HARVARD | BUSINESS | SCHOOL

The Customer May Not Always Be Right: Customer Compatibility and Service Performance

Ryan W. Buell
Dennis Campbell
Frances X. Frei

Working Paper 16-091

The Customer May Not Always Be Right: Customer Compatibility and Service Performance

Ryan W. Buell
Harvard Business School

Dennis Campbell
Harvard Business School

Frances X. Frei
Harvard Business School

Working Paper 16-091

Funding for this research was provided in part by Harvard Business School.

The Customer May Not Always Be Right: Customer Compatibility and Service Performance

Ryan W. Buell, Dennis Campbell, Frances X. Frei
Harvard Business School, Harvard University, Boston, Massachusetts 02163
[rbuell@hbs.edu, dcampbell@hbs.edu, ffrei@hbs.edu]

Abstract:
This paper investigates the impact of customer compatibility – the degree of fit between the needs of customers and the capabilities of the operations serving them – on customer experiences and firm performance. We use a variance decomposition analysis to quantify the relative importance of customer, employee, process, location, and market-level effects on customer satisfaction. In our models, which explain roughly a quarter of the aggregate variance, differences among customers account for 96-97% of the explainable portion. Further analysis of interaction-level data from banking and quick service restaurants reveals that customers report relatively consistent satisfaction across transactions with particular firms, but that some customers are habitually more satisfied than others. A second set of empirical studies provides evidence that these customer-level differences are explained in part by customer compatibility. Customers whose needs, proxied by differences in demographics and product choices, diverge more starkly from those of their bank's average customers report significantly lower levels of satisfaction. Consistently, banks that serve customer bases with more dispersed needs receive lower satisfaction scores than banks serving customer bases with less dispersed needs. Finally, a longitudinal analysis of the deposit and loan growth of all federally insured banks in the United States from 2006-2017 reveals that customer compatibility affects a firm's financial performance. Branches with more divergent customers grow more slowly than branches with less divergent customers. Institutions serving customer bases with more dispersed needs have branches that exhibit slower growth than those of institutions serving customer bases with less dispersed needs.

*[Keywords**:** Service operations, customer compatibility, satisfaction, profitability]*

1. Introduction

It has been well established that it is advantageous for companies to satisfy their customers. Satisfied customers are more loyal (Anderson 1994, Oliver 1980), purchase more frequently (LaBarbera and Mazursky 1983), have a higher willingness to pay (Homburg et al. 2005), and are more profitable over the long run than dissatisfied customers (Hallowell 1996, Heskett et al. 1997). As such, managers continuously strive to improve service outcomes and reduce the variability in the operating systems they oversee. From an operational perspective, these efforts are often inwardly directed. Employees are incented, processes are honed, and facilities are developed with an eye toward raising satisfaction levels, and in turn, firm performance. However, we suggest that an outward consideration – in particular, the fit between the needs of a firm's customers and the capabilities of its operation, which we term *customer compatibility* – may have even greater implications for the level and consistency of satisfaction a service operation is able to deliver, and its resulting financial performance.

In a manufacturing operation, inputs to the production process are sourced with compatibility in mind. Service operations differ fundamentally in that an essential input to the process – the customer – gets to choose which operation he or she engages. A wealth of theory and empirical evidence suggests that focused operations are more efficient and profitable (Skinner 1974, Tsikriktsis 2007). An extension of this idea is that customers whose needs are more aligned with the tradeoffs inherent in an operating model will impose less operational complexity and may, in turn, be served more effectively than customers whose needs are less aligned. Operations and industrial organization theory suggests that the process by which customers sort themselves among firms in a local market is rational and optimal, resulting in the pairing of each individual with the operating model best equipped to serve his or her needs (Cohen and Whang 1997, Sutton 1986, Tsay and Agarwal 2000). However, information asymmetries and the limited availability of local options may inhibit this process in practice.

In contexts where information is asymmetric – as is the case among service providers and prospective customers – equilibrium assumptions in signaling game models generally posit that firms will exhibit separating behaviors, investing in signals that accurately reveal their types to less-informed parties (Anand and Goyal 2009, Debo and Veeraraghavan 2010, Lariviere and Padmanabhan 1997). However, recent experimental research suggests that decision makers overwhelmingly opt for pooling behaviors in such settings – akin to the ill-fitting firm sending costly signals to prospective customers that its offerings are more aligned with their needs than they actually are (Schmidt and Buell 2017). The presentation of accurate operating details to prospective customers in practice may be constrained by the fact that most service organizations separate the groups responsible for acquiring and serving customers (Fitzsimmons and Fitzsimmons 2006), and these groups may have different incentives regarding customer compatibility. In messaging, managers tasked with acquiring customers may be strategic and selective about the information they choose to present, and behavioral research suggests that customers may interpret omitted details in biased and overly favorable ways (Kivetz and Simonson 2000). Likewise, recent research demonstrates how online reviews that convey the sample means of past experiences have a limited capacity to facilitate social learning, leading to biased overestimates of a service's latent quality (Besbes and Scarsini 2018), which could further undermine the decision making fidelity of prospective customers. Consistent with these dynamics, empirical work has characterized service as an experience good, which customers are unable to fully assess until after it has been delivered to them (Israel 2005). Taken together, these dynamics suggest that customers may lack the information required to reliably identify the service provider that is best equipped to meet their needs.

Even if customers were able to identify the best-fitting firms for themselves, the limited availability of local options is another force that may further inhibit the optimality of their matches with the providers they select. In many service settings such as retail banking, customers tend to select providers on the basis of geographic convenience (Devlin and Gerrard 2004, Martenson 1985), and not every provider competes in every market. Complicating matters, varying competitive dynamics can lead firms to attract customers with different preferences for service in different markets (Buell et al. 2016), and owing to the complexities and costs of managing multi-market operations, firms have a limited ability to customize service to suit each market's distinctive preferences (Berger et al. 2007, Erel 2009). As such, the same firm can attract customers with varying needs and preferences from market to market, and its operating system may be more compatible with the needs of some customers than it is with the needs of others.

Managers may experience less compatible customers as sources of variability that must either be reduced or accommodated (Frei 2006). Such customers may impose arrival, request, capability, effort, and preference variability, which simultaneously influences the company's capacity to deliver service efficiently, as well as the customer's evaluation of the experience (Campbell and Frei 2010, Chase 1978, Frei 2006, Ho and Zheng 2004, Karmarkar and Pitbladdo 1995, Sampson and Froehle 2006), both of which may have bottom line implications for the firm. To accommodate this variability, some firms engage in segmentation strategies, tailoring service on the basis of observed differences in customer needs and preferences (Guajardo and Cohen 2018). Although it is well understood that customer-imposed variability has the capacity to influence the performance of service operations, no work to date has examined the nature of this variation, quantifying the degree to which it may be a persistent and predictable phenomenon, stemming from differences in customer compatibility. In five empirical studies, we conduct such an analysis and make three contributions to the operations management literature.

First, we decompose the variance of 58,294 face-to-face retail-banking transactions, quantifying the relative significance of market, location, process, employee, and customer-level factors in explaining differences in visit satisfaction. In our models, which account for roughly a quarter of the aggregate variance in customer satisfaction, customer-level differences account for the vast majority of the explainable variance (95.9% - 96.8% of the explainable variance, 21.9% - 25.1% of the total variance, depending on the model). Differences among employees, processes, branches and markets account for the remainder. Further analysis reveals that markets, locations, processes, and employees – elements of the operating system that are traditionally considered to be within the manager's control – exhibit limited between-group variance and considerable within-group variance. For example, each employee may deliver relatively similar levels of satisfaction on average, but individual employees provide highly variable satisfaction from one transaction to the next. In contrast, customers exhibit relatively high between-group variance and relatively low within-group variance. Average levels of satisfaction differ markedly from one customer to the next, but individual customers tend to report similar satisfaction from transaction to transaction. In a complementary analysis of more than 3.7 million customer interactions with quick service restaurants (QSRs), presented in the online appendix, we show how this pattern generalizes beyond retail banking. The results reveal that although individual customers experience a range of satisfaction in their interactions with different brands, their satisfaction interacting with particular brands, and especially with particular brand locations, tends to be habitual. These converging results highlight how differences among customers serve as a critical driver in determining the outcomes of service interactions.

Second, we provide evidence that these customer-level differences are explained in part by customer compatibility – the degree of fit between the needs of individual customers and the capabilities of the operations serving them. We begin by introducing an empirical methodology that can be used to quantify customer compatibility. Next, we leverage the methodology using different publicly-available and proprietary data sources to test the effects of customer compatibility on customer experiences. Analyzing survey data collected by J.D. Power from 145,761 customers interacting with 164 retail banking institutions during a five-year period, we find that customers whose demographic characteristics, and in turn needs, diverge from the needs of their bank's average customer are less satisfied with the service they receive on a broad array of operating dimensions. A parallel analysis that uses the portfolio of products chosen by 49,582 customers of a single banking institution to measure customer divergence, provides converging, causal evidence that customer incompatibility drives dissatisfaction. By extension, we show that firms serving more heterogeneous customer bases – whose demographic characteristics, and in turn needs, are more dispersed – have customers who are significantly less satisfied overall. We find the negative effects of customer divergence are most pronounced for firms with customer bases whose needs are less dispersed – suggesting that focused service providers are most vulnerable to the negative effects of customer incompatibility.

Third and finally, we provide evidence that customer compatibility has a substantive effect on a firm's financial performance. A longitudinal analysis of branch-level deposit growth over a twelve-year period reveals that controlling for a host of factors, branches with more divergent (less compatible) customers exhibit slower deposit growth than other branches of the same institution. Against the backdrop of an average annual deposit growth rate of 6.7%, increasing a branch's customer divergence by one standard deviation resulted in a 1.2% decrease in annual deposit growth, relative to other branches of the same firm, a 17.9% reduction from baseline rates. We further find that the branches of firms with more dispersed customer bases experience considerably slower branch-level deposit growth over time. In our analysis, increasing customer dispersion from that of a first percentile (very focused) firm to that of the median (moderately focused) firm resulted in a 13.9% decrease in annual branch-level deposits, a 72.7% reduction from baseline rates. Subsequent analyses, presented in the online appendix, show that the links between customer compatibility and financial performance extend to the growth of these institutions' loan portfolios – demonstrating the broad impact of customer compatibility on a firm's growth prospects.

Taken together, these results highlight the important role that customer compatibility plays in determining customer experiences, and in turn, financial performance, in service organizations. Although a considerable body of research has documented factors that influence service outcomes for the customer and organization that are under the direct control of operating managers, our research highlights the complementary impact of factors that are not under managers' direct control – in particular, the degree of compatibility between the capabilities of an operation and the needs of the customers who choose to transact with it. This is an idea that has not been well explored in our literature, but it is foundational to the design and management of service operations.

2. Presentation of Studies

In four empirical studies, we investigate the effects of customer compatibility on customer experiences and firm performance. We begin by decomposing the variance of customer satisfaction that is attributable to market, branch, employee, process, and customer-level differences (Study 1). For our focal firm, we find that differences between customers account for the vast majority of the explainable variance in transaction satisfaction, and that although customers report quite different levels of satisfaction among themselves, individual customers tend to experience the firm consistently from transaction to transaction. In Study 2, we provide evidence that these satisfaction differences are explained in part by differences in customer compatibility - the degree of fit between the needs of individual customers and the capabilities of the operations serving them. We begin by introducing a methodology that can be used to quantify customer compatibility, measuring the degree of divergence between proxies for individual customer's needs – such as the customer's demographic characteristics, or the portfolio of products an individual customer has chosen – and the same proxies for the needs of the customers most typically served by the firm. We then leverage this methodology to test the impact of customer compatibility on service performance. We find that less compatible customers, measured by demographic divergence, report substantively lower levels of satisfaction, and that firms serving customer bases with more dispersed customer needs earn lower customer satisfaction scores. A parallel analysis provides converging evidence, using a product-based divergence measure, and demonstrates causally that incompatibility is a driver of dissatisfaction. Finally, Study 3 provides longitudinal evidence over a twelve-year period that differences in customer compatibility affect a firm's financial performance – companies serving less-compatible customers experience slower financial growth, measured both in terms of deposits and loans, over time.

2.1 Study 1: Customer Satisfaction Variance Decomposition

In this first study, we conduct a variance decomposition analysis to quantify the percentage of explainable variation in customer satisfaction that can be attributed to the market, branch, employee, process, and customer engaged in a retail banking transaction. There are several reasons that retail banking is an attractive context for this analysis. First, retail banking customers exhibit repeated interactions with the firm over time, often engaging with different employees in a variety of transaction types. Repeated interactions are a necessary feature for the variance decomposition methodology we employ. Second, nationwide retail banks, including the one we analyze, standardize service procedures and offer customers a discrete set of services, facilitating comparability of transactions across markets over time. Finally, given that the "products" offered by retail banks are largely commoditized, but long-term performance is predominantly driven through customer retention, considerable managerial attention is devoted to optimizing service interactions. In addition to enhancing the richness of service quality data collected by banks and third parties, this feature leads to operating systems that are highly streamlined. In this sense, examining a retail bank provides a window into a service system that incorporates many state-of-the-field innovations. As such, remaining sources of variance that emerge in our investigation can be characterized as opportunities for further investigation, rather than artifacts of an antiquated operating system.

2.1.1 Data and empirical approach. Between December 2003 and March 2011, our partner bank surveyed customers following 1,825,064 randomly-selected face-to-face service interactions. The day following each randomly-selected service transaction, customers were contacted by phone by a third-party survey firm and were asked to answer a series of questions that measured their visit satisfaction ("Please rate your overall satisfaction with your most recent visit to the … branch.") and their overall satisfaction with the firm ("Taking into account all the products and services you receive from [it], how satisfied are you with [the bank] overall?") on a scale of 1 (very dissatisfied) to 5 (very satisfied). Customers were also asked to estimate their perceived wait ("How many minutes would you say you had to wait for teller service during your most recent visit to the … branch?"). We matched this survey data with information from the bank's transaction databases to identify the employee, processes, location, and market involved in each transaction. Employees, locations, and markets were assigned unique identifiers by the focal institution. The bank's strategy group delineated market boundaries as a block of zip codes within which customers tend to transact. We note that each market is geographically isolated (Buell et al. 2016, Olivares and Cachon 2009), which facilitates our empirical approach.

Processes were classified by the focal institution into 12 categories: balance inquiries, bank product purchases, cashing bonds, cashing outside checks, cashing internal checks, owners cashing checks, deposits, miscellaneous credits, miscellaneous fees, payments, purchasing cash instruments, purchasing wire transfers, verifying funds, and withdrawals. For each customer-employee interaction, we capture the number, duration, and type of each process used. We characterize transactions by the complete set of processes conducted in service of the customer. For example, separate transactions where customers deposited a check and made a balance inquiry would be considered to have used identical processes. A subsequent customer who only deposited a check would be characterized as engaging in a different process.

As was typical of banks during the window of our analysis, the focal institution had attracted new customers both through its own marketing efforts, and by acquiring competing banks, whose customers had not explicitly chosen to transact with the focal institution. Since acquired customers may be less compatible than non-acquired customers, for conservatism, we only retained observations from non-acquired customers – those who had chosen to transact with the institution. Moreover, due to the data retention policies of our focal institution, transaction-level data were only available back to January 2009. Since variance decomposition analyses rely on repeated observations, we only retained observations from customers for whom there were at least two observations with complete transaction data (since January 2009) and at least one additional preceding observation with a survey response (since January 2003). This selection procedure resulted in a total of 58,294 transaction-level observations, representing 679 markets, 3,536 locations, 27,112 employees, 133 processes, and 27,247 customers. Summary statistics for these observations are provided in the online appendix. Importantly, each observed visit was distinct, separated by an average of 387.23 days (*SD* = 207.18 days). The presence of these time buffers mitigates the likelihood that facets of one observed interaction directly influenced the outcome of a subsequent one.

Although variance decomposition methodologies have been leveraged to address management questions outside of operations, most notably in the industrial organization and strategy literatures (McGahan and Porter 1997, Rumelt 1991, Schmalensee 1985), they have never been applied to analyze how dimensions of a firm's operating system affect customer outcomes. There are two factors that complicate such an analysis. First, in order to successfully decompose numerous dimensions of variance, repeat observations are required at every level of the analysis. To our knowledge, all of the current empirical work on customer satisfaction relies on data that is either cross-sectional, or lacks a sufficient degree of dimensionality to address all levels of a firm's operating system. Second, the nature of customer interactions with firms diverges from the requirements of optimization-based approaches for variance decomposition. In order to efficiently execute the optimization required to calculate maximum likelihood estimates using Analysis of Variance (ANOVA) and Components of Variance (COV) approaches, each level of data must be hierarchically nested. However, customer-firm interactions do not adhere to such a structure. Although employees typically work at particular locations and particular locations are statically situated within specific markets, process-level factors are identical across markets and, generally speaking, each market is capable of implementing all processes. Furthermore, a customer can engage with different employees, using various processes, in multiple locations across numerous markets. Using frequentist approaches to estimate the maximum likelihood function across each dimension on a sufficiently large dataset with such a structure would either be intractable or reach a terminal solution at an untenably slow pace (Jackman 2000). We overcome this challenge by using a Markov chain Monte Carlo multivariate generalized linear mixed models (MCMCglmm) approach (Hadfield 2010). By using sampling rather than optimization, this approach enables the calculation of such models with relative ease.

Our analysis relies on variations of the following model, estimated with a Gaussian functional form over 1,000,000 iterations, with a burn-in phase of 500,000 to reduce dependence on the pre-convergence period. Convergence traces for the base model are provided in the online appendix:

$$Y_{ieplmt} = \mu + A_i + B_e + C_p + D_l + E_m + \alpha WAIT_{it} + \beta SAT_{it-1} + \epsilon_{ieplmt} \quad (1)$$

Y_{ieplmt} represents the satisfaction reported by customer *i* after interacting with employee *e*, to engage in process *p*, in location *l*, in market *m*, at visit time *t*. Visit satisfaction is measured as an ordinal variable, ranging from 1 (extremely low) to 5 (extremely high). μ represents the average visit satisfaction across all visit observations, A_i represents customer-specific random effects, B_e represents employee-specific random effects, C_p represents process-specific random effects, D_l represents location-specific random effects, and E_m represents market-specific random effects. ϵ_{ieplmt} characterizes the residual. In an MCMCglmm variance decomposition analysis, random effects are used to specify the variance components that are being estimated, and fixed effects are used to introduce control variables into the estimation.

Owing to the fact that visit satisfaction is likely to be influenced by the length of time customers have to wait to conduct the transaction (Taylor 1994), in some specifications, we introduce a fixed effects variable, $\alpha WAIT_{it}$, capturing the length of time the customer reported waiting for service (in minutes). Furthermore, customer perceptions of service outcomes are likely to be influenced by latent feelings about the firm and their relationship with it. Hence, a helpful control for evaluating the degree to which various elements of a firm's operating system account for differences in a customer's satisfaction with a particular visit is the ex-ante level of satisfaction the customer brought to the interaction. Accordingly, in some specifications, we also introduce a fixed effect, βSAT_{it-1}, which controls for the customer's overall level of satisfaction as reported in their previous observed interaction with the firm.

