., & Verbeke, A. (2001). Subsidiary-specific advantages in multinational enterprises. *Strategic Management Journal, 22*(3), 237-250.

Salgado, O. (2011). Barriers to the accomplishment of a subsidiary's strategic role: How Location and Corporate Networks Influence Subsidiary Performance. *Revista de Administração Contemporânea, ANPAD, 15*(2), 261-282.

Salgado, O., Shi, Y., & Banks, J. (2012). Emerging capabilities in manufacturing companies: taxonomy of multinationals' transformation. *Journal of Business and Globalization, 9*(3), 275–310.

Shaffer, M., Harrison, D., & Gilley, K. (1999). Dimensions, Determinants, and Differences in the Expatriate Adjustment Process. *Journal of International Business Studies, 30*(3), 557-581.

Shahrokh, J. (2012). *A study of appropriate staffing mix in the subsidiaries of multinational companies*. Ann Arbor, MI: UMI Dissertations Publishing.

Shin, S., Morgeson, F., & Campion, M. (2007). What you do depends on where you are: understanding how domestic and expatriate work requirements depend upon the cultural context. *Journal of International Business Studies, 38*(1), 64-83.

Singh, D. (2012). Emerging economies and multinational corporations. *International Journal of Emerging Markets, 7*(4), 397-410.

Tarique, I., Schuler, R., & Gong, Y. (2006). A model of multinational enterprise subsidiary staffing composition. *International Journal of Human Resource Management, 17*(2), 207-224.

Thakor, M. (1996). Brand origin: Conceptualization and review. *The Journal of Consumer Marketing, 13*(3), 27-42.

Toh, S. M., & DeNisi, A. (2005). A local perspective to expatriate success. *Academy of Management Executive, 19*(1), 132-146.

UNCTAD. (2012). *United Nations Conference on Trade & Development.* Retrieved 2014, from http://unctad.org/

Van Marrewijk, A. (2011). Cross Cultural Management: Hybridization of Dutch – Indian Work Practices in Geographically Distributed IT Projects. *International Journal of Business Anthropology, 2*(2), 15-35.

Wadhwa, K., & Reddy, S. S. (2011). Foreign Direct Investment into Developing Asian Countries: The Role of Market Seeking, Resource Seeking and Efficiency Seeking Factors. *International Journal of Business and Management, 6*(11), 219-226.

World Economic Forum. (2013). *Global Competitiveness Report.* Retrieved 2013, from http://www.weforum.org/reports/global-competitiveness-report-2013-2014

AN EMPIRICAL ASSESSMENT OF PERFORMANCE AND SCALABILITY OF DECENTRALIZED DISK STORAGE FOR REAL-TIME BUSINESS APPLICATIONS

Dennis C. Guster, Saint Cloud State University
dcguster@stcloudstate.edu

Mark B. Schmidt, Saint Cloud State University
mbschmidt@stcloudstate.edu

Erich P. Rice, Saint Cloud State University
rier1201@stcloudstate.edu

ABSTRACT

The growth of the Internet has increased the need for effective databases. One is Cassandra, an open source database by the Apache Software Foundation, which claims high scalability and is able to run on commodity hardware. Because mechanical hard drives are the bottleneck in the data retrieval process, it is reasonable to investigate optimization by storing data on multiple disks, distributed across multiple devices. This methodology suggests a reduction of data access time by using Cassandra. Research is needed to determine advantages that can be obtained by using distributed databases. This study obtained data from a basic configuration of Cassandra, and the test bed revealed that a distributed database using additional nodes could reduce latency and add efficiency. However, as more nodes were added to Cassandra, diminishing returns were observed and the addition of nodes added only slightly to the efficiency of the database.

INTRODUCTION

The Internet has revolutionized the basic business model in which customers need to be physically near the product they wish to purchase, or be on a mailing list for a catalog that then allows them access to it through the mail. Its worldwide connectivity has for many businesses expanded their customer base from maybe 10,000 people residing in their local region to millions of people around the world who have web access. While this new global commerce model is most gratifying from a profit perspective to the individual companies, it necessitates the support of huge pools of e-commerce users, who need to extract data on both a read and write level from a backend database structure.

To put the ramifications of this massive growth in perspective, a quick review of the hardware architecture needed to support it is in order. Before the rapid worldwide growth of the Internet,

the maximum number of potential users of a system was typically limited to the size of the connected private network that supported it; this typically was limited to perhaps a few thousand users. However, the new global Internet based e-commerce web model has resulted in user groups in the millions. Supporting a user pool of this magnitude is not trivial and opens a question in regard to the functionality of the system related to reliability, performance, security and scalability. Further, the training level of personnel required to run such a system is much greater than in just a regional or private network model. The primary metric to be evaluated in this paper is performance or more specifically the number of read or insert operations performed in a given amount of time by the database system. To help put those metrics in perspective a brief discussion concerning response time is warranted. Because the primary delivery mechanism for business on the Internet is e-commerce, which is usually web based, a target latency of about three seconds or less is often the goal (Brown, Guster, & Krzenski, 2007). Achieving this metric can be challenging and can become more elusive as the size of the database increases, as well as when the volume and complexity of inquiries increase. Therefore, the design of the architecture used to store the data and the staging (queuing) process when extracting that data becomes an integral component in meeting the three second delay target. For a more detailed discussion see Zimmermann, Wei-Shinn, and Wei-Cheng (2004) that addresses this question and delineates how the queuing structure could be optimized in a web-based GIS application.

Given the available hardware then what is the most efficient and cost-effective way to optimize the stored data, which will eventually be extracted, processed and then forwarded to a web client? While faster alternatives exist (like solid state drives) the traditional mechanical hard drive technology is still the primary architecture for large scale data storage. Because these drives are mechanical in nature they are typically the slowest part of the information retrieval process. In the traditional sense, involving a few large mechanical drives, the target performance goal of "three seconds or less" cannot be achieved in intensive web applications that involve an end user pool which generates a "millions-of-hits" restriction point, the work by Otey, Parthasarathy, Wang, Veloso, and Meira (2004) looked at ways around this bottleneck.

Any business that has been able to effectively use e-commerce and thereby expand their customer base worldwide, must address this "millions-of-hits" mechanical bottleneck. It therefore makes sense to investigate methods that optimize large-scale storage. Further, because of its widespread usage and thus cost effectiveness, mechanical storage technology is still the dominant architecture and it seems logical to investigate methods that improve upon its characteristics. There are numerous studies that indicate using multiple nodes and creating a distributed database can be an effective solution (Brown et al., 2007; Elnikety, Tracey, Nahum, & Zwaenepoel 2004; Kanitkar, 2000; Kanitkar & Delis, 2002). Historically, while the basic methodology offered a solution, it often involved extensive setup and required a considerable amount of time from high-level personnel, which limited its cost-effectiveness. The widespread use of the LINUX operating system, which in addition to enhancing performance, security, and scalability created an excellent platform to deploy shareware, has created greater efficiency possibilities. There have been several shareware products that can easily reside on top of existing LINUX nodes (often in a cloud architecture), which make the harnessing of distributed database logic more practical, less personnel intensive, and hence more cost effective.

Perhaps the most popular of these shareware database products is Cassandra. Lakshman and Malik (2010) described Cassandra as a distributed storage system capable of managing huge amounts of structured data distributed across many commodity servers, and providing a high availability of service and fault tolerance. Cassandra was designed to operate in a way so it could be used on hundreds of nodes and could be distributed across multiple data centers, or Clouds (Cockcraft, 2011). When operating on this scale, it is not unusual for hardware components to fail almost continuously. Cassandra is designed to manage these drive failures ensuring the reliability and scalability of the software systems that rely on this service, in the case of Netflix, it chose a Cloud driven solution through the Amazon Web Service (AWS) (Cockraft, 2011). Although Cassandra resembles a relational database in that it incorporates database design and implementation strategies, Cassandra is not a relational database. It instead provides a simple data model that supports dynamic control over data layout and format, while providing increased redundancy and greater fault tolerance (Cockcraft, 2011).

Therefore, the purpose of this paper is to investigate the appropriateness and effectiveness of using the Cassandra model to help alleviate the "millions-of-hits" scenario, which might be encountered by a business using e-commerce on a global scale. Specifically, the system will be evaluated in regard to performance gains and the ease with which it can be implemented. To collect meaningful data, a series of experimental trials were carried out on a test bed of existing nodes within an autonomous system set up as a Cloud infrastructure, the tests were carried out under differing loads to help ascertain the performance gains that might be realized by using the Cassandra distributed database system.

BACKGROUND AND REVIEW OF CURRENT PRACTICE

The idea behind a distributed database system is certainly not a new or novel one. The idea and implementation of a distributed database system can be traced back at least to the late 1970s and the work of Bernstein, Rothnie, Goodman, & Papadimitriou (1978). As the amount of data produced and the need to extract and manipulate it with ever increasing speed and accuracy has increased, new and more efficient ways have been developed to deal with this ever-increasing cascade of data. As Abadi (2012) pointed out, though the initial wave of potential end users were limited to large organizations such as multinational corporations and the Federal government, the newer distributed database technologies have allowed a greater number and variety of organizations to utilize them. The benefits that even small or medium sized firms can now realize from a distributed database environment can far outweigh the costs of creating and managing the system. As Oszu and Valduriez (1991) pointed out, the benefits that such a system could provide are better system performance, ease of managing replicated data, improved reliability of transactions, and hopefully an easier and more economical method of scalability.

In today's data driven world, nothing is more important to an organization than the accumulation and preservation of useful business relevant data. The storing and easy access of that pertinent data is often times of critical importance to the continued operation of a going concern. Among

the important developments in accessing and utilizing applications has been the rise of virtualization, through tools such as VMware or Microsoft's Hyper-V, and it has been noted that virtualization has created greater efficiencies of operation (Hemminger, Rogers, & Guster 2010; Safonov, Guster, & Hemminger 2011). Thus, virtualization also has a place within the distributed database realm as well. In Oracle for example, the ability to utilize a virtual private database or VPD, allows the database administrator a level of finer granularity control over both row and column level access (Afyouni, 2006). Within the distributed database realm, various ideas have been developed to create a framework around which the virtualized database space can be managed, unlike Oracle in which an individual instance or "view" of a portion of the database is given to a user. For instance, Xu, Jing, Yongwei, Xiaomeng, & Guangwen (2008) presented a methodology for a virtual database management system or VDM, which would allow the organization to integrate different data resource types and varying data sets into a cohesive unit.

The VDM architecture could also allow users of the virtual database environment to utilize a standardized method such as structured query language or SQL to manage and query the database structure. Xu et al. (2008) also pointed out that the VDM environment might be best suited to scientific areas of research or data collection, such as data integration across organizations or spatial information grids which require a high degree of cross referencing. Although the data type can cause varying differences in how it is stored and indexed within the database, due to the Internet, the typical delivery method for the indexed data is through a web interface of some kind. Thus, the earlier discussed "three second rule" comes into play and queuing and storage of the data becomes an important piece to how the system operates, and whether it meets end user expectations. Zimmerman et al. (2004) discussed how a web based Geographic Information System (GIS) could be structured that best optimizes the end user experience, by queuing the data within the distributed database structure.

Although distributed database systems have been shown to provide a great many benefits to web or cloud based architectures (Cockcraft, 2011), the best way to facilitate that architecture is still under consideration. A peer-to-peer architecture works well on a cluster level, and although there are other options, Bonifati, Panos, Aris, &Kai-Uwe (2008) provide an excellent analysis of how to best integrate a peer-to-peer architecture with a database centric model. Only through this type of effort can the true benefits of a distributed database system be realized. Other research has focused on the CAP or consistency, availability, and partition tolerance concept. Abadi (2012) points out that within the CAP methodology there are tradeoffs, and that in designing a distributed database system really only two of the three desired criteria can be reconciled.

Abadi (2012) also explains that it is important to look beyond the CAP model, which he views as too simplistic, and instead define it as PACELC. The acronym stands for (P)artition, how the system trades off between (A)vailability and (C)onsistency; (E)lse, when the system is running under normal conditions without partitions, how does the system tradeoff between (L)atency and (C)onsistency? The ELC portion of the methodology only applies to those systems that replicate

data, such as a replication (REPL) database. Nevertheless, these types of balancing decisions are needed when determining the best option for an organization's distributed database structure.

This brings us to Cassandra, which the Apache Software Foundation developed as an open source option for high-level distributed database applications (Apache, 2009). The Cassandra distributed database platform claims to provide a high level of system performance, coupled with the ability to scale up the implementation by adding additional nodes to the distributed database network. It can be implemented through either a hardware cluster, or through the means of a cloud infrastructure as was tested in Cockcraft (2011). Cassandra's ability to replicate itself across datacenters and across the cloud has lead many large companies to utilize it for their distributed database needs, as was shown in the Netflix case example (Cockcraft, 2011). The amount of data being created and stored by large Internet firms such as Netflix has grown into the petabyte range, or 10 to the 15^{th} power bytes of information, and ways to scale up the storage apparatus of Cassandra through the continued addition of nodes is a main selling point to its use. Cockcraft (2011) points out that, at least in the Netflix example, the increase in performance was an almost linear increase in performance as new nodes were brought online.

Cassandra can be configured also to promote a higher level of system availability, though it compromises data consistency, it will however allow the user to select the degree of this tradeoff. Thus, data stored within the Cassandra structure can be configured to replicate across N different peers within its host cloud or cluster, while employing a gossip protocol to ensure each node will maintain its state in relation to its peers (Featherston, 2010). For these reasons many have chosen to utilize the Cassandra distributed database system for their own implementations, and those of their clients (DataStax, 2013).

METHODOLOGY

The test bed for the experimental portion of the paper was configured within an existing cloud infrastructure located within the authors' lab. From a hardware perspective each node was virtualized on a Dual Core CPU at 2.2GHz (AMD Opteron 6174) with 32GB of RAM and 2.6TB of hard drive space using Cronos VMware NFS. Each node was replicated only once, which would be considered the minimum if fault tolerance was desired as it would be in a real world enterprise level setting. The basic processing stack was set to 50 threads which is a reasonable starting point to assess the effectiveness of multiprocessing. The block of data to be inserted (written) or read consisted of five columns of data in a table and each column was 34 bytes in length. This was designed to mimic a basic inquiry/response transaction processing system.

The metrics chosen to determine the efficiency of using the distributed database were primarily based on latency. The basic metric was: median latency, which is a measure of how long any given transaction might have to wait to be processed and the amount of elapsed time, which is a measure of how long it will take to complete either the read or insert session. Latencies at the

95th and 99.9th percentile were also reported to provide an idea of how the system might function under duress, akin to the "millions-of-hits" scenario mentioned earlier. Also, the session time (time it took to run a specific trial) was reported as well. Sessions of varying intensities were used to determine how Cassandra might scale as the workload and number of nodes used was increased. The sessions varied from a minimum of one million records to a maximum of 15 million records. A single client against a Cassandra database containing eight nodes generated the initial tests detailed in the section below.

RESULTS

The results are reported below in Table 1. These results indicate that the eight-node Cassandra database did a pretty good job of scaling from one million reads/inserts to ten million reads/inserts. While data was obtained for inserts at the fifteen million-level, the same experiment conducted at that level for reads generated so many errors that the data was not available. In terms of session time (reported as m:ss) as the workload increased, as one would expect, so did the session time. In fact, the five million times are about five times greater than the one million times results, and that linear pattern generally holds true at the ten and fifteen million levels.

Interestingly, the median latency remained in the 1.4 to 1.8 second range across all intensity levels. Further, the results obtained up to the 95th percentile were encouraging which ranged from 2.6 to 3.2 seconds. For example, if one's target latency was three seconds or less (as mentioned earlier as a threshold of end user expectations) then the current configuration would be able to handle that up to ten million records with a 95% confidence factor. However, there are numerous outliers at the 99.9th percentile at all intensity levels. All values exceeded 20 seconds and the decay was really noticeable at the fifteen million level where a value greater than 800 seconds was observed. For many applications (such as e-commerce) having even one outlier of the magnitudes observed at the 99.9th percentile would be unacceptable. In e-commerce customers expect a maximum wait time of three to five seconds, if that threshold is exceeded they may simply migrate to a competitor never to return.

In evaluating Table 1, it is important to understand that there are multiple processors available to handle the read/insert requests. Therefore, some fairly substantial workloads can be accommodated by the system provided that the queues do not get overloaded. In the case of the four workload density trial, only the last test (fifteen million reads) experienced queue overloading issues. The first three workload levels all exhibited similar values in the median latency and 95th percentile latencies. However, the 99.9th percentile latencies are lower for the one million, and roughly the same for the five and ten million record tests. The real difference is in the session time which increases significantly, but not quite linearly as the workload increases. So in other words, the service rate remains relatively constant but how long it takes to service the request increases with the load.