2.1.2 Drivers of customer satisfaction differences. As shown in **Table 1**, customer-specific random-effects account for a substantial relative percentage of the explainable variation in visit satisfaction. In

Column (1), customer, employee, process, location, and market effects account for 96.0% (25.1%), 0.0% (0.0%), 0.4% (0.1%), 2.7% (0.7%), and 0.8% (0.2%), respectively, of the explainable (total) variance in visit satisfaction. The fixed effect controls introduced in Columns (2-4) make relatively small differences. In the fully-specified model, shown in Column (4), we find that customer, employee, process, location and market effects account for 96.8% (21.9%), 0.6% (0.1%), 0.7% (0.2%), 0.0% (0.0%), and 2.0% (0.4%), respectively, of the explainable (total) variance in visit satisfaction.

	(1) Satisfaction	(2) Satisfaction	(3) Satisfaction	(4) Satisfaction
% of total variance				
Customer	25.1%	24.3%	22.6%	21.9%
Employee	0.0%	0.0%	0.0%	0.1%
Process	0.1%	0.2%	0.1%	0.2%
Location (branch)	0.7%	0.3%	0.8%	0.0%
Market	0.2%	0.4%	0.0%	0.4%
Error	73.9%	74.8%	76.5%	77.3%
% of explainable variance				
Customer	96.0%	96.5%	95.9%	96.8%
Employee	0.0%	0.1%	0.1%	0.6%
Process	0.4%	0.7%	0.4%	0.7%
Location (branch)	2.7%	1.2%	3.5%	0.0%
Market	0.8%	1.5%	0.2%	2.0%
Fixed Effects	None	Wait Time	Lagged Overall Sat.	Wait Time, Lagged Overall Sat.

Table 1: Variance components of visit satisfaction, n=58,294 (Study 1). These results reflect sources of variability calculated from estimates of Columns (1-4), with 1,000,000 iterations, a burn-in phase of 500,000 iterations, and non-informative priors. Means and standard deviations for each estimated variance component are provided in the online appendix.

Although it is not surprising that customer differences should account for some percentage of the aggregate variance in customer satisfaction, the magnitude of the effect across models is worthy of consideration. We suggest several possibilities. First, it is important to note that roughly three-quarters of the total variance in visit satisfaction is left unexplained by our model. Service interactions and satisfaction are not fully systematic and predictable – an unexpected employee mistake, a random occurrence that puts a customer in an uncharacteristically bad mood, or an episodic process failure, all of which can wreak havoc on an individual's service experience, would reside in the error term. Similarly, interactions between and within the included dimensions – for example, specific employee/customer pairs resulting in particular delight or despair (Schneider et al. 2013) – would not be captured in our model. Due to a limited number of repeat observations at the customer and employee levels, we lacked sufficient power to include such random effect interactions in our analysis. Second, although we included a lagged overall satisfaction fixed-effect variable in Columns (3-4), a stronger control variable representing the customer's expectations upon commencing the interaction would likely mitigate the relative dominance of customer-specific effects. Our control leverages the last observed measure of overall satisfaction, but in most cases, numerous unobserved interactions took place between the focal transaction and the last observation of overall satisfaction. Third, retail banking processes are quite standardized, and low variation attributed to firm-controlled dimensions may indicate service consistency brought about by strong process control. In organizations with more highly variable operating systems, we would likely expect to see a larger percentage of the variance loading on employees, processes, locations and markets. Nevertheless, the results reveal that differences between customers play an important role in explaining differences in service outcomes.

	Between-group variance	Within-group variance	Total obs. with n>1	Number of groups	Mean obs./group
Customer	0.242	0.157	58,294	27,247	2.14
Employee	0.102	0.415	336,446	41,834	8.04
Process	0.094	0.480	343,056	179	1,916.51
Location (branch)	0.016	0.471	343,065	3,700	92.72
Market	0.011	0.478	343,093	712	481.87

Table 2: Between and within-group variance in visit satisfaction (Study 1). Between and within-group variance above was calculated using the full dataset of customer responses for which each group of interest, for every level, had more than one observation, and the customer, employee, process, location, and market were identified.

2.1.3: Between and within-group variance in visit satisfaction. In **Table 2**, we examine the between and within-group variance of visit satisfaction for customers, employees, processes, locations and markets. Since calculating between and within-group variance is conducted one dimension at a time, we began with the initial dataset of 1,825,064 randomly-selected face-to-face service interactions, and for each dimension, analyzed the full set of observations with an identified customer, employee, process, location, and market, as well as at least one repetition in the data. The results reveal that elements of the firm's operating system that are traditionally considered to be under the firm's control exhibit relatively low between-group variance, σ_b^2, and relatively high within-group variance, σ_w^2: employees (σ_b^2 = 0.102; σ_w^2 = 0.415), processes (σ_b^2 = 0.094; σ_w^2 = 0.480), locations (σ_b^2 = 0.016; σ_w^2 = 0.471), and markets (σ_b^2 = 0.011; σ_w^2 = 0.478). For example, although comparing the aggregate satisfaction produced by two employees may yield very consistent results, comparing the satisfaction produced by an individual employee from transaction to transaction would result in significant variation. In contrast, customers exhibit relatively high between-group variance (σ_b^2 = 0.242), the highest across all groups in our analysis, and relatively low within-group variance (σ_w^2 = 0.157), the lowest across all groups in our analysis.

On average, this suggests that although comparing the aggregate satisfaction experienced by two customers may result in discrepancies, tracking the satisfaction of an individual customer from transaction to transaction would yield relative consistency. The pattern of results is consistent with the idea that visit satisfaction differences are less a function of the employee, process, location, or market involved in the delivery of the transaction than they are a function of the customer who walked in the door. A complementary analysis, performed with data from 3.7 million customer-level QSR transactions and presented in the online appendix, provides converging evidence in a different domain. Individual customers exhibit considerable variation in satisfaction across their many interactions with different QSR brands, but they exhibit relatively consistent levels of satisfaction when they interact with particular brands, and extremely consistent levels of satisfaction when they interact with specific locations of particular brands. These results suggest that satisfaction differences within a firm emanate in large part from differences among customers. They also suggest that the satisfaction experienced by individual customers across transactions tends to be habitual in nature, such that some customers are routinely more satisfied with the service provided by the firm than others.

2.2 Study 2: Customer Compatibility and Service Satisfaction

One possible explanation for the customer-level differences observed in the first study is that the firm's operating system may be aligned to serve the needs of some customers better than others. If this were the case, then to the extent that the capabilities of a firm's service operations converge over time with the needs demanded by the customers it most typically serves, we might expect that customers who diverge markedly from the firm's core customer base may derive less satisfaction from its operating system. In Study 2, we test this possibility by analyzing the extent to which satisfaction (with a firm's product offerings, operating processes, people, and interaction design) varies among customers whose demographics (e.g., population density, per capita income, median age, average household size, proportion of the population between 18-34, and proportion of the population that owned their home) suggest their needs may be more or less compatible with those of a firm's most typical customers. Furthermore, to the extent that customer compatibility is associated with differences in service outcomes, we would further expect that firms serving customer bases whose needs are more dispersed would have customers who are less satisfied on these dimensions than firms serving customer bases whose needs are less dispersed. In Study 2, therefore, we additionally test this possibility by analyzing how the degree of customer dispersion in the markets served by a firm is predictive of customer satisfaction with various facets of its operation. We also investigate the interdependence of these effects.

2.2.1 Data and empirical approach. We conduct this study using data from the 2007-2011 J.D. Power Retail Banking Satisfaction Surveys, which during the time of our analysis captured customer experience evaluations from 149,389 customers of 166 retail banking institutions. During the fourth quarters of 2006-2010, J.D. Power surveyed 16,646, 16,654, 27,833, 42,279, and 45,977 customers, respectively, asking each a battery of questions regarding their satisfaction with various facets of their experience interacting with their retail bank. In addition to capturing their overall level of satisfaction, customers were asked to evaluate the firm's product offerings (e.g., variety of banking services available, innovation of new services offered, competitiveness of interest rates, range of services that can be performed by tellers, with an ATM, online, through the IVR, or with a telephone agent), operating processes (e.g., ease of opening an account, ease of making changes to an account, effectiveness of communications about accounts, timeliness, usefulness, and ease of understanding of account statements, and amount of time spent waiting for service transactions), people (e.g., courtesy, friendliness, and knowledge of phone and in-person agents), and interaction design (e.g., hours, number of branches, ease of accessing branches, number and locations of ATMs, design of statements, and clarity of online instructions and IVR menu prompts). Mean responses are aggregated across respondents for each dimension. The survey also asked each respondent to provide his or her overall level of satisfaction with the bank. We analyze the effects of customer compatibility on customer evaluations of their experiences with each facet of the bank's operation.

As a proxy for each respondent's needs, we use zip code-level demographic data from the 2007-2011 American Community Survey (ACS). In particular, for each respondent's zip code, we capture the population density, per capita income, median age, and average household size. These particular demographic characteristics have been identified in the marketing literature as the most important drivers of customer demand in chain organizations (Gupta and Chintagunta 1994, Hoch et al. 1995, Kalyanam and Putler 1997, Mulhern and Williams 1994). Owing to the fact that we are conducting this study in the retail banking industry, we further capture data on the proportion of the population aged 18-34, the age range of people most likely to use direct banking channels due to their openness to new technology (Cortiñas et al. 2010), and the proportion of the population that own their own homes, which serves as a proxy for the complexity of their financial needs and has been a fixture of empirical investigations of the drivers of customer satisfaction in banking (Levesque and McDougall 1996).

One potential concern with this empirical strategy is whether the inferences being drawn about the needs of customers under analysis are accurate, given that their identification is subject to two layers of abstraction: 1) the average demographics of a zip code may not map precisely to the demographic characteristics of an individual who lives within that zip code, and 2) although demographic characteristics have been shown to correlate with customer needs, they are admittedly an imperfect proxy. Therefore, as a pilot test of the idea that zip-code level demographics serve as a sufficiently high-fidelity proxy for customer needs to facilitate our analysis, we surveyed 497 participants (42.5% female, M_{age}=35.61, *SD*=10.74) on the Amazon Mechanical Turk platform, who consented to participate in a 5-minute survey about their relationship with their primary banking institution in exchange for $1.00 (Buhrmester et al. 2011; Mason and Suri 2012). We asked each participant a battery of questions about the channels through which they interacted with their bank (Buell et al. 2010), the products they held with their bank, and the relative importance of different drivers of satisfaction with their bank (Baumann et al. 2007, Levesque and McDougall 1996). We also asked them to identify the zip code in which they lived, their household income, age, household size, and whether they owned their home. Results presented in the online appendix reveal that 1) zip code level demographic data from the ACS are predictive of customer-level demographics in the corresponding zip code, and that 2) customer and zip code-level demographics are predictive of customer-level differences in channel usage, product usage, and drivers of banking satisfaction. This pattern of results lends support to our empirical approach.

To approximate the capabilities of each bank's operation, we use data from the 2006-2010 Federal Deposit Insurance Corporation (FDIC) Summary of Deposits database, which provides a snapshot as of June 30 each year of the location of every bank branch for every licensed retail bank in the United States. Merging these data with the demographic information in the American Community Survey enables us to calculate for each bank, the average demographic characteristics of people living in the zip codes where the bank has branches. To the extent that the capabilities of an operation are reflective of the needs of the customers it serves, we might expect the average demographics across a bank's branch network to be a reasonable proxy for the targeted capabilities of its operation. This identification strategy hinges on four well-supported assumptions: 1) bank branches tend to attract customers who live and work in the surrounding area (Devlin and Gerrard 2004), 2) population characteristics vary across regions, 3) customer needs and relevant service quality dimensions in banking vary by demographic (Stafford 1996), and 4) banks standardize their operating models across markets (Berger et al. 2007, Erel 2009).

Having proxies for the needs of each customer and the capabilities of each operation enables us to calculate the degree of divergence for each respondent to the J.D. Power Retail Banking Satisfaction Study. Consistent with our prior results, to the extent that customer compatibility – the degree of fit between the needs of customers and the capabilities of an operation – influences customer satisfaction, we hypothesize that customers who exhibit a greater degree of divergence from their bank's average customer will report lower levels of satisfaction with their bank.

Hypothesis 1 (H1): Divergence between the capabilities of an operation and the needs of a customer is negatively associated with customer satisfaction.

Access to these data also enables us to investigate a closely related question: how the heterogeneity of needs represented by a firm's customer base influences the overall level of satisfaction its operation can deliver. Leveraging the data described above, we are able to measure how much demographic dispersion exists across the markets where each bank has branches. Consistent with the theory that customer compatibility is associated with service outcomes, we hypothesize that banks with customer bases whose needs are more broadly dispersed will offer services that are less compatible on average with the distinct needs of each of its customers, leading to lower levels of overall satisfaction. However, since in the limit, maximal dispersion should result in an operating model that is optimized to meet the needs of an "average" customer across all markets (e.g., a firm that is endeavoring to be all things to all people), we predict that the negative effects of dispersion on satisfaction will decrease at a diminishing rate.

Hypothesis 2 (H2): Dispersion among the needs of customers served by an operation is negatively associated with the level of satisfaction that operation can deliver, but at a diminishing rate.

Finally, although we hypothesize that customer divergence and customer dispersion separately influence the degree of compatibility among firms and customers, and in turn, the satisfaction that results from their interaction, customer divergence and customer dispersion emanate from different sources. Customer divergence is largely a function of customer choices (for example, the firm with which a customer chooses to transact), while customer dispersion is largely a function of the firm's strategic choices (for example, the markets the firm chooses to enter). Furthermore, it seems unlikely that the effects of customer divergence and customer dispersion are independent of one another. In particular, a highly-focused operation with low customer dispersion is likely to deliver experiences that result in sharply lower customer evaluations when it serves a divergent customer – a customer who is a poor fit for a highly-focused firm will be significantly less satisfied than a highly compatible customer of the same firm. In contrast, an operation that endeavors to be all things to all people, with high customer dispersion, should not exhibit sharply lower satisfaction when it serves a highly divergent customer. If a firm is optimized around the needs of a highly dispersed customer base, then every customer is more divergent on average, and the declines in satisfaction that arise from customer divergence should not be so large. Consistently, we hypothesize:

Hypothesis 3 (H3): Customer dispersion positively moderates the negative association between customer divergence and customer satisfaction.

We next describe how we define and calculate customer divergence and customer dispersion for our analyses. Summary statistics for each metric are provided in the online appendix. Owing to our reliance on the locations of physical branches to calculate these metrics, we dropped data from two banks, which during the time of our analysis competed exclusively online. We additionally dropped the observations of customers living in zip codes for which a full panel of demographic data from the ACS was not available. Hence, we perform our tests of H1-H3 on data from the remaining 164 banks, which comprised 145,761 customer-level observations.

Customer divergence is intended to describe the gap between the needs of a particular customer and the capabilities of the operation serving them. To operationalize such a metric for each customer living in zip code *i* who completed the J.D. Power Retail Banking Satisfaction survey about their interactions with firm *j,* during year *t*, we began by calculating a normalized divergence metric, $ND_{c_{ijt}}$, for each of the six focal demographic characteristics described above. For each demographic characteristic, we captured the difference between the demographic characteristic's mean value in the customer's home zip code, c_{it}, and the mean value across all zip codes where the firm had branches that year, $\bar{c}_{jt}$. We normalized the absolute value of each difference by dividing it by the standard deviation of the demographic characteristic across all zip codes in which the firm had branches that year.

$$ND_{c_{ijt}} = \frac{\left|c_{it} - \bar{c}_{jt}\right|}{\text{Standard Deviation}_{c_{jt}}} \tag{2}$$

The resulting normalized divergence metrics, therefore, behave akin to z-scores, quantifying the degrees to which each customers' zip codes exhibited demographic characteristics (and, in turn, needs) that differed from those of the customers most typically served by their banks. Consistent with prior work (Campbell et al. 2009), we summed these six normalized divergence metrics to create an aggregate measure of customer divergence, DIV_{iit}, for each survey respondent:

$$DIV_{ijt} = \sum_{c=1}^{6} ND_{c_{ijt}} \tag{3}$$

The resulting aggregate measure serves as a proxy for the degree of compatibility between the needs of each customer, and the capabilities of the operating system that serves them, with higher levels of divergence representing lower levels of compatibility.

Customer dispersion describes the degree to which a firm serves a customer base with heterogeneous needs. Some companies provide service in markets where customer needs are relatively homogenous, while others serve markets where needs are more dispersed. To capture this variation, we created a normalized dispersion metric for each firm, *j*, during each year *t.* We began by calculating coefficients of variation for each of the six previously described demographic characteristics across all zip codes in which each firm competed during each year:

$$CV_{c_{jt}} = \frac{\text{Standard Deviation}_{c_{jt}}}{\text{Mean}_{c_{jt}}} \tag{4}$$

The coefficient of variation for each demographic characteristic for a particular bank measures the degree of dispersion of that demographic characteristic across the bank's entire retail branch network. Consistent with prior literature, we summed the six coefficients of variation to create an aggregated dispersion metric for each firm, capturing the degree of customer heterogeneity each faced across all the markets it served in a given year (Campbell et al. 2009):

$$DIS_{jt} = \sum_{c=1}^{6} CV_{c_{jt}} \tag{5}$$

Institutions facing a higher degree of customer dispersion have customers whose demographic characteristics, and in turn, needs, are less compatible with one another – complicating the task of delivering satisfying service experiences to every customer. Although we rely on these additive models as our primary metrics of normalized divergence and dispersion throughout this paper, we test and implement additional aggregation strategies in the online appendix.

2.2.2 Customer divergence and service satisfaction (H1). We test the effects of customer divergence on service satisfaction by modeling the self-reported satisfaction levels of 145,761 customers with various dimensions of the operating system (e.g., product offerings, operating processes, people, interaction design, and overall satisfaction), SAT_{ijt}, as a function of the level of divergence between the needs of each customer, and the capabilities of the institution serving them, DIV_{ijt}. To better isolate the effect of interest from omitted correlated factors that could lead to biased estimators, we control for a handful of institution, region, and time-based covariates.