TABLE 1: MEDIAN LATENCIES/SESSION TIMES FOR A CASSANDRA DATABASE AT VARIOUS RECORD INTENSITIES

	one million		five million		ten million		fifteen million	
	insert	read	insert	read	insert	read	insert	read
median latency	1.4	1.7	1.6	1.7	1.6	1.8	1.7	NA
latency 95 percentile	2.6	2.9	2.8	3	2.8	3	3.2	NA
latency 99.9 percentile	38	22.8	81	91.7	73.8	91.1	824.4	NA
session time	0:58	0:54	5:01	4:57	9:53	9:16	17:00	NA

While generally speaking, the latencies do not differ greatly for reads versus inserts in a given category, inserting was typically quicker. In part, the slight difference between reads and inserts, as can be seen in Figure 1, may be related to the buffering process. For reads, each time a new and different record was read in a random fashion, just that record could be cached. However, with inserts the whole new record could be placed in a buffer without waiting and then actually inserted at the appropriate disk location. Also, the overhead of transferring data in a distributed node environment with a one-gigabit per second network connection may have exacerbated the latencies of the read process, which could not be treated as a "block" of data.

There was no inherent automated tuning process configured within the Cassandra implementation, so even though millions of records were processed, the system did not learn how to improve upon the basic read process. In other words, some databases allow for dynamic growth while others do not. For example, with Microsoft SQL Server this is handled most of the time as a percentage of the current size. So as the size of the data block increases there is an expected delay that occurs. The bigger the data block gets, the longer the gap and more infrequent the delay occurs. So therefore, database learning can also be a factor in how efficiently it will operate. Initial queries can thus take longer as the query optimizer creates more efficient plans to effectively fulfill them. As the plans developed by the optimizer improve, the performance can and should also improve. This is a concept that merits future investigation into how overall database efficiency may be affected.

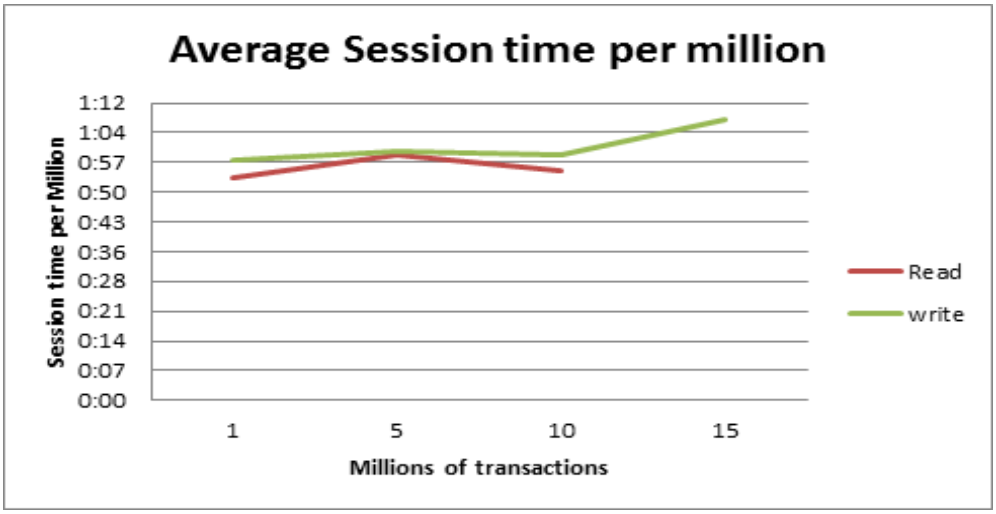


FIGURE 1: GRAPHICAL REPRESENTATION OF TABLE 1

DISCUSSION

From the data contained in Table 1 it is clear that generally speaking the Cassandra distributed database scales fairly well up to the ten million-intensity levels. At the fifteen million level the read experiment generated so many errors that it was unusable and a 99.9th percentile latency on the insert level of 800+ seconds indicated that the configuration was not adequate for inserts as well. The observed latencies at the one through ten million levels were encouraging, but need to be analyzed in the scope of a target response time. These targets may vary depending on the application supported. For example, if the application being considered is e-commerce then response time is critical. In this case, a total end-to-end response time of three seconds or less is critical to meet a customer's expectations, though the quicker the better. If this target metric is not met then customers may get impatient and migrate to another site, perhaps never to return. Therefore, a profit making enterprise would be well advised to carefully evaluate the performance gains Cassandra could provide. Based on the data reported herein, the metrics need to be evaluated in regard to the number of transactions that will exceed the desired threshold.

Given the data observed above one million records and a maximum database latency target of three seconds or less the metrics for both reads and inserts meet the target on the median and 95th percentile metrics. However, at this level the target is greatly exceeded at the 99.9th percentile. Because computer systems are dynamic in nature and workload for the most part does not follow a normal distribution it is very difficult to completely eliminate the outliers that are measured by the 99.9th percentile in this paper (Guster, 2002; Guster, Robinson, & Safonov 2005; Guster, Robinson, & Sundheim 2008; Rabl et al., 2012). Therefore, tuning a database can be tricky. Is it acceptable to have less than 40 out of one million transactions exceed the target by a factor of 10? Certainly a company needs to weigh the added cost. Perhaps, it might spend $50,000 to tune it and reduce the number of outliers from less than 40 to less than 20, would this be an acceptable return on investment? That would be for the company to decide. Further, as previously stated,

computer systems are dynamic and need to be monitored. It might not be unusual to observe an increase in outliers at a future date, which can often happen when software elements in the system have been upgraded or patched. While the magnitude of the median latencies observed at the one million level seem in line with a three second target, with an application such as an e-commerce website the more important metric is actually end-to-end delay.

In other words, one needs to take into account all of the components required to complete a transaction between a client and a server. In other words, one needs to take into account all of the components required to complete a transaction between a client and a server. A good representative example of this model is offered by Fleming (2004). This model can be summarized as follows: User Application Delay + CPU Delay on the Local Computer + Net work interface card (Ethernet) Delay on the Local Computer + Delay passing through the Network Switching devices + Delay passing across the network itself + Delay passing through the Network Switching devices on the receiving side + Net work interface card (Ethernet) Delay on the receiving Computer + Application service Delay + Disk controller Delay + Delay in the Disk Read Process, (…then the delay in traversing the path in reverse for the reply). Because there are many components which interact with one another even a small delay value like ~.0025 seconds can be significant and therefore, one needs to look at the total response time model which is quite complex and involves a number of components.

In another experimental study that dealt with reducing latency through adding nodes addressed the concept of end-to-end delay (Guster, O'Brien, & Lebentritt, 2013). While the Internet provides pretty good performance, considering its size and scope, one can expect delays of various magnitudes based on its workload at the time the transaction is consummated. For example, given that the network delay on the Internet in the U.S. could be described by the median value .5 seconds in each direction it is important to optimize each of the parameters. Further, one needs to realize that this whole algorithm is based on queuing theory, which means that there is an interaction among all the parameters in play. In other words, a delay of .0005 instead of .0001 at the first parameter won't simply result in .0004 seconds of additional response time. Rather, it will propagate through the entire algorithm and the delay will get a little longer with added wait time at each successive parameter. To put this in perspective, if one assumes a geometric progression through all 12 parameters in the algorithm above the result in total added delay would be close to one second (.8192) (Guster et al., 2013).

Besides the potential to scale effectively while maintaining reasonable performance the capabilities of Cassandra to support fault tolerance are quite useful. Furthermore, Cassandra is able to accomplish this feat by performing on inexpensive commodity hardware. Its design also allows for easy tuning of the original configuration. This situation is ideal for small/medium size companies that want to adopt Cassandra for its cost effectiveness and ease of use. Because of the ease that replication can be invoked it would be effective to use Cassandra as a central core of any disaster recovery plan. It could easily be configured to replicate within a company's existing cloud or within a cloud at a remote site making it simple for a company to have copies of their

mission critical data at several widely distributed remote locations (Guster, et al. 2013; Cockcraft, 2011).

CONCLUSIONS AND FUTURE STUDY

The data contained herein, was designed to ascertain how an eight-node Cassandra database would scale at high intensity workloads. Although the results were interesting and generally indicated respectable scaling there are still numerous complexities that need to be explored when implementing Cassandra. First, for the sake of simplicity in this study a single client generated the workload, which is not realistic. Using multiple clients would certainly be more realistic and certainly would alter the intensity and distribution of the incoming workload. Second, this data was collected using a configuration that featured only one replica. Given the distributed nature of most companies' data profile research is needed that addresses multiple nodes distributed world-wide. Third, while the workload was fairly intense, research is needed that will look at a particular workload and then try to tune the existing test bed to better accommodate that intensity. Fourth, the number of nodes was limited to eight. Although based on the latencies observed this appeared to be a logical starting point, in a production world one would expect the number of nodes to far exceed this value. While the results obtained herein answer some scaling related questions, there are still numerous unanswered questions. Therefore, future research will need to address configuration, workload and tuning variables to ascertain the effectiveness of Cassandra in various applications.

REFERENCES

Abadi, D. J., (2012). Consistency tradeoffs in modern distributed database system design: cap is only part of the story. *Computer*, 45 (2), 37-42.

Afyouni, H. A. (2006). *Database security and auditing: Protecting data integrity and accessibility*. Boston, MA: Course Technology, Cengage Learning.

Apache Software Foundation (2009). *Cassandra*. Retrieved from http://cassandra.apache.org

Bernstein, P., Rothnie, J. B., Goodman, N., & Papadimitriou, C. A. (1978). The concurrency control mechanism of sdd-1: a system for distributed databases (the fully redundant case). *IEEE Transactions on Software Engineering*, 4 (3), 154-168.

Bonifati, A., Panos K. C., Aris M. O., & Kai-Uwe S., (2008). Distributed databases and peer-to-peer databases: past and present. *ACM SIGMOD Record Archive*, 37 (1), 5-11. ACM New York, NY, USA doi>10.1145/1374780.1374781.

Brown, C., Guster, D. C., & Krzenski, S. (2007). Can distributed databases provide an effective means of speeding up web access times?. *Journal of Information Technology Management,* 18 (1), 1-15.

Cockcraft, A. (2011). Migrating netflix from a datacenter oracle to global cassandra. *Cassandra Summit Talk 2011*. Retrieved http://www.slideshare.net/adrianco/migrating-netflix-from-oracle-to-global-cassandra.

DataStax Corporation. (2013). Benchmarking top nosql databases. Retrieved from http://datastax.com/wp-content/uploads/2013/02/WP-Benchmarking-Top-NoSql-Databases.pdf.

Elnikety, S., Tracey, J., Nahum, E., and Zwaenepoel, W. (2004). A method for transparent admission control and request scheduling in e-commerce web sites. *Proceedings of the 13th International Conference on World Wide Web*, 276-286.

Featherston, D. (2010). Cassandra: Principles and applications. Retrieved from http://disi.unitn.it/~montreso/ds/papers/Cassandra.pdf.

Fleming, D. (2004). Network response time for efficient interactive use. *Proceedings of the 20th Computer Science Seminar, Addendum-T2-1*. RIP, Hartford Campus, April, 24.

Guster, D. C. (2002). A comparison of stochastic models for interarrival times of packets in a computer network. *Optimal Information Modeling Techniques*: IRM Press.

Guster, D. C., Robinson, D. H., & Safonov, P. I. (2005). Packet inter-arrival distributions in computer network workloads. *Encyclopedia of Information Science and Technology*. Hersey, PA: Idea Group.

Guster, D. C., Robinson, D., & Sundheim, R. A. (2008). Evaluating computer network packet distributions. *Encyclopedia of Information Science and Technology*. Hersey, PA: Idea Group.

Guster, D. C., O'Brien, A., & Lebentritt, L. (2013). Can a decentralized structured storage system such as cassandra provide an effective means of speeding up web access times. *Proceedings of Midwest Instructional Computing Symposium*.

Hemminger, C., Rogers, D. C., & Guster, D. C. (2010). Planning and managing the data center to green computing. *International Journal of Business Research*, 10 (4), 105-113.

Kanitkar, V. (2000). Collaborative and real-time transaction processing techniques inclient-server database architectures. *Polytechnic University*, 61 (04B), 2036.

Kanitkar, V. and Delis, A. (2002). Distributed query processing on the grid. *IEEE Transactions on Computers*, 51 (3), 269-278.

Lakshman, A. and Malik, P. (2010). Cassandra: a decentralized structured storage system. *ACM SIGOPS Operating Systems Review*, 44 (2), 35-40.

Rabl, T., Sadoghi, M., Jacobsen, H-A., Gomez-Villamor, S., Muntes-Mulero, V., & Mankovskii, S. (2012). Solving big data challenges for enterprise application performance management. Retrieved http://vldb.org/pvldb/vol5/p1724_tilmannrabl_vldb2012.pdf.

Ozsu, M. T., & Valduriez, P. (1991). Distributed database systems: where are we now? *Computer*, 24 (8), 68-78.

Otey, M. E., Parthasarathy, S., Wang, C., Veloso, A., & Meira, W. (2004). Parallel and distributed methods for incremental frequent itemset mining. *IEEE Transactions on Systems, Man & Cybernetics: Part B*. 34 (6), 2439-2450.

Safonov, P. I., Guster, D. C., & Hemminger, C. (2011). Employing host virtualization and symmetric multi-processing as a strategy for improving performance in computationally intense problems. *Issues in Information Systems*, XII (1), 357-365.

Xu, W., Jing L., Yongwei W., Xiaomeng, H. & Guangwen, Y. (2008). VDM: virtual database management for distributed databases and file systems. *Seventh International Conference on Grid and Cooperative Computing*, 309- 315.

Zimmermann, R., Wei-Shinn, K., & Wei-Cheng, C., (2004). Efficient query routing in distributed spatial databases. GIS '04 Proceedings of the 12th annual ACM: *International Workshop on Geographic Information Systems*, 176 - 183 ACM New York, NY, USA ©2004 table of contents ISBN:1-58113-979-9 doi>10.1145/1032222.1032249.

THE COMPARATIVE RELEVANCE OF RESOURCE HETEROGENEITY AND RESOURCE MOBILITY TO STRATEGIC MANAGEMENT

Rogerio S. Victer, Fairleigh Dickinson University
rvicter@fdu.edu

ABSTRACT

The objective of this paper is to offer a refined understanding of the relationship between resource and performance. Drawing from the resource-based theory (RBT) of competitive advantage, this study compares the relative relevance of core resource characteristics to the implementation of strategic activities. This approach provides an important clarification of the best managerial path towards differential performance. Should managers focus primarily on pursuing resource heterogeneity or preventing resource mobility? Results generated by this study point to the fact that resource heterogeneity has a greater impact over performance than resource mobility. The paper concludes that strategies focused on raising barriers to imitation are comparatively less relevant than strategies focused on promoting innovation.

INTRODUCTION

The resource-based theory (RBT) of competitive advantage recommends managers to identify and acquire critical resources capable of generating sustainable competitive advantage. According to RBT, managers are expected to accumulate resources that are both valuable and rare (Coff et al., 2006). In practice, accomplishing this managerial task comprises different implementation activities. Managers are required to gain access to critical resources as well as to increase barriers to their dissemination. RBT shows that strategic management consists of two fundamentally independent but complementary resource-oriented activities: one activity supports value creation whereas the other precludes value erosion. Yet we still do not know which one of these activities is the most strategic for business firms. Should managers focus primarily on pursuing resource heterogeneity or preventing resource mobility?

The current paper intends to address this core question as a way to reaffirm the relevance of RBT to the strategic management discipline. However, instead of measuring how much a critical resource influences competitive advantage, which has been the main approach adopted in this research tradition (Newbert, 2007), the main objective of the present study is to compare and contrast the performance relevance of different characteristics of the same resource. This allows RBT to be more specific and useful to management, instead of being an abstract theoretical model with little practical relevance (Priem & Butler, 2001). In order to accomplish this objective, the paper proposes to revisit the RBT theoretical framework as a condition to better explain the subtleties underlying the relationship between resource and performance.

Suggestion for change focus on a single argument: the main factor explaining superior performance is not *resource identity* but *resource attribute*. Differently from most studies in this tradition, this paper advocates that it is not the resource itself but some of its key economic traits that actually generate differential benefits to firms (Markman et al., 2004; Newbert, 2008; Victer, 2014). The managerial task relies on the ability to recognize those resources that are strategic (Barney, 1986) and also, and possibly more importantly, on the ability to quantify which aspects of the resource are the most relevant to practice. This approach goes beyond the traditional task of identifying the characteristics resources must possess (Barney & Arikan, 2001) in order to objectively rank them according to their respective influence on business outcomes. Having a clear indication on which resource attributes are the most relevant to performance increases the overall effectiveness of a firm's strategy, since managers can tailor their actions towards leveraging existing resources or developing new resources with the appropriate profile.

The formal comparison of resource attributes is not only relevant to managerial practice, but also to theory development as well. Given the complexity involved in the operationalization of foundational constructs, RBT has had difficulties in consolidating its academic relevance (Markman et al., 2004). Empirical measurements of core variables as well as their inter-relationships are part of a broader process of theory development. In order to be academically relevant, RBT must clarify the criteria in which strategic resources can be assessed according to separate and independent conditions (Priem & Butler, 2001). In this respect, the empirical agenda serves to enhance the appropriateness of the conceptual model as the reflection of a synthetic theory, thereby generating a more precise understanding of the nature and magnitude of the relationship between resource and performance.