Specifically, we control for the geographic distance between the centroid of the customer's zip code and the closest branch of their primary banking institution. Although the distribution of distance is highly skewed by customers who have long-distance relationships with their banks, the median distance in our dataset was 1.25 miles. $CLOSENESS_{ijt}$ is an indicator variable denoting whether the customer lives within this range. We note our results are substantively similar when a continuous distance measure is substituted. $COMP_{ijt}$ represents the number of separate institutions competing in the customer's zip code during the year of observation – a measure of time-varying customer choice. We additionally control for characteristics of each customer's primary banking institution, including counts of its branches, $BRCOUNT_{jt}$, and the natural logarithm of 1+ its annual deposits, $\ln DEP_{jt}$, both drawn from the FDIC Summary of Deposits database. We incorporate fixed effects to account for unobservable sources of endogeneity that could otherwise bias our results. To account for time-invariant aspects of each institution, such as its service model, employee training regimen, channel strategy, and branch network, we include a firm-level fixed effect. Since our analysis is conducted on data gathered over a five-year period, this fixed effect does not completely subsume controls for branch counts and deposits. We also include a zip-code level fixed effect, which captures time-invariant differences across zip codes in customer service expectations, and indicator variables denoting the year the survey response was collected, which captures region-invariant differences in customer attitudes toward banks over time. In the fullest specification, we use the following fixed-effects OLS model:

$$\begin{aligned} SAT_{ijt} &= \gamma_0 + \gamma_1 DIS_{ijt} + \gamma_2 CLOSENESS_{ijt} + \gamma_3 COMP_{ijt} \\ &+ \gamma_4 BRCOUNT_{jt} + \gamma_5 \ln DEP_{jt} + \alpha_i + \beta_j + \delta_t + \epsilon_{ijt} \end{aligned} \tag{6}$$

γ_1 is the primary coefficient of interest. If $\gamma_1 < 0$, then consistent with H1 and our theory of customer compatibility being an important determinant of service outcomes, divergence between the needs of a particular customer and the capabilities of the operation serving them is associated with diminished customer satisfaction.

Table 3 presents the results estimated with robust standard errors, with multi-layer clustering at the institution and zip code level, to account for potential correlation in the error terms among customers living in the same zip code, or among customers who select the same institution (Cameron et al. 2011). Such an estimation technique fails to produce a constant, so we estimate marginal effects and produce predictive plots and confidence intervals based on estimates with robust standard errors clustered at the zip code level. We present results with alternate clustering techniques, which are substantively similar, in the online appendix. Columns (1-3) demonstrate a robust negative relationship between customer divergence and overall satisfaction (γ=-0.011, p<0.05). Reducing customer divergence from the 99th to the 1st percentile corresponded with a 1.63% increase in overall satisfaction – equivalent to 7.0% of the standard deviation of overall satisfaction. Reducing customer divergence from the maximum to the minimum in our sample (the effect of changing from the least compatible to the most compatible customer) corresponded with a 23.2% increase in customer satisfaction – equivalent to 100% of the standard deviation of overall satisfaction. The magnitude and direction of focal effects is similar for customer evaluations of product offerings (γ=-0.008, p<0.05), operating processes (γ=-0.010, p<0.05), people (γ=-0.008, p<0.05), and interaction design (γ=-0.023, p<0.01), providing converging support for the hypothesis that customer divergence – an incompatibility between the needs of customers and the capabilities of an operation – adversely effects service outcomes.

	(1) Overall satisfaction	(2) Overall satisfaction	(3) Overall satisfaction	(4) Product offerings	(5) Operating processes	(6) People	(7) Interaction design
Customer divergence	-0.014***	-0.012**	-0.011**	-0.008**	-0.010**	-0.008**	-0.023***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.007)
Geographic closeness		0.065***	0.063***	0.028*	0.055***	0.061***	0.262***
		(0.014)	(0.014)	(0.014)	(0.012)	(0.014)	(0.014)
Banking instituions in zip		0.006	-0.001	0.000	-0.005	-0.000	-0.005
		(0.008)	(0.007)	(0.007)	(0.007)	(0.010)	(0.008)
Instituion branch count		0.000	-0.000	-0.000	0.000	0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)		0.003	0.013**	0.012*	0.009**	0.001	0.006
		(0.002)	(0.006)	(0.007)	(0.004)	(0.006)	(0.004)
2007 indicator			-0.062**	-0.182***	-0.113***	-0.046	-0.225***
			(0.026)	(0.025)	(0.027)	(0.028)	(0.032)
2008 indicator			0.050	-0.124***	-0.048	0.057	-0.152***
			(0.042)	(0.045)	(0.036)	(0.036)	(0.050)
2009 indicator			0.116*	0.176***	0.256***	0.233***	0.231***
			(0.066)	(0.066)	(0.050)	(0.054)	(0.050)
2010 indicator			0.198*	0.372***	0.391***	0.305***	0.302***
			(0.109)	(0.108)	(0.086)	(0.105)	(0.074)
Observations	141,429	141,429	141,429	141,015	141,359	120,835	141,360
Number of zip codes	13,039	13,039	13,039	13,031	13,041	12,378	13,042
Adjusted R-squared	0.083	0.082	0.082	0.085	0.081	0.092	0.067

Table 3: Customer Divergence and Service Satisfaction (Study 2). Consistent with H1, customers with higher levels of divergence exhibit lower satisfaction with service outcomes across a broad array of operating dimensions. All models include institution and zip code-level fixed effects. Robust standard errors, with multi-layer clustering by institution and zip code, are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

2.2.3 Customer dispersion and satisfaction (H2). We test the effects of customer dispersion on service satisfaction by modeling the self-reported satisfaction levels of the same respondents with various dimensions of the operating system (e.g., product offerings, operating processes, people, interaction design, and overall satisfaction), SAT_{ijt} , this time as a function of the level of dispersion across the firm's customer base, DIS_{ijt} . To account for the possibility that customer dispersion negatively impacts customer experiences, but at a diminishing rate, as hypothesized in H2, we additionally include a quadratic term, DIS_{ijt}^{2} . Owing to the relatively slow evolution of the composition of a bank's customer base (through the opening, closing, and acquisition of branches in different geographies), customer dispersion is largely a fixed characteristic of the firm over time. Hence, most of the customer dispersion differences among firms' customer bases would be subsumed by the inclusion of institution-level fixed effects. Consequently, we model the effects of customer dispersion on satisfaction using the following linear fixed effects model, with zip code-level fixed effects and year-based indicators, and standard errors clustered at the institution level:

$$SAT_{ijt} = \left(\delta_0 + \delta_1 DIS_{ijt} + \delta_2 DIS_{ijt}{}^2 + \delta_3 BRCOUNT_{jt} + \delta_4 \ln DEP_{jt} + \alpha_i + \gamma_t + \epsilon_{ijt}\right) \qquad (7)$$

δ_1 and δ_2 are the primary coefficients of interest. If $\delta_1 < 0$, then consistent with H2 and our theory of customer compatibility being an important determinant of service outcomes, dispersion among the needs of customers served by an operation is negatively associated with the level of satisfaction that operation can deliver. Moreover, if $\delta_2 > 0$, then consistent with H2, the dispersion of needs represented by a firm's customer base negatively affects customer satisfaction at a diminishing rate. Indeed, in **Table 4**, the fully-specified model in Column 3 reveals that consistent with H2 and our theory of customer compatibility being an important driver of service experiences, overall satisfaction falls in customer dispersion (δ=-0.408, p<0.05), but at a diminishing rate (δ=0.047, p<0.05). **Figure 1** plots these results graphically. The remaining columns demonstrate a consistent pattern of results that offers converging support for H2. As with overall satisfaction, increasing customer divergence reduces customer evaluations of operating processes, people, and interaction design, albeit at diminishing rates.

	(1) Overall satisfaction	(2) Overall satisfaction	(3) Overall satisfaction	(4) Product offerings	(5) Operating processes	(6) People	(7) Interaction design
Customer dispersion	-0.922***	-0.436***	-0.408**	-0.161	-0.257**	-0.405**	-0.396**
	(0.209)	(0.166)	(0.172)	(0.108)	(0.111)	(0.165)	(0.174)
Customer dispersion²	0.109***	0.051**	0.047**	0.020	0.031**	0.050**	0.052**
	(0.025)	(0.020)	(0.021)	(0.013)	(0.014)	(0.020)	(0.022)
Geographic closeness		0.096***	0.091***	0.034**	0.071***	0.082***	0.305***
		(0.019)	(0.018)	(0.014)	(0.013)	(0.019)	(0.014)
Banking instituions in zip		0.018**	0.003	0.004	-0.002	0.002	-0.003
		(0.009)	(0.008)	(0.007)	(0.008)	(0.010)	(0.008)
Instituion branch count		-0.000***	-0.000***	-0.000**	-0.000**	-0.000***	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)		-0.002	-0.006	0.001	-0.004	-0.014**	-0.002
		(0.002)	(0.010)	(0.007)	(0.006)	(0.007)	(0.006)
2007 indicator			-0.016	-0.160***	-0.080***	-0.014	-0.194***
			(0.029)	(0.024)	(0.026)	(0.028)	(0.032)
2008 indicator			0.107**	-0.100**	-0.010	0.095***	-0.116**
			(0.045)	(0.044)	(0.037)	(0.034)	(0.050)
2009 indicator			0.195***	0.186***	0.299***	0.285***	0.271***
			(0.056)	(0.051)	(0.043)	(0.046)	(0.043)
2010 indicator			0.048	0.247***	0.272***	0.168*	0.248***
			(0.114)	(0.086)	(0.081)	(0.090)	(0.076)
Observations	141,429	141,429	141,429	141,015	141,359	120,835	141,360
Number of zip codes	13,039	13,039	13,039	13,031	13,041	12,378	13,042
Adjusted R-squared	0.097	0.094	0.093	0.092	0.089	0.103	0.079

Table 4: Customer Dispersion and Service Satisfaction (Study 2). Consistent with H2, customers with higher levels of dispersion exhibit lower levels of overall satisfaction and lower levels of satisfaction with operating processes, though the effect diminishes in higher levels of dispersion. Robust standard errors, with multi-layer clustering by institution and zip code, are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

Interestingly, piecewise linear regression analyses, presented in the online appendix, reveal that despite the negative associations among dispersion and customer evaluations at most levels, within the top quartile of dispersion, the slopes on overall satisfaction and satisfaction with people are positive and significant. These results are consistent with the idea that among the least focused institutions, an *everything to everyone* strategy may, on some dimensions, outperform companies that are just highly unfocused. However, consistent with H2, focused firms, which have the least dispersed customer bases, consistently receive the most favorable evaluations from customers. Holding all else constant, increasing customer dispersion from the 1st percentile to the median is associated with a 2.2% decline in overall satisfaction, 9.7% of the standard deviation of customer satisfaction across firms. Owing to the non-linearity, increasing customer dispersion from the 1st percentile to the 99th percentile is associated with a similar 2.3% decline in overall satisfaction, equivalent to 9.9% of the standard deviation of customer satisfaction across firms.

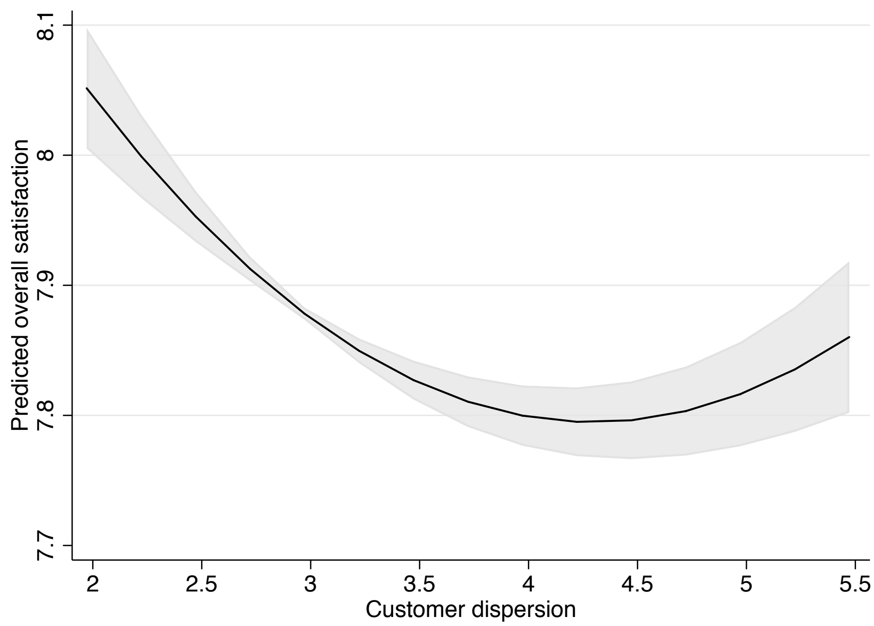

Figure 1: Customer dispersion and overall satisfaction (Study 2). Consistent with H2, customers with higher levels of dispersion exhibit lower satisfaction with service outcomes across a broad array of operating dimensions. Figure is plotted within 99% support of the data, with a 95% confidence interval band, estimated with robust standard errors clustered by zip code, shown in light grey.

2.2.4 Interdependent effects of customer divergence and customer dispersion on satisfaction (H3). The results presented in the preceding sections suggest that customer divergence and customer dispersion both have negative effects on how customers experience an operation. In this section, we test whether, as hypothesized in H3, the effect of customer divergence depends on the degree of customer dispersion. To test this proposition, we introduce an interaction between customer divergence and customer dispersion into Model (8), using the following fixed-effects specification, with standard errors clustered by institution:

$$SAT_{ijt} = \zeta_0 + \zeta_1 DIV_{ijt} + \zeta_2 DIV_{ijt} \times DIS_{jt} + \zeta_3 CLOSENESS_{ijt} + \zeta_4 COMP_{ijt} \\ + \zeta_5 BRCOUNT_{jt} + \zeta_6 \ln DEP_{jt} + \alpha_i + \beta_j + \gamma_t + \epsilon_{ijt} \tag{8}$$

ζ_1 represents the effects of customer divergence on highly focused firms (where there is no customer dispersion. If $\zeta_1 < 0$, then consistent with H1, highly-divergent customers of focused firms exhibit declines in satisfaction that are directionally consistent with the effects documented earlier across all customers. ζ_2 represents the incremental effect of customer divergence among customers of more dispersed firms. If $\zeta_2 > 0$, then consistent with H3, focused firms are more prone to the deleterious effects of customer divergence.

	(1) Overall satisfaction	(2) Overall satisfaction	(3) Overall satisfaction	(4) Product offerings	(5) Operating processes	(6) People	(7) Interaction design
Customer divergence	0.068**	-0.035***	-0.031***	-0.023**	-0.019**	-0.013	-0.031***
	(0.028)	(0.010)	(0.009)	(0.009)	(0.008)	(0.009)	(0.012)
Divergence x dispersion	-0.021**	0.009***	0.008***	0.006**	0.004*	0.002	0.003
	(0.009)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Geographic closeness		0.065***	0.063***	0.028*	0.055***	0.061***	0.262***
		(0.014)	(0.014)	(0.014)	(0.012)	(0.014)	(0.014)
Banking instituions in zip		0.006	-0.001	0.001	-0.005	-0.000	-0.005
		(0.008)	(0.008)	(0.007)	(0.007)	(0.010)	(0.008)
Instituion branch count		0.000	-0.000	-0.000	0.000	0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)		0.003	0.012**	0.012*	0.009**	0.001	0.006
		(0.002)	(0.006)	(0.007)	(0.004)	(0.006)	(0.004)
2007 indicator			-0.063**	-0.183***	-0.114***	-0.046	-0.226***
			(0.027)	(0.025)	(0.027)	(0.028)	(0.032)
2008 indicator			0.049	-0.125***	-0.049	0.057	-0.153***
			(0.042)	(0.045)	(0.036)	(0.036)	(0.050)
2009 indicator			0.114*	0.175***	0.255***	0.233***	0.230***
			(0.066)	(0.066)	(0.050)	(0.054)	(0.051)
2010 indicator			0.189*	0.366***	0.387***	0.303***	0.298***
			(0.109)	(0.107)	(0.086)	(0.105)	(0.074)
Observations	141,429	141,429	141,429	141,015	141,359	120,835	141,360
Number of zip codes	13,039	13,039	13,039	13,031	13,041	12,378	13,042
Adjusted R-squared	0.100	0.082	0.082	0.085	0.081	0.092	0.067

Table 5: The interdependent effects of Customer Divergence and Customer Dispersion on Service Satisfaction (Study 2). Consistent with H3, customer dispersion positively moderates the relationship between customer divergence and service satisfaction on most dimensions. All models include zip code and institution-level fixed effects. Robust standard errors, with multi-layer clustering by institution and zip code are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

Indeed, **Table 5** suggests that focused firms are more prone to the negative effects of customer divergence than non-focused firms on most dimensions. In the fully specified models, customer divergence among highly focused firms has a negative effect on overall satisfaction (ζ=-0.031, p<0.01), satisfaction with product offerings (ζ=-0.023, p<0.05), satisfaction with operating processes (ζ=-0.019, p<0.05), and satisfaction with interaction design (ζ=-0.031, p<0.01). Interestingly, and consistent with H3, the results further indicate that the differential effect of customer divergence varies between highly-focused firms (with zero dispersion) and less-focused firms (with non-zero dispersion). Divergent customers of more dispersed (less focused) firms are more satisfied overall (ζ=0.008, $p<0.01$), and are more satisfied with the less focused firm's product offerings (ζ=-0.006, $p<0.05$) and operating processes (ζ=-0.004, $p<0.10$) than are equivalently divergent customers of highly focused firms.

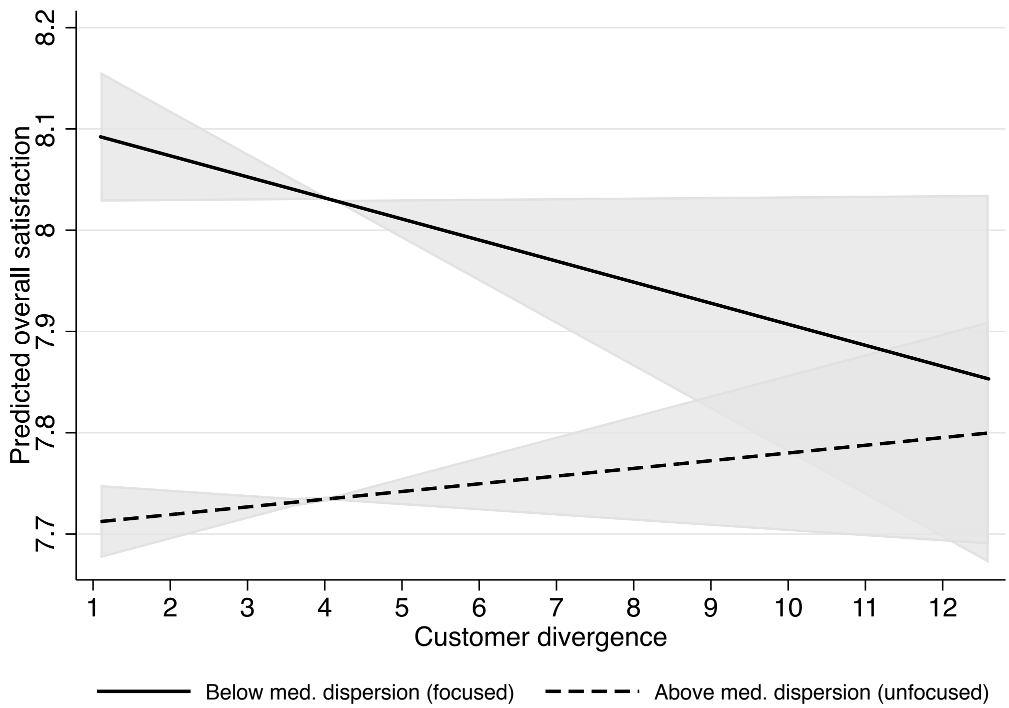

Figure 2: Plots of the interdependent effects of Customer Divergence and Customer Dispersion on Service Satisfaction (Study 2). Consistent with H3, the negative effects of customer divergence are most pronounced among more focused firms, with less dispersed customer bases. Figure is plotted within support of 99% of the data, with a 95% confidence interval band, estimated with robust standard errors clustered by zip code, shown in light grey.