In order to measure and compare the relative relevance of resource attributes to performance, we rely on data generated by the AIDS/HIV pharmaceutical industry and adopt technological knowledge embedded in drugs as a proxy for a strategic resource. Technological knowledge is a particularly critical resource in science-based industries (e.g., pharmaceutical, chemistry, and information technology). Access to technological knowledge is the precondition for any pharmaceutical firm to commercialize a drug. Technological knowledge is an example of a case in which the resource is not homogeneous. There are many kinds of technological knowledge and they differ from each other in the manner they resolve a problem or propose a solution for a particular challenge. Results generated by this study corroborate the assumption that technological knowledge is a highly differentiated resource and point to the fact that some knowledge characteristics are more relevant to outcomes than others. This empirical evidence suggests that the most effective way to generate a sustainable competitive advantage is to promote sources of resource heterogeneity instead of preventing sources of resource mobility.

THEORY

The reason a resource contributes to a firm's competitive position is related to its characteristics (Collis & Montgomery, 1995). Managers need to be able to identify which traits of the resource

most influence the process of gaining and sustaining competitive advantage. A resource is not a homogeneous, undifferentiated construct. Variation occurs even within the same class of a key resource. Consequently, the predictor of performance is not the resource *per se*, but its core relevant attributes. This is because the strategic relevance of a resource partly reflects changes in evolutionary forces and temporal conditions within particular industries (Aragon-Correa & Sharma, 2003; Polidoro & Toh, 2010; Sirmon et al., 2007). The resource is not relevant in absolute terms, but contingent to environmental constraints. Resource attributes might gain or lose relevance depending on the results of a complex dynamic competition happening both at the factor and product markets.

The difference in approach might be subtle, but is quite relevant to the theoretical project. Instead of assuming that the resource is strategic due to its inherent value and long-term rareness, the theory builds more precision by highlighting the conditions in which a resource becomes relevant. It clarifies the reason why the resource is important by defining specific paths in which value is created and sustained. Therefore, the research interest is not to isolate the resource from the context of its application as done in traditional RBT studies. On the contrary, what matters the most is the understanding of the functionality of the resource in a given historical condition. A resource becomes strategic to the degree in which its characteristics improve or deteriorate performance of end products according to its availability as well as its ability to meet expectations of end consumers.

This specific approach does not refute the core assumption of RBT. Actually, it intends to strengthen it through the clarification of the means by which the resource influences performance. Although the original idea was based on the principle that valuable and rare resources are those that give the firm competitive advantage, here we suggest that a resource gains strategic relevance in the proportion of being valuable and rare. The major consequence resulting from this reasoning is to consider resources' relevance more flexible and contingent. Not only the environment affects the valuation process, but managers can also exert influence on firm's resource profile. Management can choose alternative strategic actions on how to best leverage a particular resource attribute. As part of this process, it is essential that firms constantly assess the relevance of resources according to situations in which they are or are not relevant to business outcomes.

Traditional RBT defines resources as valuable when they enable a firm to implement strategies that improve its efficiency and effectiveness or when they exploit opportunities or neutralize threats (Barney, 1991). In the same line of reasoning, RBT defines resources as rare when they are not possessed or controlled by others, simply because firms implementing similar strategies would be unable to differentiate their outcomes (Barney, 1991). This way of thinking generates a binary model of competitive advantage in which the strategic resource either present or absent. Here our task is to suggest a less trivial model by focusing on the variation in strategic relevance of a critical resource. This means that we give more attention to the *degree of relevance*, instead of assuming *absolute relevance* of a resource. The methodological implication is to be less focused on measuring the direct relationship between resource and performance and more

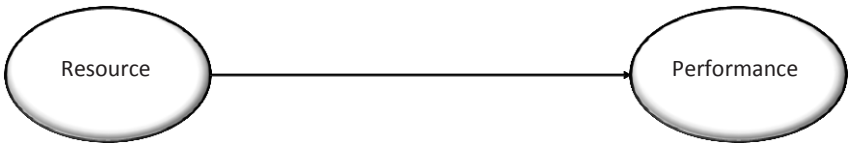
concerned with measuring how the resource becomes relevant to performance through the mediation of its core attributes. This means that performance is conditioned by the attributes of heterogeneity and immobility.

This subtle difference is illustrated in *Figure 1* below, which depicts the difference between the received and the proposed approach in RBT. In one, the resource is assumed to be heterogeneous and immobile, which allows it to influence performance; while in the other, a resource is strategic only in the degree to which its attributes of heterogeneity and immobility are relevant to performance. The model applied in the current paper promotes the latter, less popular approach. In line with this rationale, we suggest two core propositions based on RBT classical arguments regarding the heterogeneity and immobility of resources. The propositions are articulated in such a way as to incorporate fundamental dichotomy between creation and erosion of value, such that:

Proposition #1: *Resource heterogeneity increases performance*.

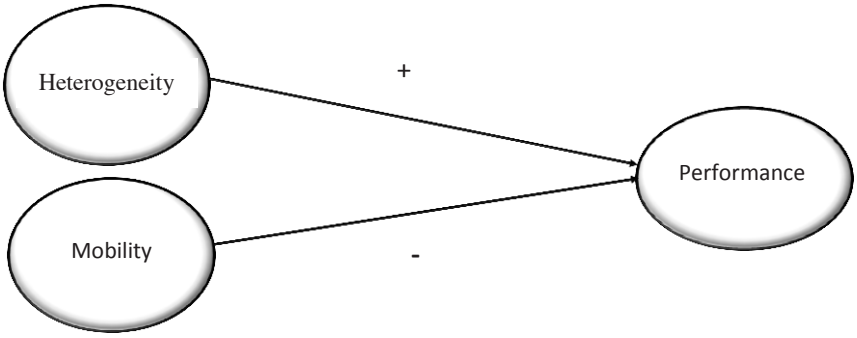
Proposition #2: *Resource mobility decreases performance*.

a) Received Model



Heterogeneous and Immobile

b) Proposed Model



Resource

FIGURE 1. COMPARISON BETWEEN THE RECEIVED AND THE PROPOSED RBT MODEL

HYPOTHESES

Both heterogeneity and mobility are complex notions. They require multiple empirical constructs to capture their aggregated influence on performance. Here we suggest measuring each one of them through the combination of three distinct measurable dimensions. Heterogeneity is measured by *distinctiveness*, *maturation*, and *integration*. These dimensions represent ways in which there is a process of generation of value to the end consumer. Mobility is measured by *commonness*, *diffusion*, and *replacement*. These dimensions represent ways in which there is a process of erosion of value in the eyes of the end consumer. Each one of these separate dimensions delineates a basic hypothesis relevant to the relationship between the resource and performance.

Resource Heterogeneity

Resources are not only unevenly distributed across firms, but they are also differentiated according to their usefulness in creating value to customers (Sirmon et al., 2007). The notion that resources are heterogeneous is derived from the fact that the same resource can be unique in specific and diverse ways. Resource heterogeneity reflects variation in the resource profile or composition sufficiently relevant to affect outcomes. Even the same resource might have different production levels (Ricardo, 1817), which generates the notion of superior resources (Peteraf, 1993). We suggest three ways to capture the path how resources gain the attribute of heterogeneity, as follows:

1. Distinctiveness

A critical resource generates value essentially by having a direct functional application in the resolution of a problem relevant to the end consumer (Christensen & Raynor, 2003). Resources embedded in products are effective due to their technical competence in achieving a particular objective. It carries with it a particular methodology to deal with a trouble. Even the same resource can have multiple intrinsic distinctions on how to deal with a challenge. From this perspective, it is important to recognize the existence of different kinds of the same resource, such that:

Hypothesis #1: *Variations of the same resource have distinct influence on performance*.

2. Maturation

Heterogeneity of the same resource is further magnified through modifications and improvement in the resource's functionalities. The strategic element here is not necessarily to create a new resource, but to better perfect the already established methodology for action. This process is made possible by the result of learning and experience. A new generation of a resource has potentially the opportunity to magnify its original benefits or minimize its shortcomings. Consequently, the continuous development of a resource creates additional advantages to the end

consumer. This means the ability to overcome weaknesses in previous resource characteristics and improve particular effects, such that:

Hypothesis #2: *A new generation of the same resource has a positive effect on performance*.

3. Integration

Resource functionality can be also improved through the combination of distinct variations of the same resource. Bundles of resources of the same class can also generate an additional portion of value because of the effect of complementarity. Combining resources together in the same integrated design is a way to enhance overall effectiveness of the end product. This form of combination can be potentially performed by the customer during the consumption process. Each end consumer has the ability to customize products to his or her own particular tastes or needs. Yet firms can promote combination of resources during the manufacturing and delivery stages of the value chain, thereby generating value through convenience and speed of use. Additional sophistication and complexity of formulation can even make the integration of resources by the consumer unfeasible for practical reasons. As a result, the integration of different resources as part of the product design can potentially increase performance through simple addition or more complex synergistic interactions. Therefore, we suggest that:

Hypothesis #3: *The combination of different kinds of the same resource has a positive effect on performance*.