These results are interesting, because they suggest that although focused operating models are beneficial for service outcomes in general, a customer whose needs diverge from the capabilities of the focused operation may exhibit lower levels of satisfaction. In contrast, for the least focused firms, which exhibit lower overall levels of satisfaction in general, customer divergence tends to result in no change in satisfaction (**Figure 2**).

2.2.5 Product-based divergence analysis. Although a strength of the previous analysis is that it analyzes the satisfaction of more than 140,000 customers interacting with 164 different retail banking institutions, it is limited in the sense that customer needs are proxied by the aggregate demographic characteristics of the zip codes where each customer resides. Although the pilot study discussed above, and presented in full in the online appendix, provides evidence in support of this approach, unlocking the potential to perform industry-level analyses with data that are relatively easy to obtain, the results also reveal that an even higher fidelity measure maps directly to each individual. Therefore, as an additional test of H1, we collaborated with the focal retail bank from Study 1, gathering data from 49,582 customers who engaged in at least two retail banking transactions during 2015 and 2016 where they were randomly selected to be surveyed about their experiences. The focal bank asked these customers a series of questions, assessed on a 1-5 scale, to capture their satisfaction with relevant facets of each interaction (e.g., overall satisfaction, knowledge, helpfulness and friendliness of the representative, ease of the interaction, intended loyalty, etc. – see the online appendix for the full list of questions). Aggregated satisfaction was relatively high among the surveyed customers during the period of analysis (*M*=4.51, *SD* = 0.92).

	(1) Satisfaction	(2) Satisfaction	(3) Satisfaction	(4) Satisfaction	(5) Satisfaction	(6) Satisfaction Δ	(7) Satisfaction Δ
Divergence	-0.008***	-0.008***	-0.008***	-0.007***	-0.008***		
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)		
Divergence Δ from first interaction						-0.003**	-0.007*
						(0.002)	(0.004)
Ln(average balance)			0.002	0.005***	0.004**	0.003	0.005
			(0.002)	(0.002)	(0.002)	(0.003)	(0.004)
Tenure		-0.000	-0.001	0.003***	0.001*	0.005***	0.004
		(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Tenure²		0.000***	0.000***	0.000	0.000***	-0.000***	-0.000*
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	4.586***	4.569***	4.565***	4.957***	4.528***	-0.053	-0.103
	(0.021)	(0.021)	(0.023)	(0.025)	(0.021)	(0.073)	(0.088)
Observations	49,582	49,582	47,750	47,574	95,324	26,053	12,626
Customers	49,592	49,592	47,750	47,454	48,683	26,053	12,626
Transaction included	First	First	First	Second	Both	Second	Second
Sample	All	All	All	All	All	All	Same products
R-squared	0.005	0.007	0.006	0.008	0.007	0.003	0.004

Table 6: Customer divergence leads to diminished satisfaction in a product-based analysis. Consistent with H1, customer divergence is negatively associated with customer satisfaction. Moreover, Columns (6-7) provides evidence that increases in customer divergence, that are exogenous to the customer, leads to declines in satisfaction. All models include counts of the number of products held in each of the bank's 13 product categories, and time-based fixed effects. Robust standard errors, clustered at the customer level, are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

To measure customer divergence for each customer, we captured counts of each of the different types of products the customer had at the time they were surveyed, as well as the average count of each type of product held by all customers of the bank at the time of the survey. Following the strategy detailed in Equation (2) therefore, we could assess how each customer's portfolio of products – an indicator that is revelatory of their specific banking needs – compared with the product holdings of the bank's average customer. For each product c offered by the bank at time t, we calculated a normalized divergence metric, ND_{it}, which evaluated how the count of that product held by customer i compared with the holdings of the average customer of the bank, $\overline{c}_t$, normalized by the standard deviation of the count of that type of product held by all customers of the bank during that time period. To the extent that the bank's capabilities in meeting particular needs should be honed and improved through interactions with its customers, we would expect that it will most capably deliver on the basket of needs typified by its average customer. Therefore, the normalized divergence metric for each product was aggregated as in Equation (3), to create an overall product-based divergence metric for each customer, DIV_{it} (*M*=19.85, *SD*=7.32). Larger values of this measure (99th percentile=45.55) indicated customers whose needs, as exhibited by their chosen portfolio of products, diverged most markedly from the bank's capabilities, while smaller values (1st percentile=11.23), indicated customers whose needs were more closely aligned with the bank's capabilities.

To control for additional facets of the customer's relationship with the bank that may affect differences in their satisfaction at the time of the survey, SAT_{it}, we further capture time-varying facets of the relationship, including the natural logarithm of 1+ their total balance across all accounts they held, $\ln\left(BAL_{it}\right)$, the non-linear effects of tenure, $TENURE_{it}$, and $TENURE_{it}^2$, counts of the number of products they held in each of the bank's 13 product categories, $PRODCT_{it}$ and a month and year indicator, X_t, to account for time-varying differences in service performance of the bank over time. We estimate the following cross-sectional model with robust standard errors, clustered by customer, for the first, second, and both observations:

$$\begin{gathered} SAT_{it} = \kappa_0 + \kappa_1 DIV_{it} + \kappa_2 \ln\left(BAL_{it}\right) + \kappa_3 TENURE_{it} \\ +\kappa_4 TENURE^2{}_{it} + PRODCT_{it} + X_t + \epsilon_{it} \end{gathered} \tag{9}$$

To the extent that divergence between the capabilities of an operation and the needs of a customer is negatively associated with customer satisfaction, as predicted in H1, we would expect that $\kappa_1 < 0$. Indeed, Columns (1-5) in **Table 6** offer converging evidence in support of H1. Controlling for other factors, customers exhibiting a higher degree of product-based divergence report lower levels of satisfaction in their first interaction (κ=-0.008, p<0.01), second interaction (κ=-0.007, p<0.01), and when both are jointly considered (κ=-0.008, p<0.01). In the fully-controlled specification in column (3), considering only the first measured interaction, an increase in customer divergence from the 1st to the 99th percentile in the sample corresponds with a 0.28 drop in satisfaction – a 6.1% decline from baseline rates, which is equivalent to 30% of the standard deviation in satisfaction. Moreover, as a test of the notion that the link between customer divergence and diminished satisfaction is causal, we refine Equation (9) by modelling the change in satisfaction for customer *i* from the first to the second interaction we observe, as a function of the change in customer divergence from the first to the second interaction, again estimated with robust standard errors, clustered by customer:

$$
\begin{aligned}
\left(SAT_{it} - SAT_{it-1}\right) = \lambda_0 + \lambda_1\left(DIV_{it} - DIV_{it-1}\right) + \lambda_2 COUNT_{it} + \lambda_3 \ln BAL_{it} \\
+\lambda_4 AGE_{it} + \lambda_5 AGE_{it}^{\,2} + \lambda_6 TENURE_{it} + PRODCT_{it} + X_t + \epsilon_{it}
\end{aligned}
\tag{10}
$$

Importantly, we perform this analysis only on the subsample of 26,053 customers for whom at least a year elapsed between their first and second interaction to ensure sufficient separation between observations. Column (6) demonstrates that as customer divergence increases, customer satisfaction falls (λ=-0.003, p<0.05). Column (7) further refines the identification of a causal effect by examining observations only from the 12,626 customers who held the same products from the first observation to the second. In this way, the change in customer divergence being modelled is completely exogenous to the customer, attributable only to changes in the composition of the focal bank's customer base. The results provide evidence that's consistent with a causal relationship, wherein customer divergence causes diminished satisfaction (λ=-0.007, p<0.10).

2.3 Study 3: Customer Compatibility and Firm Performance

The results of the previous studies suggest that differences among customers drive a meaningful portion of the variation in customer satisfaction, and that customer compatibility – fit between the needs of customers and the capabilities of the operation serving them – accounts for a portion of these differences. To the extent that firms that pursue more compatible customer relationships are able to deliver more satisfying service interactions, an important question is whether customer compatibility drives firm performance.

2.3.1 Data and empirical approach. To address this question, we leverage the methodology used in Study 2 to calculate annual metrics of divergence for each branch of each bank insured by the FDIC from 2006-2017, as well as annual metrics of customer dispersion for each banking institution over the time period. To accomplish this, we use the 2007-2011, 2008-2012, 2009-2013, 2010-2014, 2011-2015, and 2012-2016 American Community Survey's 5-year estimates, tabulated at the zip code level, to capture annual variation in the demographic composition of the market around each branch, as well as year-to-year variation in the composition of each firm's customer base. To measure financial performance, we use annual branch-level deposit information from the 2006-2017 FDIC Summary of Deposits databases. On June 30 of every year, every branch of every federally-insured retail bank in the United States reports its balance to the FDIC. To the extent that compatible customers are more satisfied with their service provider, and satisfied customers are more profitable over the long run than dissatisfied customers (Hallowell 1996, Heskett et al. 1997), we hypothesize that branches where customers exhibit a greater degree of divergence from the characteristics of customers more typically served by the firm will exhibit slower deposit growth than branches where customers are less divergent.

Hypothesis 4 (H4): Divergence between the capabilities of an operation and the needs of the customers it serves is negatively associated with a firm's financial performance.

Moreover, consistent with the theory that customer compatibility is a primary driver of service performance, we hypothesize that banks with customer bases whose needs are more broadly dispersed will offer services that are less aligned on average with the distinct needs of each of its customers, leading to diminished financial performance. However, consistent with the logic outlined in Section 2.2.1, we predict that dispersion among the needs of customers will negatively affect financial performance at a diminishing rate.

Hypothesis 5 (H5): Dispersion among the needs of customers served by an operation is negatively associated with financial performance, but at a diminishing rate.

We further hypothesize that the interaction between customer divergence and customer dispersion that influences the level of satisfaction an operation can deliver will likewise influence the level of financial performance the firm can achieve. Specifically, we hypothesize that the negative effects of customer divergence on firm performance will loom largest for firms with more focused, less dispersed customer bases.

Hypothesis 6 (H6): Customer dispersion positively moderates the negative association between customer divergence and a firm's financial performance.

Our analysis proceeds in parallel with the analyses introduced in Section 2 above. We begin by modeling financial performance as a function of customer divergence (H4). Next, we model financial performance as a function of customer dispersion (H5). In the final analysis, we consider the interdependent effects of customer divergence and dispersion (H6). In all analyses, financial performance is operationalized as the difference in the natural logarithm of deposit balances from one year to the next, $Ln\left(DEPOSITS_{ijt}\right)-Ln\left(DEPOSITS_{ijt-1}\right)$. Hence, all coefficients presented in the log-linear models that follow can be interpreted as the percentage change in deposit growth expected by changing the focal coefficient by 1 unit.

2.3.2. Customer divergence and firm performance (H4). For the branch in zip code *i*, belonging to institution *j*, during year *t*, we model deposit growth as a function of branch, institution, and zip-code level factors, including customer divergence, DIV_{ijt}, whether the branch was opened during the 12-year period of analysis, $NEWBR_{ij}$, whether the branch was acquired during the period of analysis, $ACQBR_{ij}$, the lagged natural logarithm of branch-level deposits, $\ln BRDEP_{iit-1}$, the number of branches the institution had during the year, $BRCOUNT_{jt}$, the natural logarithm of total deposits for the year, $\ln DEP_{jt}$, and measures of the time-varying demographics in the zip code, *i* (population density, per capita income, median age, average household size, the proportion of the population between 18 and 34, and the proportion of the population that owned their home). We also control for zip code-level fixed effects and a vector of year indicators. In the fullest specifications, we reduce our analysis to the 52 largest institutions that had above-median branch counts, and include an additional vector of institution-level fixed effects. The following linear fixed effects model is estimated using robust standard errors clustered at the branch level.

$$\begin{aligned}\ln\left(DEPOSITS_{ijt}\right)-\ln\left(DEPOSITS_{ijt-1}\right)=\eta_0+\eta_1 DIV_{ijt}+\eta_2 NEWBR_{ij}+\eta_3 ACQBR_{ij}+\eta_4 \ln BRDEP_{ijt-1}+\\ \eta_5 BRCOUNT_{jt}+\eta_6 \ln DEP_{jt}+\eta_7 POPDENSITY_{ijt}+\\ \eta_8 PCAPINC_{ijt}+\eta_9 MEDAGE_{ijt}+\eta_{10} HHSIZE_{ijt}+\\ \eta_{11} PROP18t34_{ijt}+\eta_{12} PROPOWN_{ijt}+\alpha_i+\beta_j+\gamma_t+\epsilon_{ijt}\end{aligned} \tag{11}$$

In the model above, if $\eta_1 < 0$, then consistent with H4, customer divergence is negatively associated with financial performance. Consistent with H4, **Table 7** Column (1) demonstrates that customer divergence is negatively associated with financial performance as measured by branch level deposit growth (η=-0.006, *p*<0.01), results that are robust and consistent in Column (2) after controlling for institution level fixed-effects (η=-0.006, *p*<0.01). Increasing a branch's level of customer divergence by one standard deviation results in a 1.4% decrease in branch deposit growth, a 20.9% reduction from baseline deposit growth rates of 6.7% per year, suggesting that customer divergence has a very meaningful effect on attracting core deposits. Column 3 demonstrates that deposit growth during the time period was sharply lower among branches that were first opened during the period of analysis (η=-0.011, *p*<0.01) and among branches that were acquired during the period of analysis (η=-0.031, *p*<0.01). These patterns appear consistent with the idea that new branches and branch acquisitions represented institutional forays into more divergent markets, where customers may have been less compatible on average with the capabilities of the operation. Indeed, in the split sample analysis in Column (4), which exclusively focuses on branches that were neither opened nor acquired during the period of analysis, the negative effect of customer divergence on deposit growth is robust, but diminished (η=-0.002, *p*<0.01).

Moreover, just as deposits are an important determinant of financial performance in retail banking, so too are loans. In an analysis documented in the online appendix, we demonstrate that this same pattern of results holds for growth of net loans. Increasing average branch-level divergence by a standard deviation reduces the institution-level annual growth in net loans by 0.92% – a 13.3% decline relative to baseline growth rates of 6.9% per year. These results suggest that the financial performance impact of customer incompatibility may not be isolated to deposits, but may instead be more holistic.

2.3.3. Customer dispersion and firm performance (H5). Similarly, we model deposit growth as a function of customer dispersion, DIS_{ijt}, the quadratic term DIS^2, and the same branch and zip code level controls described above. The one exception is that consistent with our earlier dispersion analyses, we do not include an institution-level fixed-effect, owing to the fact that customer dispersion is largely a time-invariant, institution-level characteristic. The following linear fixed effects model is estimated using robust standard errors clustered at the branch level.

$$\begin{aligned}\ln\left(DEPOSITS_{ijt}\right)-\ln\left(DEPOSITS_{ijt-1}\right) = \theta_0 + \theta_1 DIS_{ijt} + \theta_2 DIS^2{}_{ijt} + \theta_3 \ln BRDEP_{ijt-1} + \\ \theta_4 BRCOUNT_{jt} + \theta_5 \ln DEP_{jt} + \theta_6 POPDENSITY_{ijt} + \\ \theta_7 PCAPINC_{ijt} + \theta_8 MEDAGE_{ijt} + \theta_9 HHSIZE_{ijt} + \\ \theta_{10} PROP18t34_{ijt} + \theta_{11} PROPOWN_{ijt} + \alpha_i + \beta_j + \gamma_t + \epsilon_{ijt}\end{aligned} \quad (12)$$

	(1) ΔLn(Deposits)	(2) ΔLn(Deposits)	(3) ΔLn(Deposits)	(4) ΔLn(Deposits)	(5) ΔLn(Deposits)	(6) ΔLn(Deposits)
Customer divergence	-0.006***	-0.006***	-0.006***	-0.002***		-0.013***
	(0.000)	(0.001)	(0.000)	(0.000)		(0.002)
New branch indicator			-0.011***			
			(0.001)			
Acquired branch indicator			-0.031***			
			(0.001)			
Customer dispersion					-0.110***	
					(0.003)	
Customer dispersion^2					0.022***	
					(0.001)	
Divergence x dispersion						0.003***
						(0.001)
Ln(1+Lagged branch deposits)	-0.106***	-0.123***	-0.108***	-0.051***	-0.108***	-0.123***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
Branch count	0.000***	-0.000***	0.000***	0.000***	0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Total institution deposits)	0.000	0.073***	-0.000	-0.000	0.005***	0.072***
	(0.000)	(0.005)	(0.000)	(0.000)	(0.000)	(0.005)
Population density (per sq. mile)	0.000***	-0.000	0.000***	0.000	0.000***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Per capita income	0.000***	-0.000	0.000***	-0.000***	0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Median age	0.001***	0.001	0.001***	0.002***	0.001**	0.001
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)
Average household size	0.023***	0.006	0.024***	0.029***	0.019***	0.005
	(0.005)	(0.007)	(0.005)	(0.004)	(0.005)	(0.007)
Proportion of population aged 18-34	0.025	0.056	0.027	0.036	0.014	0.049
	(0.029)	(0.044)	(0.029)	(0.026)	(0.029)	(0.044)
Proportion of population owning homes	0.024	0.010	0.025*	-0.007	0.021	0.010
	(0.015)	(0.025)	(0.015)	(0.013)	(0.015)	(0.025)
Constant	1.056***	0.058	1.083***	0.487***	1.117***	0.096
	(0.029)	(0.086)	(0.029)	(0.028)	(0.029)	(0.086)
Observations	941,890	477,384	941,890	464,860	941,890	477,384
Sample	All branches	Above-median branch count branches	All branches	Non-acquired, non-new branches	All branches	Above-median branch count branches
Institution fixed effects	No	Yes	No	No	No	Yes
Adjusted R-squared	0.128	0.166	0.130	0.051	0.130	0.166

Table 7: Customer divergence, customer dispersion, and financial performance (Study 3). All models include zip code-level fixed effects, and (2) and (6) include institution-level fixed effects. Year indicator variables are included in all models, but are withheld from the table for parsimony. Robust standard errors, clustered by institution are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

If $\theta_1 < 0$ and $\theta_2 > 0$, then following its relationship with customer satisfaction, and consistent with H5, customer dispersion is negatively associated with a firm's level of financial performance, albeit at a decreasing rate. **Table 7**, Column (5) provides evidence consistent with the patterns observed earlier, and with H5, that customer dispersion diminishes deposit growth (θ=-0.11, *p*<0.01), but at a diminishing rate (θ=0.022, *p*<0.01). Consistent with the results of Study 2, piecewise linear regression analyses presented in the online appendix reveal that deposit growth falls with dispersion in the bottom three quartiles, and rises in the top quartile – an upward-sloping relationship that does not arise for net loan growth. Increasing an institution's level of dispersion from the 1st percentile (a very focused firm) to the 50th percentile (a less-focused firm with a median level of customer dispersion) reduces deposit growth by 13.9% (from 19.1% to 5.2%). Increasing dispersion in the same way reduces net annual loan growth by 2.1% (from 7.4% to 5.3%).