Resource Mobility

The proliferation of a critical resource neutralizes its intrinsic value. Increasing accessibility to valuable resources affect the relative attractiveness of the products in which they are embedded. Mobility is not an absolute dimension of the resource, but a derived one which is affected by the nature of competition in the factor market. The quantity of players possessing or controlling similar resources ultimately affects the resource's relative value. The notion that resources are mobile (or conversely sticky) serves as a metaphor to represent how much a critical resource can be captured simultaneously by many players operating in the same market. If the resource is tangible, one can imagine an actual process of spatial distribution of the resource. If the resource is intangible, one should consider that resources are distributed only figuratively through a process of cultural or cognitive transmission. Many different processes can underlie the actual effect of a resource's increasing accessibility to direct or indirect rivals. We suggest three ways to capture the paths how resources gain the attribute of mobility, as follows:

4. Commonness

A product's performance advantage highly depends on the rarity of its underlying resources. Even functionally relevant resources needs to be exclusive, unique, or "one of a kind" as a condition for being valuable (Barney, 1991). If the same resource is shared by numerous existing rivals, then this resource, although intrinsically efficient and useful, is unlikely to generate performance differentials or any other type of business-related advantage. Relevant but common resources are, at best, simple sources of competitive parity, such that:

Hypothesis #4: *The number of products based on the same kind of resource has a negative effect on performance*.

5. Diffusion

Imitation by rival firms is the most common threat to the relevance of a particular resource class. Imitation from rivals dissipates the economic relevance of the resource applied in products. The critical resource circulates and gets diffused, reducing its uniqueness or originality even in case of property rights. The complexity and specificity of critical resources also tend to play a relevant role as barriers to imitation (McEvily & Chakravarthy, 2002). They basically increase resource stickiness by increasing the costs of transferring resources across organizational boundaries. We can assume that the temporal flow of resources across firms generates a particular pattern of replication and adoption, affecting the speed in which competition introduces new products with the same existing resource.

Hypothesis #5: *The age of the resource has a negative effect on performance*.

6. Replacement

In addition to imitation, another relevant aspect of the resource's mobility attribute is the impact generated by substitutes. In addition to direct rivalry, resources are also confronted with the threat of replacement. This means that an existing product applying a specific resource faces competition not only from products based on the same resource, but also from those products built on alternative kinds of the same resource, but with similar benefits. In fact, the dominance of any particular resource carries with it a threat of replacement, because a problem to be solved requires a solution independent of the method of solving it (Schumpeter, 1934). Even when rivals cannot perfectly imitate a firm's critical resource, they can often create alternatives with compatible functionality. Substitution can sometimes be even more detrimental to incumbents than imitation or direct rivalry, because it threatens not only the product, but also a firm's existing resource profile (Polidoro & Toh, 2010).

Hypothesis #6: *The number of products using alternative kinds of the resource has a negative effect on performance*.

METHODOLOGY

In order to measure the relationship between these resource attributes and performance, we rely on empirical data generated by branded drugs commercialized in the AIDS/HIV pharmaceutical industry. Technological knowledge embedded in drugs is used as a proxy for a critical resource. The data is organized as a panel data set (Hsiao, 1986) of global sales of antiretroviral drugs along a decade (1999-2008). The distinctiveness of the panel data applied here is that it contains complete measures of each drug's market trajectory. Consequently, there is a favorable condition for clarification of underlying causal interrelationships between the independent variables and the dependent variable through the application of a time series methodology (Finkel, 1995). The

panel data covers 26 different brands or approximately 90% of all branded drugs commercialized during the selected period. Technological knowledge is subject to patent and, therefore, should be considered useful, non-trivial, and unique. Consequently, technological knowledge can be assumed to be a valuable and immobile resource (de Carolis, 2003). However, the most relevant research issue here is not whether firms having access to patented technological knowledge have a competitive advantage over firms without access to it, since this is in fact a pre-condition for commercialization of a final pharmaceutical product. Rather, what matters the most for the purpose of this study is to learn how different technological knowledge generate distinct levels of product efficiency. Ultimately, the main question is: which knowledge attributes are the most important? We suggest capturing performance through the market value of a product as represented by *log* of global sales per year, and operationalize the attributes of knowledge through drugs' mechanisms of action (MOA). According to the National Institute of Allergy and Infectious Diseases (NIAID 2012), there are six major types of drugs used to treat HIV patients. These drugs are grouped by how they interfere with steps in the virus replication. They are Nuclear Reverse Transcriptase Inhibitors (NRTI), Non-Nuclear Reserve Transcriptase Inhibitors (NNRTI), Protease Inhibitors (PI), Fusion or Entry Inhibitors (EI), and Integrase Inhibitors (II). New MOAs have been discovered recently, but since they were not applied in drugs available in the market before 2009, they are not formally considered in the analysis.

Knowledge attributes are operationalized as follows:

a) *Distinctiveness* is captured through a categorical variable representing the underlying MOA of a drug;
b) *Maturation* is captured through a categorical variable representing whether the MOA underlying a drug is from the first or second generation;
c) *Integration* is captured through a categorical variable representing whether the drug is composed on one or more MOA;
d) *Scarcity* is captured through a continuous variable representing the number of drugs using the same MOA;
e) *Diffusion* is captured through a continuous variable representing the age of the MOA applied by a drug; and
f) *Replacement* is captured through a continuous variable representing the number of drugs using alternative MOAs.

In order to control for the uniqueness of each drug product and any other relevant influence on product's market value, the model also controls for the brand-specific effect and for the systematic temporal effect. The brand-specific effect captures everything related to the product that is not directly related to the mechanism of action adopted by the embedded technological knowledge, such as the quality of a particular biological or chemical compound as well as the effectiveness of the marketing campaign, among other potentially relevant aspects that are specific to a particular brand. In addition, any other relevant effect influencing the market value of a drug is captured through the auto-correlation of the residuals. The model adopts a first-order autoregressive methodology (*AR-1*), which lies in the assumption that disturbances on residuals are considered to depend only on its own previous value (the Markov property) and a random,

"white noise" component. The coefficient of the autocorrelation of the residual (ρ) reflects the aggregated influence of other potentially relevant and persistent factors influencing the performance of drugs that is not directly controlled by the adopted independent variables. This procedure also serves to correct biases generated by the non-stationary nature of the panel data according to the Prais-Winston procedure (Wooldridge, 2002).

RESULTS

Figure 2 displays the coefficient of determination (R^2) for partial models as a way to analyze each separate hypothesis. Based on these results, hypotheses #1, 2, 3, and 5 are accepted, while hypotheses #4 and 6 are rejected. *Exhibit 1* displays the result for the full model, which also corroborates main findings, even when all variables are considered simultaneously. Before discussing more details, it is important to highlight that, given that sales is *log*-transformed, but variables representing knowledge attributes are not, interpretation of the results requires a relatively simple transformation. We can easily interpret the regression coefficient in an *OLS model* as the expected change in log of the dependent variable with respect to a one-unit increase in the independent variable, holding all other variables at a fixed value. The natural way to do this is to exponentiate the regression coefficient, since exponentiation is the inverse of a logarithm function.