2.3.4. Interdependent effects of customer divergence and customer dispersion on firm performance (H6). Consistent with the above, and the preceding section, we model the interdependent effects of customer divergence and customer dispersion on financial performance using the following linear fixed effects model, estimated using robust standard errors clustered at the branch level.

$$\begin{aligned}\ln\left(DEPOSITS_{ijt}\right)-\ln\left(DEPOSITS_{ijt-1}\right) = {} & \iota_0+\iota_1 DIV_{ijt}+\theta_2 DIS_{ijt}\times DIV_{ijt}+\iota_3\ln BRDEP_{ijt-1}+\\ & \iota_4 BRCOUNT_{jt}+\iota_5\ln DEP_{jt}+\iota_6 POPDENSITY_{ijt}+\\ & \iota_7 PCAPINC_{ijt}+\iota_8 MEDAGE_{ijt}+\iota_9 HHSIZE_{ijt}+\\ & \iota_{10} PROP18t34_{ijt}+\iota_{11} PROPOWN_{ijt}+\alpha_i+\beta_j+\gamma_t+\epsilon_{ijt}\end{aligned} \quad (13)$$

If $\iota_1 < 0$ and $\iota_2 > 0$, then consistent with H6 and the patterns observed with customer satisfaction, a branch's level of customer divergence is negatively associated with deposit growth, and more focused firms (with lower levels of customer dispersion) suffer from customer divergence more. **Table 7** Column (6) demonstrates a pattern of results consistent with H6 and the pattern observed for satisfaction, wherein the effects of customer divergence on financial performance are most acute among focused firms. A very focused firm exhibits significant declines in deposit growth as customer divergence increases (l=-0.013, *p*<0.01). However, the effects are attenuated for less focused firms, with higher degrees of customer dispersion (l =0.003, *p*<0.01). A one standard deviation increase in customer divergence for a branch of a maximally focused firm would experience a 1.8% reduction in deposit growth (a 29.1% reduction from baseline rates), whereas a similar increase in customer divergence for a firm with a median level of dispersion would experience a 1.2% reduction in deposit growth (a 16.6% reduction from baseline rates).

3. General Discussion

In this paper, we have empirically documented the effects of customer compatibility – the degree of fit between the needs of individual customers and the capabilities of the operations serving them – on customer experiences and firm performance. For our focal firm, we have shown that differences among customers account for the vast majority of the explainable variance in transaction satisfaction, while differences among employees, processes, branches, and markets account for relatively little variance. Our analysis further suggests that although differences in satisfaction exist among customers, individual customers tend to report relatively consistent levels of satisfaction from one transaction to the next. On the other hand, employees, processes, branches, and markets tend to deliver inconsistent levels of satisfaction from transaction to transaction, offering support for the idea that differences in service outcomes emanate in large part from differences among customers (Study 1).

In Study 2, we introduce a methodology for empirically quantifying customer compatibility, which we leverage to investigate the experiences of 145,761 customers interacting with 164 banking institutions over a five-year period. Using demographic proxies for customer needs, we provide evidence that these customer-level differences are explained in part by customer compatibility (Study 2). Customers whose needs are more aligned with those of their service providers' most typical customers report systematically higher levels of satisfaction on a broad array of operating dimensions than customers whose needs are more divergent. Consistently, we find that firms that serve customer bases whose needs are more dispersed have customers who are less satisfied overall than firms serving more focused customer bases. Our results further suggest that the negative effects of customer incompatibility are especially acute for firms with more focused customer bases. A parallel analysis that uses the portfolio of products chosen by 49,582 customers of a single banking institution to measure customer divergence provides converging, causal evidence that customer incompatibility reduces customer satisfaction.

Finally, in line with the patterns of relationships detected among customer compatibility and customer experiences, our investigation reveals that customer compatibility has a significant and meaningful effect on a firm's financial performance (Study 3). Bank branches serving customers whose needs are more divergent from the needs of their firm's most typical customers exhibit slower deposit and loan growth than branches serving more compatible customers. Institutions serving customer bases with more dispersed needs exhibit slower deposit and loan growth than institutions serving customer bases whose needs are more compatible with one another.

These results lend support to a broad theoretical literature that advocates reducing the variation imposed on operating systems by customers (Karmarkar and Pitbladdo 1995), though by identifying customer compatibility as a persistent and systematic customer-level characteristic, it suggests a new set of strategic tools for doing so, which could serve as a fruitful area for future study. Indeed, the notion of customer compatibility builds on the concept of operating segments introduced by Frei and Morriss (2012), which are groups of customers who share priorities about the dimensions of service that matter most and least to them. Customer compatibility extends this idea by quantifying the degree of fit between the needs of customers, and the capabilities of operations, providing the first empirical evidence of how this degree of fit systematically and persistently impacts performance.

Future research can investigate the efficacy of various strategies for improving customer compatibility, either through the selection of more compatible customers, or through the customization of offerings to better align with individuals' needs. For example, managers may choose to pursue strategies that reduce incompatibility through customer dispersion by prudently selecting markets where customers' needs exhibit greater homogeneity or by designing service offerings that are tightly aligned with the needs of the customers they serve. Complimentarily, managers may seek strategies to reduce customer divergence by being more transparent with prospective customers about the tradeoffs in their operating models (Buell and Choi 2019), or by designing service offerings that can be more readily customized to suit individuals' needs (Guajardo and Cohen 2018). Future work could explore the costs and benefits of such strategies.

Interestingly, in some service contexts like retail banking, the decision to improve customer experiences by reducing customer dispersion may run afoul of competing strategic priorities. For example, for a bank, pursuing a dispersed customer base might help diversify risk in its loan portfolio, and expanding into divergent markets might enhance the its value proposition as a financial intermediary by connecting heterogeneous populations of investors and borrowers. Future research could examine the strategic tradeoffs inherent in actively managing customer compatibility. Moreover, future research could seek to develop models that explain a greater degree of the variance in customer satisfaction. Devising higher fidelity satisfaction models could both facilitate the distinction between compatible and incompatible customers, as well as assist in the customization of service to better align the capabilities of an offering with the needs of particular customers. Additionally, although the present research documents the effects of customer incompatibility, it does not investigate its origins. Disentangling the drivers of customer incompatibility, for example, whether it emanates more from information asymmetries, a lack of options in a customer's local market, or even from acquisitions wherein a customer's compatibility is shifted through no fault of their own, could help shine light on which approaches for managing it might hold the most promise. We hope that the present research spurs more work in our field that considers how service performance can be enhanced by improving the degree of fit between the needs of customers and the capabilities of the operations that serve them.

References

Anand KS, Goyal M (2009) Strategic Information Management Under Leakage in a Supply Chain. *Manage. Sci.* 55(3):438–452.

Anderson E (1994) Cross-Category Variation in Customer Satisfaction and Retention . *Mark. Lett.* 5(1):19–30.

Baumann C, Burton S, Elliott G (2007) Determinants of customer loyalty and share of wallet in retail banking. *J. Financ. Serv. Mark.*

Berger AN, Rosen RJ, Udell GF (2007) Does market size structure affect competition? The case of small business lending. *J. Bank. Financ.* 31(1):11–33.

Besbes O, Scarsini M (2018) On Information Distortions in Online Ratings. *Oper. Res.* Forthcomin.

Buell RW, Campbell D, Frei FX (2010) Are self-service customers satisfied or stuck? *Prod. Oper. Manag.* 12(6):679–697.

Buell RW, Campbell D, Frei FX (2016) How Do Customers Respond to Increased Service Quality Competition? *Manuf. Serv. Oper. Manag.* 18(4):1–33.

Buell RW, Choi M (2019) *Improving customer compatibility with operational transparency*.

Buhrmester M, Kwang T, Gosling SD (2011) Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspect. Psychol. Sci.* 6(1):3–5.

Cameron CA, Gelbach JB, Miller DL (2011) Robust inference with multiway clustering. *J. Bus. Econ. Stat.*

Campbell D, Datar SM, Sandino T (2009) Organizational Design and Control across Multiple Markets : The Case of Franchising in the Convenience Store Industry. *Account. Rev.* 84(6):1749–1779.

Campbell D, Frei F (2010) Market Heterogeneity and Local Capacity Decisions in Services. *Manuf. Serv. Oper. Manag.* 13(1):2–19.

Chase RB (1978) Where does the customer fit in a service operation? *Harv. Bus. Rev.* 56(6):137–142.

Cohen MA, Whang S (1997) Competing in product and service: A product life-cycle model. *Manage. Sci.* 43(4):535–545.

Cortiñas M, Chocarro R, Villanueva ML (2010) Understanding multi-channel banking customers. *J. Bus. Res.* 63(11):1215–1221.

Debo L, Veeraraghavan S (2010) *Prices and congestion as signals of quality*

Devlin JF, Gerrard P (2004) Choice criteria in retail banking: an analysis of trends. *J. Strateg. Mark.* 12(1):13–27.

Erel I (2009) The effect of bank mergers on loan prices: Evidence from the United States. *Rev. Financ. Stud.* 24(4):1068–1101.

Fitzsimmons JA, Fitzsimmons MJ (2006) *Service Management: Operations, Strategy and Information Technology* Fifth Edit. (McGraw-Hill Irwin, Boston, MA).

Frei F, Morriss A (2012) *Uncommon Service: How to Win by Putting Customers at the Core of Your Business* (Harvard Business Review Press, Boston).

Frei FX (2006) Breaking the trade-off between efficiency and service. *Harv. Bus. Rev.* 84(11):93–101.

Guajardo JA, Cohen MA (2015) *After-Sales Services Differentiation in Product-Service Bundles*

Gupta S, Chintagunta P (1994) On Using Demographic Variables to Determine Segement Membership in Logit Mixture Models. *J. Mark. Res.* 31(1):128–136.

Hadfield JD (2010) MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* 33(2):1–22.

Hallowell R (1996) The relationships of customer satisfaction, customer loyalty, and profitability: an empirical study. *Int. J. Serv. Ind. Manag.* 7(4):27.

Heskett J, Sasser E, Schlesinger L (1997) *The Service Profit Chain* (The Free Press, New York).

Ho TH, Zheng YS (2004) Setting customer expectation in service delivery: An integrated marketing-operations perspective. *Manage. Sci.* 50(4):479–488.

Hoch S, Kim B, Montgomery A, Rossi P (1995) Determinants of Store-Level Price Elasticity. *J. Mark. ...* 32(1):17–29.

Homburg C, Koschate N, Hoyer WD (2005) Do Satisfied Customers Really Pay More? A Study of the Relationship Between Customer Satisfaction and Willingness to Pay. *J. Mark.* 69(2):84–96.

Israel M (2005) Services as experience goods : An empirical examination of consumer learning in automobile insurance. *Am. Econ. Rev.* 95(5):1444–1463.

Jackman S (2000) Estimation and Inference via Bayesian Simulation: An Introduction to Markov Chain Monte Carlo. *Am. J. Pol. Sci.* 44(2):375–404.

Kalyanam K, Putler DS (1997) Incorporating Demographic Variables Brand Choice Models : An in Indivisible Alternatives Framework. *Mark. Sci.* 16(2):166–181.

Karmarkar U, Pitbladdo R (1995) Service Markets and Competition. *J. Oper. Manag.* 12:397–411.

Kelley SW, Donnelly, Jr. JH, Skinner SJ (1990) Customer Participation in Service Production and Delivery. *J. Retail.* 66(3):315–335.

Kivetz R, Simonson I (2000) The Effects of Incomplete Information on Consumer Choice. *J. Mark. Res.* 37(4):427–448.

LaBarbera P, Mazursky D (1983) A Longitudinal Assessment of Consumer Satisfaction / Dissatisfaction: The Dynamic Aspect of the Cognitive Process. *J. Mark. Res.* 20(4):393–404.

Lariviere MA, Padmanabhan V (1997) Slotting Allowances and New Product Introductions. *Mark. Sci.* 16(2):112–128.

Levesque T, McDougall GHG (1996) Determinants of customer satisfaction in retail banking. *Int. J. Bank Mark.* 14(7):12–20.

Martenson R (1985) Consumer choice criteria in retail bank selection. *Int. J. Bank Mark.* 3(2):64–74.

Martin C, Pranter C (1989) Compatibility Management : Customer-To-Customer Relationship. *J. Serv. Mark.* 3(3):5–15.

Mason W, Suri S (2012) Conducting behavioral research on Amazon's Mechanical Turk. *Behav. Res. Methods* 44(1):1–23.

McGahan AM, Porter ME (1997) How Much Does Industry Matter, Really? *Strateg. Manag. J.* 18:15–30.

Mulhern FJ, Williams JD (1994) A Comparative Analysis of Shopping Behavior in Hispanic and Non-Hispanic Market Areas. *J. Retail.* (3):231–251.

Olivares M, Cachon GP (2009) Competing retailers and inventory: An empirical investigation of General Motors' dealerships in isolated U. S. markets. *Manage. Sci.* 55(9):1586–1604.

Oliver R (1980) A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *J. Mark. Res.* XVII:460–469.

Rumelt RP (1991) How Much Does Industry Matter? *Strateg. Manag. J.* 12(3):167.

Sampson SE, Froehle CM (2006) Foundations and implications of a proposed unified services theory. *Prod. Oper. Manag.* 15(2):329–343.

Schmalensee R (1985) Do Markets Differ Much? *Am. Econ. Rev.* 75(3):341–351.

Schmidt W, Buell RW (2017) Experimental Evidence of Pooling Outcomes Under Information Asymmetry. *Manage. Sci.* 63(5):1586–1605.

Schneider B, Parkington JJ, Buxton VM (2013) Employee and Customer Perceptions of Service in Banks. 25(2):252–267.

Skinner W (1974) The focused factory. *Harv. Bus. Rev.* 52(May-June):113–122.

Stafford MR (1996) Demographic discriminators of service quality in the banking industry. *J. Serv. Mark.* 10(4):6–22.

Sutton J (1986) Vertical product differentiation: Some basic themes. *Am. Econ. Rev.* 76(2):393–398.

Taylor S (1994) Waiting for service: The relationship between delays and evaluations of service. *J. Mark.* 58(2):56–69.

Tsay AA, Agarwal N (2000) Channel dynamics under price and service competition. *Manuf. Serv. Oper. Manag.* 2(4):372–391.

Tsikriktsis N (2007) The Effect of Operational Performance and Focus on Profitability: A Longitudinal Study of the U.S. Airline Industry. *Manuf. Serv. Oper. Manag.* 9(4):506–517.

Online appendix for “The Customer May Not Always Be Right: Customer Compatibility and Service Performance”

A1. Variance Decomposition Analysis

	Mean	Std. Dev.	Min	Max
Outcome summary statistics				
Satisfaction with interaction	4.73	0.63	1	5
Overall satisfaction with bank	4.56	0.74	1	5
Interaction summary statistics				
Transactions/interaction	2.87	2.18	1	28
Interaction duration (minutes)	3.63	15.54	0	558
Customer perceived wait (minutes)	3.73	12.59	0	98
Processes involved in interaction				
Balance inquiry	17.7%	38.2%	0	1
Bank product purchase	0.1%	2.6%	0	1
Cash bond	0.0%	2.0%	0	1
Cash outside check	0.3%	5.3%	0	1
Cash inside check	16.4%	37.0%	0	1
Cash check (owner)	0.0%	0.8%	0	1
Make deposit	88.6%	31.8%	0	1
Miscellaneous credit	0.0%	1.2%	0	1
Miscellaneous fee	0.0%	1.0%	0	1
Miscellaneous	39.6%	48.9%	0	1
Payment	1.2%	10.8%	0	1
Purchase cash instrument	9.0%	28.6%	0	1
Purchase wire	0.0%	2.2%	0	1
Verify funds	4.3%	20.2%	0	1
Withdrawal	0.3%	5.0%	0	1
Frequency statistics				
Observations/customer	2.14	0.40	2	6
Observations/process	438.30	2326.40	1	23,048
Observations/employee	2.15	1.82	1	42
Observations/branch	16.49	10.89	1	91
Observations/market	85.85	67.87	1	570

Table A1: Summary Statistics for variance decomposition analysis, n=58,294 (Study 1).

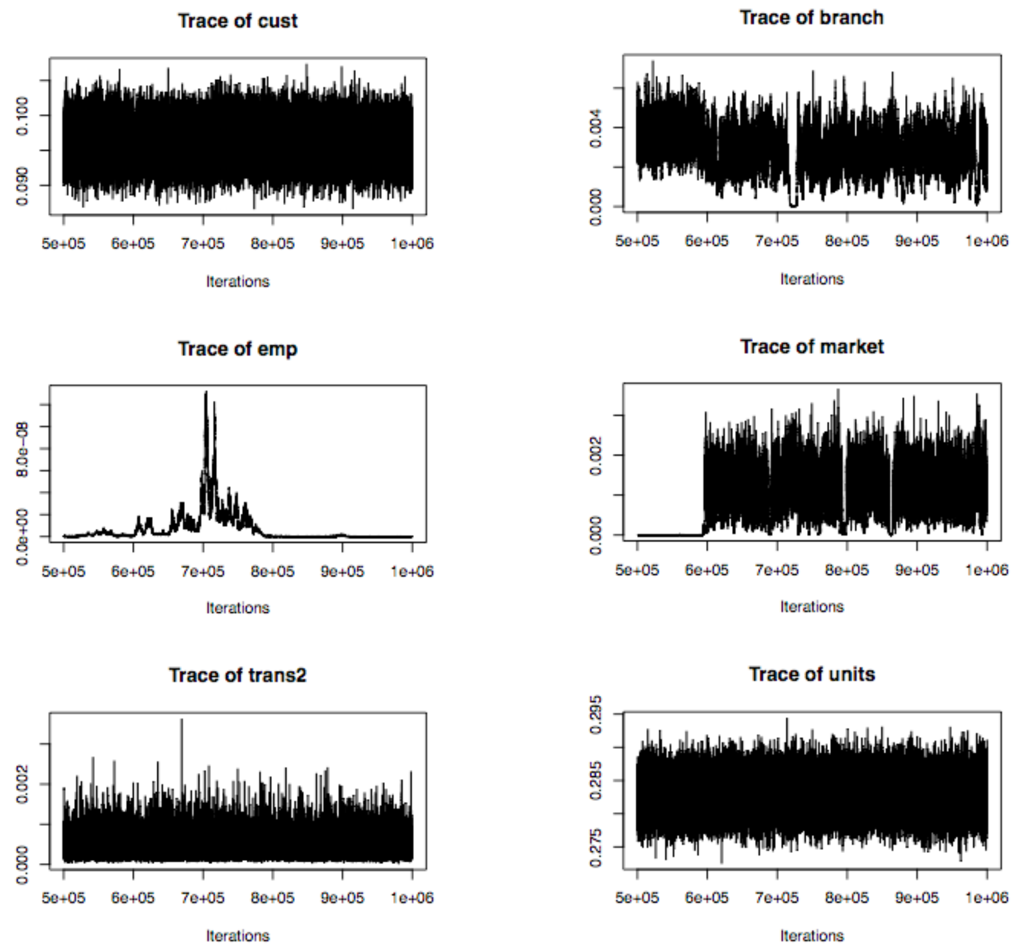

Figure A1: Convergence Traces of the Variance Components, based on Model 1 (Study 1A). These plots are based on an analysis with 1,000,000 iterations and a burn-in phase of 500,000 iterations.

	Model 1			Model 2			Model 3			Model 4		
	Mean	SD	Ratio	Mean	SD	Ratio	Mean	SD	Ratio	Mean	SD	Ratio
Customer	9.62E-02	2.64E-03	36.5	8.87E-02	2.51E-03	35.3	8.55E-02	2.72E-03	31.4	7.91E-02	2.56E-03	30.9
Process	4.35E-04	2.57E-04	1.7	6.09E-04	3.28E-04	1.9	3.33E-04	2.26E-04	1.5	5.39E-04	3.07E-04	1.8
Employee	6.32E-09	1.33E-08	0.5	6.79E-05	1.93E-04	0.4	5.62E-05	1.64E-04	0.3	4.85E-04	6.49E-04	0.7
Location	2.71E-03	1.11E-03	2.4	1.13E-03	9.64E-04	1.2	3.13E-03	9.41E-04	3.3	1.38E-08	8.32E-08	0.2
Market	8.35E-04	6.36E-04	1.3	1.41E-03	5.06E-04	2.8	1.61E-04	4.16E-04	0.4	1.59E-03	4.56E-04	3.5
Error	2.83E-01	2.63E-03	107.8	2.73E-01	2.55E-03	106.9	2.90E-01	2.79E-03	103.9	2.79E-01	2.72E-03	102.6
Fixed Effects	None			Wait Time			Lagged Overall Sat.			Wait Time, Lagged Overall Sat.		

Table A2: Means and Standard Deviations of the Variance Components, n=58,294 (Study 1A). These results reflect sources of variability calculated from estimates of Models (1-4), with 1,000,000 iterations, a burn-in phase of 500,000 iterations, and non-informative priors.

A2. Supplementary Analysis: Between and Within-Group Variance of QSRs

To assess whether the patterns presented in Study 1 generalize beyond our focal firm and the retail banking industry, and to assess whether the previously-observed satisfaction differences among customers might be explained by general differences in customer affect (e.g., perhaps some customers are habitually happier than others across all service interactions), we were able to obtain data from more than 3.7 million QSR transactions, conducted by 107,320 customers at 88,087 locations of 848 QSR brands operating in the United States in 2017. Customers had signed up to use a location-based smartphone app, which would randomly push surveys asking them to "please rate your overall satisfaction with your experience at *this restaurant,"* if the app detected that the user paused long enough to purchase a meal in a particular location with GPS coordinates corresponding with the site of a QSR restaurant. Importantly, this sampling strategy enabled the observation of the same customers interacting across multiple QSR brands. Consistent with the previous analysis, data were collected on the customer's level of visit satisfaction on a 5-point scale. Data were additionally collected on the identity of the restaurant, the location where the visit took place, and the identity of the customer, facilitating a similar between and within-group variance analysis, though with considerably more customer-level observations. The results in this QSR context are strikingly similar to the results of the previous analysis in retail banking.

	Between-group variance	Within-group variance	Total obs. with n>1	Number of groups	Mean obs./group
Customer within location	0.714	0.083	2,617,971	575,136	4.55
Customer within brand	0.687	0.179	3,344,299	523,461	6.39
Customer	0.468	0.441	3,709,740	107,320	34.57
Location	0.332	0.607	3,708,379	88,087	42.10
Brand	0.241	0.773	3,711,926	848	4,377.27

Table A3: Between and within-group variance in visit satisfaction for QSRs. Between and within-group variance above was calculated using the full dataset of customer responses for which each group of interest, for every level, had more than one observation, and the customer, location, and brand were identified.

Table A3 demonstrates relatively low between-group variance and relatively high within-group variance for brands (σ_b^2= 0.241; σ_w^2= 0.773) and locations (σ_b^2= 0.332; σ_w^2= 0.607); average satisfaction differences between brands and locations pale in comparison to the differences in satisfaction experienced by customers within them. Importantly, between and within-group variance for customers are comparable ($\sigma_b^2 = 0.468$; $\sigma_w^2 = 0.441$), suggesting that just as some customers offer higher or lower average evaluations of their experiences in aggregate, individual customers also tend to experience varying levels of satisfaction across their interactions with different brands and locations.

However, examining the between and within group variance of customer interactions with specific brands, or with specific brand locations, reveals a pattern of results that is strongly consistent with the results of the previous analysis. Customers interacting with specific brands exhibit high between group variance and low within group variance ($\sigma_b^2 = 0.687$; $\sigma_w^2 = 0.179$). Although there is considerable variation in the average experiences reported by different customers interacting with particular brands, individual customers tend to report great consistency in their interactions with specific brands over time. The pattern is even more striking for customers interacting with specific brand locations ($\sigma_b^2 = 0.714$; $\sigma_w^2 = 0.083$). Although some customers are a good fit and others are a bad fit for particular locations, the experiences individual customers have interacting with particular locations are remarkably consistent from one interaction to the next.

A3. Pilot Study: Mapping Zip Code Demographics to Individual Customer Needs

The primary data utilized in Studies 2 and 3 are subject to two layers of abstraction: 1) the average demographics of a zip code may not map precisely to the demographic characteristics of an individual who lives, and responds to a banking survey, within that zip code, and 2) although demographic characteristics have been shown to correlate with customer needs in prior research (Cortiñas et al. 2010, Gupta and Chintagunta 1994, Hoch et al. 1995, Kalyanam and Putler 1997, Mulhern and Williams 1994), they are only a proxy. We therefore conduct a pilot study to assess whether zip code-level demographics could provide a sufficiently high-fidelity view of varying customer needs to facilitate our empirical approach in Studies 2 and 3. In particular, we set out to test three hypotheses: 1) zip-code level demographic data from the ACS are predictive of respondent-level demographics from a particular zip code, 2) respondent-level demographics are predictive of differences in respondent-level needs in banking, and by extension, 3) zip code-level demographics are predictive of differences in respondent-level needs in banking.

A.3.1 Participants. 497 participants (42.5% female, M_{age}=35.61, *SD*=10.74) were recruited on the Amazon Mechanical Turk platform to participate in a 5-minute survey about their relationship with their primary banking institution in exchange for $1.00 (Buhrmester et al. 2011; Mason and Suri 2012).

A.3.2 Design and procedure. To facilitate tests of the three hypotheses outlined above, which are implicit in our empirical approach for Studies 2 and 3, each participant was asked a battery of questions about the channels through which they interacted with their bank (Buell et al. 2010), the products they held with their bank, and the relative importance of different drivers of satisfaction with their bank (Baumann et al. 2007, Levesque and McDougall 1996). Participants were also asked to provide the 5-digit zip code where they lived, to facilitate merging with ACS data, and to answer a series of individual-level demographic questions that mapped directly to the demographic characteristics that have been shown to correlate with customer needs in the prior research cited above.

	N	Mean	SD		N	Mean	SD
Respondent Demographics				**Drivers of satisfaction**			
Female %	497	42.45%	49.48%	Core	488	9.45	2.78
				When my bank promises to do something by a certain time, it does so	488	10.05	5.11
Age	497	35.61	10.74	My bank performs the service right the first time	488	9.02	5.65
Youngest in household	490	23.24	16.56	My bank provides its services at the time it promises to do so	488	9.93	5.18
Oldest in household	490	41.70	15.85	My bank performs the service accurately	488	7.44	5.64
Median age in household	490	32.47	13.48	My bank tells you exactly when services will be performed	488	10.81	5.19
Household income	497	49,839.60	31,537.63	Relational	488	11.39	2.52
Per capita income	497	22,917.94	18,542.38	Employees in my bank have the skills and knowledge to perform the service	488	10.32	4.90
				Employees in my bank are always willing to help	488	11.36	4.98
Household size	497	2.72	1.44	Employees in my bank are consistently courteous	488	11.62	5.14
				My bank gives me individual attention	488	11.49	5.39
Owned home	498	54.42%	49.85%	Employees of my bank understand my specific needs	488	12.17	4.90
Chanels of interaction				Tangibles	488	14.60	4.34
Face to face	497	13.25%	14.67%	My bank's physical facilities are visually appealing	488	14.53	5.28
Online	497	45.28%	28.97%	My bank's employees are neat in physical appearance	488	14.67	5.20
ATM	497	15.93%	15.95%				
Mobile banking	497	20.34%	22.54%	Enabling	488	10.46	2.62
Phone	497	5.19%	7.83%	My bank offers a complete range of services	488	10.00	5.52
				My bank has convenient branch locations	488	9.59	5.78
Products				My bank provides easily-understood statements	488	11.44	5.09
Checking account	497	94.77%	22.29%	It is very easy to get in and out of my bank quickly	488	10.80	5.57
Savings account	497	82.70%	37.87%				
Investment account	497	30.58%	46.12%	Competitive	488	10.46	5.79
Personal loan	497	23.54%	42.47%	My bank provides competitive interest rates	488	10.46	5.79
Mortgage loan	497	23.14%	42.21%				
Online billpay	497	63.38%	48.22%	Technology	488	8.10	3.71
Credit card	497	55.73%	49.72%	Using my bank's online banking is convenient for me	488	6.69	5.90
Debit card	497	91.55%	27.84%	Using my bank's mobile banking app is convenient for me	488	9.30	6.41
Safe deposit box	497	15.49%	36.22%	Using my bank's ATM is convenient for me	488	8.31	5.92

Table A4. Summary statistics for Pilot Study.

A.3.3 Behavioral and perceptual measures of banking needs. Individual-level differences in banking needs and preferences were measured in terms of channel usage, product usage, and satisfaction drivers. Summary statistics are presented in **Table A4**. To assess differences in channel usage, each participant was asked, “What percentage of your interactions with your bank are conducted in each of the following channels?” Participants then approximated percentages for face to face / within-branch interactions, online interactions, ATM interactions, mobile interactions, and phone interactions. To assess differences in product usage, each participant was asked, “Please select the types of products you have with your bank." Participants then denoted whether they held a checking account, savings account, investment account, personal loan, mortgage loan, online billpay, credit card, debit card, and safe deposit box. These product categories were selected to map directly to the product categories tracked by the focal institution with which we collaborated for the product-level analysis presented in Section 2.2.5.

Although channels of interaction and product selection are important behavioral indicators of individual differences of customer needs, there are important perceptual differences among customers as well. To assess these differences, we asked each participant to rank the relative importance of 20 different drivers of their satisfaction with their primary banking institution. We presented each participant with 17 drivers identified by Levesque and McDougall (1996), separated into five dimensions. Core quality drivers included, "When my bank promises to do something at a certain time, it does so," "My bank performs the service right the first time," "My bank provides its services at the time it promises to do so," "My bank performs the service accurately," and "My bank tells you exactly when services will be performed." Relational quality drivers included, "Employees in my bank have the required skills and knowledge to perform the service," "Employees in my bank are always willing to help," "Employees in my bank are consistently courteous," "My bank gives me individual attention," and "Employees of my bank understand my specific needs." Questions assessing tangibles included, "My bank's physical facilities are visually appealing," and, "My bank's employees are neat in appearance." Drivers of enablement included, "My bank offers a complete range of services," "My bank has convenient branch locations," "My bank provides easily-understood statements," and "It is very easy to get in and out of my bank quickly." The competitiveness driver was, "My bank provides competitive interest rates." To incorporate preferences for modern banking technologies, we additionally included three items adapted from Bauman, Burton, and Elliott (2004), which included, "Using my bank's online banking is convenient for me," "Using my bank's mobile banking app is convenient for me," and "Using my bank's ATM is convenient for me." Participants were asked to rank the importance of each of these 20 dimensions from 1 (most important) to 20 (least important), and rankings were aggregated within each construct to facilitate our analysis.

A.3.4 Demographic measures. To capture individual-level demographics, for themselves and their households, participants were asked their age, the ages of the youngest and oldest people in their households, their monthly household income, the number of people living in their household (including themselves), and whether they owned their own home. Median age for each respondent's household was imputed by calculating the median between the oldest and the youngest person living in the respondent's household. Household income was annualized, and a per-capita income estimate was imputed by dividing annualized household income value by the number of people living in the participant's household.

A.3.5 Empirical approach. To test whether zip-code level demographic data from the ACS are predictive of respondent-level demographics from a particular zip code, we modelled participant-level responses to the continuously-varying demographic questions presented in the survey as a function of corresponding ACS zip code-level estimates. To test whether respondent-level demographics are predictive of differences in respondent-level needs in banking, we use respondent-level measures described above that correspond with the aggregated measures used in Studies 2 and 3 (respondent age, household income, household size, whether they own their home, and the population density in their zip code) to conduct an Analysis of Variance on our measures of how respondents used various channels and products, and their aggregated rankings of various service attribute dimensions. We estimate models with robust standard errors, clustered by zip code.

	(1) Age	(2) Household median age	(3) Age	(4) Household income	(5) Per capita income	(6) Household size	(7) Pr(Ownership)
ACS median age	0.192**	0.208**					
	(0.078)	(0.102)					
ACS Pr(18-34)			-11.669***				
			(4.134)				
ACS per capita inc.				0.369***	0.245***		
				(0.082)	(0.062)		
ACS avg. household size						0.454***	
						(0.154)	
ACS Pr(Home ownership)							1.762***
							(0.510)
Constant	28.253***	24.440***	38.510***	37,151.849***	14,599.548***	1.554***	-0.783***
	(2.949)	(3.947)	(1.231)	(2,950.160)	(1,971.893)	(0.395)	(0.296)
Observations	495	488	495	495	495	495	495
Model	OLS	OLS	OLS	OLS	OLS	OLS	Logistic
(Pseudo) R-squared	0.010	0.008	0.010	0.048	0.061	0.017	0.019

Table A4. Regressions of respondent-level demographics regressed on zip code-level demographics from the ACS. Columns (1-6) are estimated with OLS and Column 7 is estimated with logistic regression. All models are estimated with robust standard errors, which are presented in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels respectively. The results indicate that zip-code level demographic data from the ACS are predictive of respondent-level demographics from a particular zip code.

Finally, to test whether zip code-level demographics are predictive of differences in respondent-level needs in banking, we use zip code-level measures from the ACS – the same measures used in Studies 2 and 3, including median age, proportion of the population between 18 and 34, per capita income, average household size, proportion of homeowners, and population density – to conduct an ANOVA on our measures of how respondents used various channels and products, and their aggregated rankings of various service attribute dimensions. Models are estimated with robust standard errors, clustered by zip code. This final set of tests is a replication of the approach adopted in Studies 2 and 3.

	(A) Participant-level demog.				(B) ACS zip code-level demog.			
	df	F	Prob>F	Sig.	df	F	Prob>F	Sig.
Chanels of interaction								
Face to face	F(5,457)	0.91	0.476		F(6,455)	0.93	0.473	
Online	F(5,457)	11.65	0.000	***	F(6,455)	2.91	0.009	***
ATM	F(5,457)	2.13	0.060	*	F(6,455)	2.23	0.039	**
Mobile banking	F(5,457)	9.56	0.000	***	F(6,455)	0.45	0.842	
Phone	F(5,457)	5.69	0.000	***	F(6,455)	2.86	0.010	***
Products								
Checking account	F(5,457)	2.31	0.044	**	F(6,455)	2.36	0.030	**
Savings account	F(5,457)	3.19	0.008	***	F(6,455)	2.01	0.064	*
Investment account	F(5,457)	14.15	0.000	***	F(6,455)	6.19	0.000	***
Personal loan	F(5,457)	8.90	0.000	***	F(6,455)	2.91	0.009	***
Mortgage loan	F(5,457)	19.75	0.000	***	F(6,455)	1.63	0.136	
Online billpay	F(5,457)	3.00	0.011	**	F(6,455)	0.74	0.621	
Credit card	F(5,457)	6.10	0.000	***	F(6,455)	3.95	0.001	***
Debit card	F(5,457)	1.86	0.099	*	F(6,455)	1.51	0.172	
Safe deposit box	F(5,457)	5.10	0.000	***	F(6,455)	3.54	0.002	***
Drivers of satisfaction								
Core	F(5,452)	5.01	0.000	***	F(6,450)	4.77	0.000	***
Relational	F(5,452)	2.81	0.016	**	F(6,450)	1.54	0.162	
Tangibles	F(5,452)	5.76	0.000	***	F(6,450)	1.48	0.183	
Enabling	F(5,452)	2.72	0.020	**	F(6,450)	0.43	0.858	
Competitive	F(5,452)	1.57	0.169		F(6,450)	1.85	0.087	*
Technology	F(5,452)	2.58	0.026	**	F(6,450)	0.40	0.878	

Table A5. ANOVA results provide evidence that participant and zip code-level demographic variables are predictive of respondent-level differences in channels of interaction, product usage, and drivers of satisfaction. Model F-statistics, p-values, and significance levels are shown. Models in both panels (A) and (B) are estimated with robust standard errors, clustered by zip code. *, **, and ***, denote significance at the 10%, 5%, and 1% levels respectively.

A.3.6 Analysis and Results. **Table A4** provides evidence that supports the idea that zip-code level demographic data from the ACS are predictive of respondent-level demographics from a particular zip code. Zip code level median age as reported in the ACS is predictive of respondent age (coefficient = 0.192, $p<0.05$), as well as household-level median age (coefficient = 0.208, $p<0.05$). The proportion of the population aged 18-34, a younger demographic, is negatively associated with respondent age (coefficient = -11.669, $p<0.01$). Per capita income in the ACS is positively associated with both annual household income (coefficient = 0.369, $p<0.01$) and imputed per capita income (coefficient = 0.245, $p<0.01$). Zip code-level average household size is predictive of respondent-level household size (coefficient = 0.454, p<0.01). Finally, the proportion of the population that owns their own home, as reported in the ACS is predictive of the probability of respondent-level home ownership (logistic coefficient = 1.762, p<0.01).