Below we discuss main results:

- ***Distinctiveness (Hypothesis #1):*** The first hypothesis is accepted, since the model considering the effect of MOA distinctiveness in product performance is statistically significant with a coefficient of determination (R^2) of 5.71% (at $p < .05$). This test confirms that MOAs applied in drugs differs from each other. A more detailed analysis of the result in the full model allows the conclusion that NRTI and II are not significantly different from each other, but PI, NNRTI and EI are. NNRTI are expected to generate less 96% of sales than NRTI, PI less 94%, and EI less 92%, assuming all other variables are controlled. On the other side, II is expected to generate similar levels of sales than NRTI. This result suggests that drugs based on NRTI and II are the most effective ones;

- ***Maturation (Hypothesis #2):*** The second hypothesis is also accepted. The model considering the effect of different generations of MOA is statistically significant with a R^2 of 15.31% (at $p < .001$). Results indicate that the second generation is significantly higher than the first generation. The model estimates that a new version of MOA is likely to generate an increase of sales in order of more than 510%. This is an indication that once a new generation of MOA is launched, drugs adopting it tends to capture market share quickly;

- ***Integration (Hypothesis #3):*** The third hypothesis is also accepted. The model considering the effect of the number of MOA in the same drug is statistically significant with a R^2 of 4.97% (at $p < .005$). Results indicate that multi-MOA drugs are likely to generate 380% more sales than drugs using a single MOA. This result seems to corroborate the idea that

incorporating multiple compounds in the same drug design generates considerable more value in the eyes of the patient (and/or physician) than drugs with just one compound;

- ***Commonness (Hypothesis #4)*:** The fourth hypothesis is rejected, since the number of existing products, products applying the same MOA, is not statistically significant. The intercept of the partial model is significant, indicating that the unconditional geometric average of an HIV drug is $337 million. This overall average does not significantly change with the increase (or decrease) of an additional competing resource of the same kind, indicating that direct MOA rivalry is not a relevant condition to change expected market performance of a HIV drug;

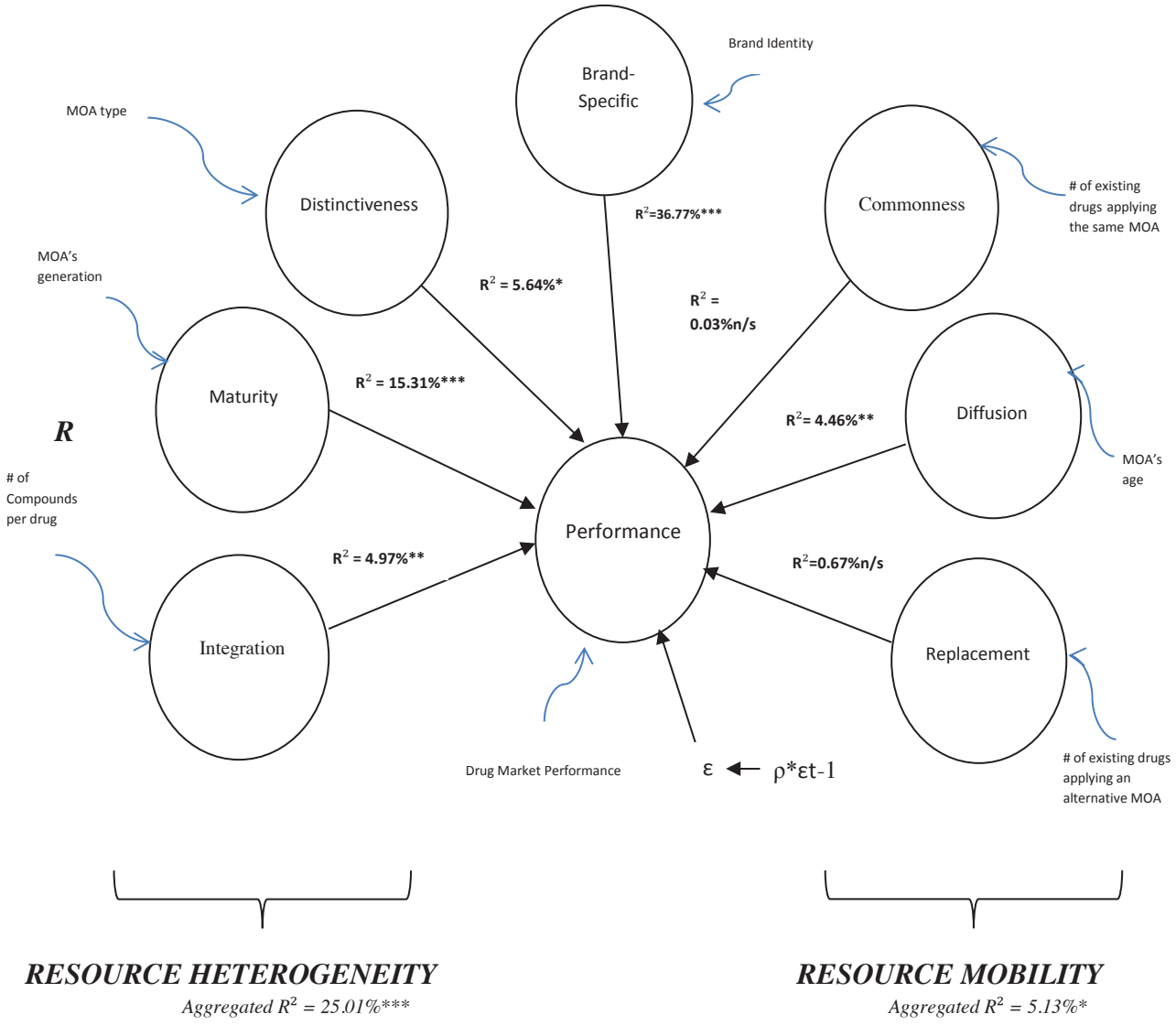


FIGURE 2. COEFFICIENT OF DETERMINATION PER CONSTRUCT/HYPOTHESIS

EXHIBIT 1. RESULTS FOR THE FULL MODEL

Variables	Coefficient	Std Error	Level of Significance
DISTINCTIVENESS			
NRTI (Constant)	9.33	0.78	0.000
PI	-3.04	0.62	0.000
NNRTI	-3.49	0.72	0.000
EI	-2.63	0.87	0.003
II	-0.33	0.52	0.721
MATURATION			
2nd Generation	1.81	0.52	0.001
INTEGRATION			
Multi-MOA	1.57	0.54	0.004
COMMONNESS			
Rivals	0.01	0.03	0.722
DIFFUSION			
Age of MOA	-0.20	0.03	0.004
REPLACEMENT			
Substitutes	-0.01	0.02	0.722
DRUG-SPECIFIC			
Kalestra	n/a		
Norvir	2.38	0.77	0.003
Reyataz	2.09	0.75	0.006
Sustiva	3.71	0.79	0.000
Viracept	0.76	0.73	0.297
Atripla	-3.13	0.91	0.001
Combivir	-0.27	0.75	0.714
Crixivan	1.48	0.73	0.044
Entriva	-1.31	0.53	0.015
Epivir	0.39	0.53	0.460
Epzicom	0.55	0.81	0.496
Fuzeon	-0.78	0.73	0.916
Invirase	0.33	1.23	0.786
Isentress	omitted		
Lexiva	1.37	0.75	0.070
Prezista	1.98	0.83	0.019
Rescriptor	-1.19	0.81	0.145
Retrovir	-1.73	0.65	0.009
Selzentry	omitted		
Trivizir	-0.86	0.76	0.259
Truvada	omitted		
Videx	-0.56	0.52	0.284
Viramune	omitted		
Viread	omitted		
Zerit	-0.58	0.49	0.284
Ziagem	omitted	0.78	0.000
# of Observations	**184**		
R^2	**0.6902*****		
ρ	**0.6188**		

Note: Some coefficients are omitted due to collinearity; n/a not applicable

- ***Diffusion (Hypothesis #5):*** The fifth hypothesis is accepted. The model considering the effect of MOA age is statistically significant with a R^2 of 4.46% (at $p < .005$). This variable's coefficient is -0.1925, meaning that a drug is expected to decrease sales in the order of 17% for every additional year in the life cycle of a MOA, possibly as the result of new entries of generics, which is not precisely measured due to the lack of available information. However, it is important to highlight that the estimated constant for this model is rather large ($44 billion of dollars);

- ***Replacement (Hypothesis #6):*** The sixth hypothesis is not accepted, since the number of existing substitutes, or the number of products applying a different MOA than product *i* in year *t*, is not statistically significant. This means that the expected performance level of a HIV drug is unlikely to change simply because a new substitute resource enters the AIDS pharmaceutical industry.

The aggregation of all six knowledge-based variables generates an equivalent R^2 of 32.25% (at $p < .001$) as displayed in *Exhibit 1* above. This means that they are capable of explaining a third of the variance of drug sales. When controlled by drug-specific effect, the consolidated R^2 double to 69.02%. Nine out of the 26 brands in the sample are significantly different from each other, whereas 17 are not. The magnitude of ρ is also high at the level of .62, showing strong degree of auto-correlation of the residuals during the period under consideration. Focusing on the six main variables of interest for this study, it is possible to assert that the most relevant knowledge attribute is maturation. It is followed by distinctiveness, integration, and diffusion. In aggregate, we can conclude that resource heterogeneity, with an aggregated R^2 of 25.01% (at $p < .001$), is more important that resource mobility, with an aggregated R^2 of 5.01% (at $p < .05$).

DISCUSSION

The objective of this paper is to better understand the complexities of the relationship between resource and performance. The study intends to support both theory development and managerial practice. From a theoretical perspective, the paper suggests a model in which the relationship between resource and performance is contextualized to a particular competitive condition within a given industry. Unlike previous papers in the RBT tradition which assumes that a resource (or its accumulate stock) have a direct effect on performance, this paper examines how distinct resource attributes contribute to performance differentiation within a historical situation. This approach has the potential of refining the original RBT assumption: it shows that it is not the resource, but certain characteristics of the resource that ultimately impacts performance outcomes. Consequently, the paper points to the relevance of the conditions that make a resource relevant to performance either of products or firms. This approach clarifies critical paths in which a resource can generates differentiated performance.