Table A5 provides evidence that respondent and zip code-level demographic variables are predictive of respondent-level differences in channels of interaction, product usage, and drivers of satisfaction. Panel (A) demonstrates that participant level demographics are jointly predictive of customer-level differences across most behavioral and perceptual categories. Collectively, these results provide evidence that converges with prior research, showing that demographic differences among customers are associated with different customer needs. Panel (B) replicates the analysis using zip code level-measures from the ACS used as proxies in Studies 2 and 3. The results are robust across many perceptual and behavioral categories, when demographic data that's been aggregated at the zip code level are used instead of respondent-level measures. These results lend credence to the empirical approach adopted in Studies 2 and 3, and provide a blueprint for similar analyses that could be done in other industries, using widely-available data on demographic differences to estimate customer divergence and customer dispersion.

A4. Robustness Tests of Customer Compatibility and Service Satisfaction Results

	N	Mean	Std. Dev.	Min	Max		N	Mean	Std. Dev.	Min	Max
Satisfaction measures						*Aggregated dispersion*	145,761	3.12	0.67	1.76	7.11
Overall satisfaction	145,596	7.88	1.88	1.00	10.00	CV population density	145,761	1.71	0.63	0.66	5.97
Product offerings	145,170	7.36	1.67	1.00	10.00	CV per capita income	145,761	0.45	0.08	0.18	0.73
Operating processes	145,517	7.80	1.59	1.00	10.00	CV median age	145,761	0.16	0.02	0.07	0.26
People	124,961	8.02	1.80	1.00	10.00	CV average household size	145,761	0.15	0.03	0.06	0.27
Interaction design	145,519	7.58	1.62	1.00	10.00	CV proportion 18-34	145,761	0.38	0.04	0.19	0.65
						CV proportion owned home	145,761	0.28	0.06	0.10	0.61
Aggregated divergence	145,761	4.01	2.63	0.27	166.41						
ND population density	145,761	0.60	1.33	0.00	146.72	Additional controls					
ND per capita income	145,761	0.62	0.55	0.00	15.13	Pr(Geographically close)	145,761	0.51	0.50	0.00	1.00
ND median age	145,761	0.71	0.64	0.00	8.75	Banking instituions in zip code	145,761	6.51	4.35	0.00	41.00
ND average household size	145,761	0.68	0.60	0.00	19.05	Institution branch count	145,761	2,145.90	2,168.80	12.00	6,779.00
ND proportion 18-34	145,761	0.63	0.68	0.00	11.11	Ln(1+Institution total deposits)	145,761	14.59	5.99	0.00	20.63
ND proportion owned home	145,761	0.77	0.56	0.00	6.26						

Table A6: Summary Statistics for the tests of customer divergence and dispersion on service satisfaction (Study 2).

A.4.1. Disaggregated measures of divergence and dispersion. We draw on prior literature in selecting the demographic measures used to calculate the divergence and dispersion metrics we use in this paper (Cortiñas et al. 2010, Gupta and Chintagunta 1994, Hoch et al. 1995, Kalyanam and Putler 1997, Levesque and McDougall 1996, Mulhern and Williams 1994). Below, we replicate the analyses from Study 2, presenting the disaggregated measures of divergence and dispersion, to identify the extent to which specific measures are more or less impactful in this retail banking context. To gain a holistic perspective on the relevance of various dimensions in this context, we separately present the tables below excluding and including zip code level fixed effects. Zip code level demographic characteristics, such as population density, per capita income, median age, average household size, the proportion of the population aged 18-34, and the proportion of the population that owned their home, which we use in the calculation of normalized divergence and dispersion, are subsumed by zip code level fixed effects, as is the general demographic divergence of particular zip codes. However, the distinct divergence between particular zip codes and institutions is not subsumed by zip code fixed effects. Moreover, the inclusion of zip code level fixed effects accounts for additional time-invariant factors that differ among zip codes, such as the competitive landscape in different markets, which has been shown to influence perceptions of service (Buell et al. 2016).

The results, which are presented in **Tables A7-A10** demonstrate the varied relevance of the demographic characteristics we leverage for the creation of the aggregated metrics, both separately and collectively, across dimensions of the service experience. In **Tables A7-A8**, normalized divergence measures of population density, per capita income, average household size, the proportion of the population aged 18-34, and the proportion of the population that owned their home are negatively associated with overall banking satisfaction. **Tables A9-A10** present the disaggregated normalized dispersion metrics, demonstrating how dispersion in per capita income, average household size, the proportion of the population aged 18-34, and the proportion of the population that owned their home is negatively associated with overall satisfaction. Interestingly, disaggregated divergence of median age has an insignificant association with overall satisfaction and dispersion of median age has a positive association with satisfaction on most dimensions – perhaps reflecting the general appeal of a banking institution that is designed to meet customer needs throughout one's lifetime. However, we note median age dispersion is negatively associated with interaction design, which is consistent with the idea that customers of different ages wish to interact in different ways with their banking institutions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Overall satisfaction	Overall satisfaction	Overall satisfaction	Overall satisfaction	Overall satisfaction	Overall satisfaction	Overall satisfaction	Product offerings	Operating processes	People	Interaction design
ND population density	-0.032***						-0.024***	-0.015***	-0.021***	-0.037***	-0.027***
	(0.008)						(0.008)	(0.006)	(0.006)	(0.008)	(0.008)
ND per capita income		-0.039***					-0.013	-0.019*	-0.017*	-0.025*	-0.008
		(0.011)					(0.010)	(0.010)	(0.010)	(0.014)	(0.008)
ND median age			0.019*				0.089***	0.052***	0.055***	0.069***	0.046***
			(0.010)				(0.012)	(0.010)	(0.010)	(0.012)	(0.010)
ND average household size				-0.026**			-0.013	-0.004	-0.006	-0.017	0.003
				(0.011)			(0.011)	(0.010)	(0.009)	(0.011)	(0.012)
ND proportion 18-34					-0.070***		-0.101***	-0.080***	-0.071***	-0.075***	-0.090***
					(0.008)		(0.011)	(0.010)	(0.009)	(0.012)	(0.010)
ND proportion owned home						-0.069***	-0.012	-0.018*	-0.012	-0.015	0.005
						(0.010)	(0.015)	(0.010)	(0.009)	(0.014)	(0.013)
Geographic closeness	-0.005	-0.006	-0.007	-0.007	-0.004	-0.007	0.001	-0.027*	-0.011	-0.029**	0.167***
	(0.017)	(0.017)	(0.018)	(0.017)	(0.017)	(0.017)	(0.016)	(0.015)	(0.015)	(0.013)	(0.012)
Banking instituions in zip	-0.002	-0.002	-0.002	-0.002	-0.002	-0.003*	-0.001	-0.003**	-0.003*	-0.005***	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Instituion branch count	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)	0.011*	0.011**	0.011*	0.011**	0.011*	0.011*	0.011*	0.013*	0.009**	0.003	0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.004)	(0.005)	(0.004)
Constant	8.165***	8.158***	8.114***	8.144***	8.176***	8.185***	8.175***	7.627***	7.906***	8.267***	7.733***
	(0.092)	(0.092)	(0.093)	(0.091)	(0.094)	(0.096)	(0.092)	(0.097)	(0.073)	(0.095)	(0.066)
Observations	145,596	145,596	145,596	145,596	145,596	145,596	145,596	145,170	145,517	124,961	145,519
Adjusted R-squared	0.024	0.024	0.024	0.024	0.024	0.024	0.025	0.020	0.026	0.030	0.034

Table A7: Disaggregated customer divergence metrics and satisfaction without zip code fixed effects (Study 2). All models include institution and year fixed effects. Robust standard errors, clustered by institution and zip code, are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Overall satisfaction	Overall satisfaction	Overall satisfaction	Overall satisfaction	Overall satisfaction	Overall satisfaction	Overall satisfaction	Product offerings	Operating processes	People	Interaction design
ND population density	-0.014**						-0.012*	-0.008	-0.012**	-0.016***	-0.017**
	(0.007)						(0.007)	(0.006)	(0.005)	(0.006)	(0.008)
ND per capita income		-0.003					0.008	-0.000	0.007	0.047	-0.056***
		(0.030)					(0.029)	(0.021)	(0.021)	(0.029)	(0.021)
ND median age			-0.013				-0.010	-0.003	-0.016	-0.021	-0.048
			(0.037)				(0.040)	(0.033)	(0.031)	(0.041)	(0.032)
ND average household size				0.023			0.029	0.030	0.048*	0.100***	-0.025
				(0.033)			(0.032)	(0.029)	(0.027)	(0.031)	(0.028)
ND proportion 18-34					-0.017		0.044	0.042	0.041	0.034	0.049
					(0.035)		(0.039)	(0.039)	(0.036)	(0.045)	(0.037)
ND proportion owned home						-0.088**	-0.088**	-0.075**	-0.060*	-0.071*	-0.096***
						(0.034)	(0.039)	(0.035)	(0.032)	(0.041)	(0.029)
Geographic closeness	0.064***	0.068***	0.067***	0.068***	0.067***	0.065***	0.064***	0.028**	0.056***	0.063***	0.261***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.015)	(0.014)	(0.014)	(0.012)	(0.014)	(0.014)
Banking instituions in zip	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	0.001	-0.005	-0.000	-0.005
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.010)	(0.008)
Instituion branch count	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)	0.013**	0.013**	0.013**	0.013**	0.013**	0.013**	0.013**	0.012*	0.009**	0.001	0.006
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	`	(0.004)	(0.006)	(0.004)
Observations	141,429	141,429	141,429	141,429	141,429	141,429	141,429	141,015	141,359	120,835	141,360
Number of zip	13,039	13,039	13,039	13,039	13,039	13,039	13,039	13,031	13,041	12,378	13,042
Adjusted R-squared	0.082	0.082	0.082	0.082	0.082	0.082	0.082	0.085	0.08	0.092	0.067

Table A8: Disaggregated customer divergence metrics and satisfaction with zip code fixed effects (Study 2). All models include zip code, institution and year fixed effects. Robust standard errors, clustered by institution and zip code, are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

	(1) Overall satisfaction	(2) Overall satisfaction	(3) Overall satisfaction	(4) Overall satisfaction	(5) Overall satisfaction	(6) Overall satisfaction	(7) Overall satisfaction	(8) Product offerings	(9) Operating processes	(10) People	(11) Interaction design
CV population density	-0.027						-0.016	-0.003	-0.006	-0.003	-0.002
	(0.023)						(0.025)	(0.018)	(0.019)	(0.024)	(0.020)
CV per capita income		-1.216***					-0.443	0.088	-0.172	-0.506*	0.123
		(0.385)					(0.336)	(0.299)	(0.254)	(0.262)	(0.413)
CV median age			3.372***				4.546***	2.648***	2.897***	5.087***	0.097
			(1.043)				(1.017)	(0.927)	(0.841)	(1.016)	(1.476)
CV average household size				-2.362*			-1.179	-0.912	-0.739	-2.036	-1.120
				(1.244)			(1.224)	(1.011)	(0.933)	(1.289)	(1.391)
CV proportion 18-34					-1.338**		-1.209***	-0.646*	-0.784**	-1.122***	-0.296
					(0.650)		(0.464)	(0.377)	(0.364)	(0.419)	(0.585)
CV proportion owned home						-2.067***	-0.863	-0.618	-0.811*	-0.859	-1.202
						(0.487)	(0.601)	(0.576)	(0.483)	(0.547)	(0.756)
Geographic closeness	-0.025	-0.031**	-0.015	-0.031**	-0.023	-0.023	-0.011	-0.041***	-0.022	-0.043***	0.160***
	(0.015)	(0.015)	(0.016)	(0.015)	(0.015)	(0.015)	(0.016)	(0.014)	(0.015)	(0.013)	(0.013)
Banking instituions in zip	-0.003	-0.001	-0.004*	-0.002	-0.003	-0.001	-0.002	-0.003**	-0.002*	-0.006***	0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Instituion branch count	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)	-0.017	-0.005	-0.019*	-0.014	-0.015	-0.009	-0.003	0.005	-0.000	-0.007	0.002
	(0.010)	(0.011)	(0.010)	(0.011)	(0.011)	(0.010)	(0.010)	(0.007)	(0.007)	(0.008)	(0.007)
Constant	8.281***	8.552***	7.785***	8.511***	8.696***	8.643***	8.401***	7.396***	8.001***	8.523***	7.885***
	(0.172)	(0.145)	(0.206)	(0.154)	(0.212)	(0.138)	(0.210)	(0.174)	(0.159)	(0.219)	(0.197)
Observations	145,596	145,596	145,596	145,596	145,596	145,596	145,596	145,170	145,517	124,961	145,519
Adjusted R-squared	0.009	0.011	0.010	0.010	0.010	0.013	0.014	0.011	0.016	0.020	0.021

Table A9: Disaggregated customer dispersion metrics and satisfaction without zip code fixed effects (Study 2). All models include institution and year fixed effects. Robust standard errors, clustered by institution and zip code, are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

	(1) Overall satisfaction	(2) Overall satisfaction	(3) Overall satisfaction	(4) Overall satisfaction	(5) Overall satisfaction	(6) Overall satisfaction	(7) Overall satisfaction	(8) Product offerings	(9) Operating processes	(10) People	(11) Interaction design
CV population density	-0.019						-0.009	0.007	0.002	0.013	0.016
	(0.021)						(0.026)	(0.020)	(0.021)	(0.025)	(0.017)
CV per capita income		-1.324***					-1.115***	-0.648**	-0.835***	-1.083***	-0.554
		(0.271)					(0.358)	(0.307)	(0.263)	(0.267)	(0.378)
CV median age			1.973**				2.445**	0.501	0.868	2.978***	-2.466**
			(0.980)				(0.966)	(0.876)	(0.825)	(1.016)	(1.113)
CV average household size				-1.929*			0.769	1.590*	1.248	-0.161	0.643
				(1.011)			(1.174)	(0.930)	(0.852)	(1.179)	(1.127)
CV proportion 18-34					-0.846		-0.532	-0.111	-0.274	-0.627	0.565
					(0.570)		(0.474)	(0.362)	(0.365)	(0.406)	(0.427)
CV proportion owned home						-1.710***	-0.721	-0.330	-0.541	-0.373	-1.316**
						(0.403)	(0.563)	(0.521)	(0.442)	(0.520)	(0.610)
Geographic closeness	0.094***	0.079***	0.095***	0.083***	0.093***	0.078***	0.080***	0.034**	0.066***	0.072***	0.291***
	(0.020)	(0.016)	(0.020)	(0.016)	(0.019)	(0.016)	(0.015)	(0.015)	(0.013)	(0.016)	(0.013)
Banking instituions in zip	0.003	0.003	0.003	0.003	0.003	0.003	0.002	0.003	-0.003	0.001	-0.003
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.010)	(0.008)
Instituion branch count	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)	-0.012	-0.002	-0.013	-0.010	-0.010	-0.007	-0.001	0.003	-0.000	-0.009	0.001
	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.009)	(0.008)	(0.006)	(0.005)	(0.006)	(0.006)
Observations	141,429	141,429	141,429	141,429	141,429	141,429	141,429	141,015	141,359	120,835	141,360
Number of zip	13,039	13,039	13,039	13,039	13,039	13,039	13,039	13,031	13,041	12,378	13,042
Adjusted R-squared	0.093	0.091	0.093	0.093	0.093	0.092	0.091	0.092	0.087	0.100	0.078

Table A10: Disaggregated customer dispersion metrics and satisfaction with zip code fixed effects (Study 2). All models include zip code, institution and year fixed effects. Robust standard errors, clustered by institution and zip code, are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

A.4.2. Alternative aggregation techniques for divergence and dispersion. We present an additive aggregation technique to aggregate the six normalized measures of customer divergence and customer dispersion presented in the main body of the paper, as shown in (1) and (2) below:

$$DIV_{ijt} = \sum_{c=1}^{6} ND_{c_{ijt}} \tag{1}$$

$$DIS_{jt} = \sum_{c=1}^{6} CV_{c_{jt}} \tag{2}$$

The rationale underlying an additive strategy is that measures of divergence and dispersion may be linear in nature, building upon one another in a continuous and independent fashion. However, prior literature has also demonstrated the efficacy of a multiplicative aggregation strategy (Campbell et al. 2009). Multiplicative models, by virtue of their interactive nature, assume that the impact of additional divergence or dispersion on performance is interactive and co-determined. Difference on one dimension is more impactful when differences on other dimensions are large as well. In this way, the multiplicative aggregation strategy is more sensitive to extremely high or extremely low levels of divergence and dispersion.

$$DIV_{ijt} = \prod_{c=1}^{6} \left(ND_{c_{ijt}} \right) \tag{3}$$

$$DIS_{jt} = \prod_{c=1}^{6} \left(CV_{c_{jt}} \right) \tag{4}$$

Since the normalized divergence measures and coefficients of variation of particular demographic characteristics are often fractional in this study, in Models (5) and (6), we diverge from Campbell, Datar, and Sandino (2009) by adding 1 to each separate measure of divergence and dispersion, to ensure continuous growth of the aggregated metrics when new dimensions of variance are incorporated:

$$DIV_{ijt} = \prod_{c=1}^{6} \left(1 + ND_{c_{ijt}} \right) \tag{5}$$

$$DIS_{jt} = \prod_{c=1}^{6} \left(1 + CV_{c_{jt}} \right) \tag{6}$$

In **Table A11** below, we replicate the overall satisfaction analyses from Study 2, leveraging the additive and multiplicative aggregation techniques side by side. The results demonstrate that both strategies show consistent effects, and the similarity in adjusted R-squared measures across models provides evidence that the two techniques are similarly efficacious in the retail banking context. Although we rely on the additive technique as our primary strategy in this manuscript, the comparable efficacy of multiplicative models suggests that it may be worthwhile for scholars and practitioners replicating these techniques in other industries to consider both approaches, so as to determine which best fits their focal context.