The main consequence of this theoretical perspective is to support an alternative version explaining how performance differentiation is possible in a competitive market. It shows that the

same critical resource might have distinct levels of efficiency according to the industrial circumstances. A resource is not either strategic or non-strategic for the firm; rather the resource becomes more or less strategic as the consequence of a competition occurring between different kinds of the same resource. Resource attributes might gain and loose relevance over time. This shows the impact that management can have in the positioning of firms from a practical perspective. Understanding the conditions in which knowledge affects performance has a direct implication to managerial decision making and execution. One of the consequences of focusing research attention on the attributes of the resource instead of focusing on the resource itself is to emphasize the relevance of implementation activities. This approach is particularly appealing to strategic management, because it allows the firm to intentionally manage and even design its resources according to particular historical needs. The specific relevance of a resource attribute informs management on the best stream of action.

Effective strategies are directly related to managerial processes which are oriented towards either promoting value creation or preventing value erosion. Firm's quest for competitive advantage motivates managers to focus on strategic actions oriented towards effectively leveraging resources. Here we have learned which resource attributes are the most relevant in the AIDS/HIV pharmaceutical industry. Empirical results indicate that the characteristics of knowledge captured by the attributes of heterogeneity (i.e., distinctiveness, maturation, and integration) are more relevant to performance than the attributes of mobility (e.g. commonness, diffusion, and replacement). This empirical evidence suggests that competitive advantage is essentially generated through the systematic improvement of the functionality of technological knowledge. Data-based evidence on the AIDS pharmaceutical industry in the beginning of the 21st Century suggests that the most effective way to generate a sustainable competitive advantage is to stimulate sources of resource heterogeneity. Although the flow of resource is also relevant phenomenon, promoting knowledge heterogeneity is considerably more relevant than preventing knowledge mobility. This result seems to diminish the relevance of strategies anchored in raising barriers to imitation comparatively to strategies anchored in promoting innovation. Although diffusion of the same resource certainly depreciates competitive advantage, as the accumulated process of new entries, it is the modification of the resource that is the most relevant factor generating changes in product outcomes.

Although this paper helps to distinguish which resource characteristic is the most relevant to performance, innumerable questions still remains unanswered regarding the reason underlying this phenomenon. Why is resource mobility less relevant than resource heterogeneity? We might start addressing this question by assuming that there are in fact many potential reasons for this finding. One potential explanation is to consider that it is not the number of rivals that really matters to performance, but their quality in pursuing effective strategies. In this respect, more important than the amount of rivalry is the particular design of each rival product. This phenomenon is being captured here through the brand-specific effect, which in fact explains a considerable amount of sales variance. The low magnitude of the effect generated by resource mobility might also be explained through the nature of the industry or its existing life cycle. This is particularly relevant to the AIDS pharmaceutical industry that deals with a new and fast expanding disease based on a virus that can mutate in response to medications. In this case,

results could not be easily generalized to other pharmaceutical sectors or other industries. Another possibility is that the selected variables representing resource mobility are not measured appropriately, which might indicate an inefficient process of measurement. The lack of data on specific generics drugs introduced during the period being covered by the analysis, especially in developing countries such as India, South Africa, and Brazil, might represent a major flaw in the current dataset. However, we argue that measuring the effect of new entries can be done successfully through the dimension of time flow. Although this means to accept a certain degree of simplification, it also promotes a powerful construct based on the depreciation or "aging" of the resource - even for an intangible and non-rival resource as technological knowledge. Knowledge aging might be in fact a more robust construct than a direct measure of competitiveness simply because it is more important to account for the net effect of resource mobility in this particular case than to account for the gross effect of separate factors in particular years.

A good example for that is the case of the brand Atripla that managed to gain market share very quickly, although it is composed of already existing MOAs. In this particular case, the performance of the brand is generated by a process involving both knowledge heterogeneity (i.e., integration of MOAs) and knowledge mobility (i.e., application of existing MOA). In other words, it is possible to say that Atripla is the consequence of a simultaneous process of imitation and improvement. This example shows the sophistication of the phenomenon being measured, since imitation might also involves some form of design improvement (McEvily & Chakravarthy 2002). By measuring the dimension of commonness through time flow allows a methodological procedure to discount for unnecessary variability of the data and directly deal with the aggregated effect of overall MOA devaluation and value erosion. Future studies will have to revisit these issues in more detail and provide additional evidence on the complex dynamics affecting the relationship between resource and performance.

REFERENCES

Aragon-Correa, J. A., & Sharma, S. (2003). A Contingent Resource-based View of Proactive Corporate Environmental Strategy. *Academy of Management Review*, 28, 71-88.

Barney, J. (1986). Strategic Factor Markets. *Management Science*, 32, 1231-41.

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17, 99-120.

Barney, J. (2001). Is the Resource-based View a Useful Perspective for Strategic Management? Yes. *Academy of Management Review*, 26, 41-56.

Barney, J., & Arikan, A. (2001). The Resource-based view: Origins and implications. *Handbook of Strategic Management*. Hitt et al. (eds), Blackwell: Oxford, UK, 124-188.

Christensen, C., & Raynor, M. (2003). *The Innovator's Solution: Creating and Sustaining Successful Growth*. Harvard Business Press: Boston, MA.

Collis, M. & Montgomery, C. (1995). Competing on Resources. *Harvard Business Review*, 73, 118-128.

Coff, R., Coff, D., & Eastvold, R. (2006). The Knowledge-Leverage Paradox: How to Achieve Scale Without Making Knowledge Imitable. *Academy of Management Review*, 31, 452-465.

De Carolis, D. (2003). Competences and Imitability in the Pharmaceutical Industry: an Analysis of their Relationship with Firm Performance. *Journal of Management*, 29, 27-50

Finkel, S. (1995). *Causal Analysis with Panel Data,* Sage University Paper Series on Quantitative Application in the Social Sciences, 07-105: Thousand Oaks, CA

Hsiao, C. (1986). *Analysis of Panel Data,* Cambridge University Press 1st ed.: Cambridge, UK.

Markman, G., Espina, M., & Phan, P. (2004). Patents as Surrogates of Inimitable and Non-Substitutable Resources. *Journal of Management*, 30, 529-544.

McEvily, S., & Chakrawarthy, B. (2002). The Persistence of Knowledge-based Advantage: an Empirical Test for Product Performance and Technological Knowledge. *Strategic Management Journal*, 23, 935-952.

National Institute of Allergy and Infectious Diseases (2012). Understanding HIV/AIDS. http://www.niaid.nih.gov/topics/hivaids/Pages/Default.aspx, accessed on January, 2014.

Newbert, S. (2007). Empirical Research on the Resource-based View of the Firm: an Assessment and Suggestions for Future Research. *Strategic Management Journal,* 28, 121-146.

Newbert, S. (2008). Value, Rareness, Competitive Advantage, and Performance: A Conceptual-level Empirical Investigation of the Resource-based View of the Firm. *Strategic Management Journal***,** 29, 745-68.

Peteraf, M. (1993). The Cornerstone of Competitive Advantage: A Resource-based View. *Strategic Management Journal*, 14, 179-191.

Polidoro, F., & Toh, P. (2010). Letting Rivals Come Close or Warding Them Off? The Effects of Substitution Threat on Imitation. *Academy of Management Journal*, 54, 369-92

Priem, R., & Butler, J. (2001). Is the Resource- based "View" a Useful Perspective for Strategic Management Research? *Academy of Management Review*, 26, 22-40,

Ricardo, D. (1817). *Principles of Political Economy and Taxation*. J. Murray: London.

Schumpeter, J. (1934). *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*, Oxford University Press: London.

Sirmon, D., Hitt, M., & Ireland, D. (2007). Managing Firm Resources in Dynamic Environments to Create Value: Looking Inside the Black Box. *Academy of Management Review*, 32, 273-292.

Victer, R. (2014). The Strategic Implications of Knowledge Attributes: Understanding the Conditions in Which Knowledge Matters to Performance. *Management Decision*, 52, 3.

Wooldridge, J.M. (2002). *Econometric analysis of cross section and panel data*, MIT Press: Cambridge, MA.

JOURNAL OF INTERNATIONAL BUSINESS DISCIPLINES

Volume 9, Number 1 May 2014

Published By:
International Academy of Business Disciplines and Frostburg State University

ISSN 1934-1822 WWW.JIBD.ORG

Printed in the USA by Commercial Press Printing Company