	(1) Overall satisfaction	(2) Overall satisfaction	(3) Overall satisfaction	(4) Overall satisfaction	(5) Overall satisfaction	(6) Overall satisfaction
Additive divergence (Model 1)	-0.011**					
	(0.005)					
Multiplicative divergence (Model 3)		-5.81e-06**				
		(2.57e-06)				
Multiplicative divergence (Model 5)			-3.32e-06**			
			(1.66e-06)			
Additive dispersion (Model 2)				-0.408**		
				(0.172)		
Additive dispersion² (Model 2)				0.047**		
				(0.021)		
Multiplicative dispersion (Model 4)					-129.012***	
					(43.485)	
Multiplicative dispersion² (Model 4)					10,916.935	
					(8,263.836)	
Multiplicative dispersion (Model 6)						-0.113**
						(0.048)
Multiplicative dispersion² (Model 6)						0.004**
						(0.002)
Geographic closeness	0.063***	0.068***	0.067***	0.091***	0.080***	0.088***
	(0.014)	(0.014)	(0.014)	(0.018)	(0.017)	(0.017)
Banking instituions in zip	-0.001	-0.001	-0.001	0.003	0.003	0.003
	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)
Instituion branch count	-0.000	-0.000	-0.000	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)	0.013**	0.013**	0.013**	-0.006	-0.004	-0.004
	(0.006)	(0.006)	(0.006)	(0.010)	(0.009)	(0.010)
2007 indicator	-0.062**	-0.060**	-0.060**	-0.016	-0.015	-0.015
	(0.026)	(0.026)	(0.026)	(0.029)	(0.030)	(0.029)
2008 indicator	0.050	0.052	0.052	0.107**	0.104**	0.107**
	(0.042)	(0.043)	(0.043)	(0.045)	(0.046)	(0.045)
2009 indicator	0.116*	0.118*	0.118*	0.195***	0.185***	0.194***
	(0.066)	(0.066)	(0.066)	(0.056)	(0.058)	(0.057)
2010 indicator	0.198*	0.202*	0.201*	0.048	0.063	0.075
	(0.109)	(0.109)	(0.109)	(0.114)	(0.104)	(0.114)
Observations	141,429	141,429	141,429	141,429	141,429	141,429
Number of zip	13,039	13,039	13,039	13,039	13,039	13,039
Adjusted R-squared	0.082	0.082	0.082	0.093	0.092	0.092

Table A11: Customer divergence and dispersion analyses using different aggregation techniques (Study 2). All columns include zip code fixed effects. Columns 1-3 additionally include institution fixed effects. Robust standard errors, clustered by institution and zip code, are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively. The results demonstrate consistency across aggregation techniques.

A.4.3. Alternative clustering techniques. In the primary analyses presented in this paper, we leverage multi-layer clustering, to account for the fact that there may be correlation in the error terms among customers living in the same zip code, or among customers who select the same institution (Cameron 2011). The results below demonstrate that the primary results are robust when standard errors are instead clustered solely at the institution or zip code level. Estimating the regressions with these alternative clustering techniques allow for the recovery of the intercept, which cannot be directly estimated with multi-layer clustering in a fixed effects model.

	(1) Overall satisfaction	(2) Overall satisfaction	(3) Overall satisfaction	(4) Product offerings	(5) Operating processes	(6) People	(7) Interaction design
Customer divergence	-0.014***	-0.012**	-0.011**	-0.008**	-0.010**	-0.008**	-0.023***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.007)
Geographic closeness		0.065***	0.063***	0.028**	0.055***	0.061***	0.262***
		(0.014)	(0.013)	(0.014)	(0.011)	(0.013)	(0.013)
Banking instituions in zip		0.006	-0.001	0.000	-0.005	-0.000	-0.005
		(0.008)	(0.007)	(0.007)	(0.007)	(0.010)	(0.008)
Instituion branch count		0.000	-0.000	-0.000	0.000	0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)		0.003	0.013*	0.012	0.009**	0.001	0.006
		(0.002)	(0.007)	(0.007)	(0.005)	(0.006)	(0.005)
Constant	8.357***	8.247***	8.097***	7.520***	7.856***	8.099***	7.815***
	(0.074)	(0.093)	(0.134)	(0.131)	(0.114)	(0.134)	(0.104)
Observations	145,596	145,596	145,596	145,170	145,517	124,961	145,519
Adjusted R-squared	0.054	0.054	0.055	0.056	0.061	0.062	0.076

Table A12: Customer divergence and satisfaction with standard errors clustered by institution (Study 2). All models include institution and zip code-level fixed effects, as well as year indicator variables, which are not displayed. Robust standard errors, clustered by institution, are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

	(1) Overall satisfaction	(2) Overall satisfaction	(3) Overall satisfaction	(4) Product offerings	(5) Operating processes	(6) People	(7) Interaction design
Customer divergence	-0.014***	-0.012**	-0.011**	-0.008*	-0.010**	-0.008	-0.023***
	(0.006)	(0.005)	(0.005)	(0.004)	(0.004)	(0.006)	(0.005)
Geographic closeness		0.065***	0.063***	0.028*	0.055***	0.061***	0.262***
		(0.016)	(0.016)	(0.015)	(0.014)	(0.017)	(0.015)
Banking instituions in zip		0.006	-0.001	0.000	-0.005	-0.000	-0.005
		(0.010)	(0.010)	(0.009)	(0.008)	(0.010)	(0.008)
Instituion branch count		0.000**	-0.000	-0.000*	0.000	0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)		0.003***	0.013***	0.012***	0.009**	0.001	0.006
		(0.001)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
Constant	8.357***	8.247***	8.097***	7.520***	7.856***	8.099***	7.815***
	(0.109)	(0.125)	(0.142)	(0.123)	(0.117)	(0.144)	(0.125)
Observations	145,596	145,596	145,596	145,170	145,517	124,961	145,519
Adjusted R-squared	0.054	0.054	0.055	0.056	0.061	0.062	0.076

Table A13: Customer divergence and service satisfaction with standard errors clustered by zip code (Study 2). All models include institution and zip code-level fixed effects, as well as year indicator variables, which are not displayed. Robust standard errors, clustered by zip code, are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

	(1) Overall satisfaction	(2) Overall satisfaction	(3) Overall satisfaction	(4) Product offerings	(5) Operating processes	(6) People	(7) Interaction design
Customer dispersion	-0.922***	-0.436**	-0.408**	-0.161	-0.257**	-0.405**	-0.396**
	(0.222)	(0.175)	(0.182)	(0.112)	(0.117)	(0.175)	(0.184)
Customer dispersion²	0.109***	0.051**	0.047**	0.020	0.031**	0.050**	0.052**
	(0.027)	(0.021)	(0.022)	(0.014)	(0.015)	(0.021)	(0.023)
Geographic closeness		0.096***	0.091***	0.034**	0.071***	0.082***	0.305***
		(0.019)	(0.018)	(0.013)	(0.012)	(0.019)	(0.014)
Banking instituions in zip		0.018**	0.003	0.004	-0.002	0.002	-0.003
		(0.009)	(0.008)	(0.007)	(0.007)	(0.010)	(0.008)
Instituion branch count		-0.000***	-0.000***	-0.000**	-0.000*	-0.000**	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)		-0.002	-0.006	0.001	-0.004	-0.014**	-0.002
		(0.003)	(0.010)	(0.007)	(0.007)	(0.007)	(0.007)
Constant	9.649***	8.698***	8.733***	7.540***	8.207***	8.857***	8.070***
	(0.402)	(0.298)	(0.263)	(0.222)	(0.195)	(0.295)	(0.310)
Observations	145,596	145,596	145,596	145,170	145,517	124,961	145,519
Adjusted R-squared	0.042	0.044	0.045	0.049	0.054	0.053	0.065

Table A14: Customer dispersion and satisfaction with standard errors clustered by institution (Study 2). All models include zip code fixed effects, as well as year indicator variables, which are not displayed. Robust standard errors, clustered by institution are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

	(1) Overall satisfaction	(2) Overall satisfaction	(3) Overall satisfaction	(4) Product offerings	(5) Operating processes	(6) People	(7) Interaction design
Customer dispersion	-0.922***	-0.436**	-0.408**	-0.161	-0.257**	-0.405**	-0.396**
	(0.222)	(0.175)	(0.182)	(0.112)	(0.117)	(0.175)	(0.184)
Customer dispersion²	0.109***	0.051**	0.047**	0.020	0.031**	0.050**	0.052**
	(0.027)	(0.021)	(0.022)	(0.014)	(0.015)	(0.021)	(0.023)
Geographic closeness		0.096***	0.091***	0.034**	0.071***	0.082***	0.305***
		(0.019)	(0.018)	(0.013)	(0.012)	(0.019)	(0.014)
Banking instituions in zip		0.018**	0.003	0.004	-0.002	0.002	-0.003
		(0.009)	(0.008)	(0.007)	(0.007)	(0.010)	(0.008)
Instituion branch count		-0.000***	-0.000***	-0.000**	-0.000*	-0.000**	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)		-0.002	-0.006	0.001	-0.004	-0.014**	-0.002
		(0.003)	(0.010)	(0.007)	(0.007)	(0.007)	(0.007)
Constant	9.649***	8.698***	8.733***	7.540***	8.207***	8.857***	8.070***
	(0.402)	(0.298)	(0.263)	(0.222)	(0.195)	(0.295)	(0.310)
Observations	145,596	145,596	145,596	145,170	145,517	124,961	145,519
Adjusted R-squared	0.042	0.044	0.045	0.049	0.054	0.053	0.065

Table A15: Customer dispersion and satisfaction with standard errors clustered by zip code (Study 2). All models include zip code fixed effects, as well as year indicator variables, which are not displayed. Robust standard errors, clustered by zip code are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

A.4.4. Collapsing to the zip code/institution/year level. As an additional robustness test, we collapse our analyses in Study 2 to the zip code/institution/year level. The results, which are presented in **Tables A16-A17**, are substantively similar, despite a significant loss of statistical power by losing roughly a third of the observations.

	(1) Overall satisfaction	(2) Overall satisfaction	(3) Overall satisfaction	(4) Product offerings	(5) Operating processes	(6) People	(7) Interaction design
Customer divergence	-0.014**	-0.011**	-0.011*	-0.006	-0.008*	-0.008	-0.020***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.005)	(0.006)	(0.006)
Geographic closeness		0.085***	0.085***	0.038**	0.071***	0.072***	0.291***
		(0.018)	(0.018)	(0.016)	(0.015)	(0.019)	(0.016)
Banking instituions in zip		0.004	-0.005	-0.006	-0.009	-0.004	-0.011
		(0.010)	(0.011)	(0.009)	(0.009)	(0.011)	(0.009)
Instituion branch count		0.000	-0.000**	-0.000**	-0.000	-0.000	-0.000**
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)		0.002	0.014***	0.013***	0.010**	0.003	0.007
		(0.001)	(0.005)	(0.005)	(0.004)	(0.005)	(0.004)
Constant	8.326***	8.239***	8.057***	7.502***	7.838***	8.078***	7.764***
	(0.120)	(0.134)	(0.154)	(0.133)	(0.128)	(0.158)	(0.134)
Observations	99,441	99,441	99,441	99,231	99,403	87,678	99,410
Adjusted R-squared	0.070	0.070	0.071	0.074	0.079	0.079	0.092

Table A16: Customer divergence and satisfaction with data collapsed at the institution, zip code, year level (Study 2). All models include institution and zip code-level fixed effects, as well as year indicator variables, which are not displayed. Robust standard errors, clustered by institution, are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

	(1) Overall satisfaction	(2) Overall satisfaction	(3) Overall satisfaction	(4) Product offerings	(5) Operating processes	(6) People	(7) Interaction design
Customer dispersion	-0.871***	-0.386***	-0.358***	-0.120*	-0.210***	-0.370***	-0.357***
	(0.059)	(0.069)	(0.070)	(0.063)	(0.059)	(0.071)	(0.061)
Customer dispersion²	0.104***	0.046***	0.042***	0.015*	0.026***	0.047***	0.047***
	(0.008)	(0.009)	(0.009)	(0.008)	(0.008)	(0.009)	(0.008)
Geographic closeness		0.114***	0.112***	0.042***	0.085***	0.093***	0.330***
		(0.018)	(0.018)	(0.016)	(0.015)	(0.019)	(0.016)
Banking instituions in zip		0.015	-0.002	-0.004	-0.007	-0.002	-0.009
		(0.010)	(0.011)	(0.009)	(0.009)	(0.011)	(0.009)
Instituion branch count		-0.000***	-0.000***	-0.000***	-0.000***	-0.000***	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)		-0.003**	-0.006	0.001	-0.005	-0.016***	0.000
		(0.001)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
Constant	9.534***	8.624***	8.648***	7.482***	8.134***	8.834***	7.963***
	(0.105)	(0.136)	(0.144)	(0.129)	(0.124)	(0.148)	(0.127)
Observations	99,441	99,441	99,441	99,231	99,403	87,678	99,410
Adjusted R-squared	0.054	0.057	0.059	0.065	0.070	0.068	0.079

Table A17: Customer dispersion and satisfaction with data collapsed at the institution, zip code, year level (Study 2). All models include zip code fixed effects, as well as year indicator variables, which are not displayed. Robust standard errors, clustered by zip code are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

A.4.5. Piecewise linear regressions. Keying in on the observation that the relationships among customer dispersion and customer evaluations of service performance are non-linear, we conducted a piecewise linear regression analysis of the relationships among customer dispersion and satisfaction. The results reveal that although the slopes among customer dispersion and evaluations of product offerings, operating processes, and interaction design are insignificantly different from zero among the quartile of firms that have the most dispersed customer bases, a positive slope exists for overall satisfaction and satisfaction with people. . This pattern of results is intriguing, since it suggests companies deliberately targeting an "everything to everyone" strategy may, on some dimensions, outperform companies that merely lack focus. It is noteworthy however, that although the downward sloping relationship between dispersion and customer evaluations is consistent across all dependent measures with the results presented in the primary analyses, the upward sloping relationship at the highest levels of dispersion is not (**Table A18**).

	(1) Overall satisfaction	(2) Overall satisfaction	(3) Product offerings	(4) Product offerings	(5) Operating processes	(6) Operating processes	(7) People	(8) People	(9) Interaction design	(10) Interaction design
Customer dispersion	-0.196**	0.122**	-0.047	0.076	-0.112**	0.078	-0.186**	0.149**	-0.180***	0.053
	(0.083)	(0.061)	(0.047)	(0.049)	(0.052)	(0.050)	(0.083)	(0.066)	(0.068)	(0.063)
Geographic closeness	0.088***	0.041	0.033**	0.010	0.074***	0.034	0.083***	0.059	0.297***	0.263***
	(0.019)	(0.033)	(0.016)	(0.031)	(0.015)	(0.030)	(0.020)	(0.040)	(0.016)	(0.034)
Banking institutions in zip code	0.000	0.002	-0.000	0.009	-0.009	0.009	-0.005	-0.001	-0.006	0.023
	(0.009)	(0.017)	(0.010)	(0.017)	(0.009)	(0.018)	(0.012)	(0.023)	(0.009)	(0.018)
Institution branch count	-0.000***	-0.000**	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000**	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Institution total deposits)	-0.007	0.032	-0.001	0.054**	-0.005	0.023	-0.017**	0.020	-0.005	0.021
	(0.010)	(0.035)	(0.008)	(0.022)	(0.007)	(0.021)	(0.007)	(0.035)	(0.006)	(0.020)
2007 indicator	-0.022	0.076	-0.169***	-0.049	-0.094***	0.040	-0.036	0.125	-0.210***	-0.097
	(0.032)	(0.089)	(0.025)	(0.072)	(0.028)	(0.076)	(0.034)	(0.097)	(0.032)	(0.076)
2008 indicator	0.074	0.273***	-0.125***	0.076**	-0.031	0.145***	0.063	0.262***	-0.146***	0.031
	(0.050)	(0.066)	(0.044)	(0.037)	(0.038)	(0.043)	(0.042)	(0.075)	(0.053)	(0.043)
2009 indicator	0.192***	0.215***	0.185***	0.256***	0.298***	0.348***	0.280***	0.344***	0.241***	0.345***
	(0.070)	(0.063)	(0.060)	(0.054)	(0.051)	(0.053)	(0.056)	(0.070)	(0.048)	(0.041)
2010 indicator	0.065	0.515	0.249***	0.923***	0.276***	0.644***	0.163**	0.620	0.183**	0.597***
	(0.100)	(0.405)	(0.080)	(0.247)	(0.075)	(0.225)	(0.081)	(0.396)	(0.072)	(0.204)
Observations	105,958	32,132	105,687	31,977	105,925	32,090	90,425	27,162	105,933	32,086
Dispersion range	Bottom 75%	Top 25%	Bottom 75%	Top 25%	Bottom 75%	Top 25%	Bottom 75%	Top 25%	Bottom 75%	Top 25%
Number of zip codes	11,982	6,058	11,971	6,036	11,982	6,051	11,304	5,505	11,984	6,052
Adjusted R-squared	0.119	0.227	0.116	0.227	0.112	0.225	0.129	0.248	0.105	0.217

Table A18: Piecewise linear regression analyses of the relationships among customer dispersion and satisfaction (Study 2). All models include zip code fixed effects, as well as year indicator variables, which are not displayed. Robust standard errors, clustered by institution are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

Category	Question	Scale
Bank Loyalty	Taking into account all the products and services you receive from them, how satisfied are you with [the bank] overall?	1= Not at all satisfied 2 3 4 5 = Extremely satisfied
Bank Loyalty	How likely are you to continue to do business with [the bank]?	1= Not at all likely 2 3 4 5 = Extremely likely
Bank Loyalty	How likely are you to recommend [the bank] to a friend or associate?	1= Not at all likely 2 3 4 5 = Extremely likely
Diagnostic: Sales	[The agent] asked questions related to your financial needs	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
Diagnostic: Sales	[The agent] showed a genuine interest in your financial goals	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
Email Service Score	[The agent] was helpful in responding to your requests	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
Email Service Score	[The agent] was friendly	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
Email Service Score	[The agent] answered your questions concisely	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
Email Service Score	[The agent] was professional	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
Email Service Score	[The agent] made you feel valued	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
Email Service Score	[The agent] answered your questions completely	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
IVR Diagnostic	How easy was the voice recognition system to use?	1= Not at all Easy 2 3 4 5 = Extremely Easy
IVR Service Score	The speed of the dialog [in the voice recognition system was appropriate]	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
IVR Service Score	The ease of obtaining specific information [in the automated phone system was appropriate]	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
IVR Service Score	The privacy and security [of the automated phone system was appropriate]	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
IVR Service Score	The choice of services offered [in the automated phone system was appropriate]	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
IVR Service Score	The number of menus needed to go through to complete your inquiry or transaction [in the automated phone system was appropriate]	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
IVR Service Score	The voice [of the automated phone system was appropriate]	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
Service Score	[The agent] was knowledgeable about the bank's products and services	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
Service Score	[The agent] made it easy to do business with [the bank]	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
Service Score	[The agent] went out of their way to please you	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
Service Score	[The agent] made you feel like they wanted your business	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
Service Score	[The agent] treated you as a valued customer	1= Strongly Disagree 2 3 4 5 = Strongly Agree 0 = Does Not Apply
Phone Satisfaction	(Added starting January 2016) Overall, how satisfied were you with [the bank's] automated phone system?	1= Not at all satisfied 2 3 4 5 = Extremely satisfied 0 = Does not apply
Call/ Email Satisfaction	[The call/ email] showed a genuine interest in your financial goals	1= Not at all satisfied 2 3 4 5 = Extremely satisfied

Table A19. Questions that comprised the customer engagement score metric used in Section 2.2.5. Scores were asked as appropriate to the interaction, the average responses to scores was taken after dropping responses where customers answered that a particular question "does not apply."

A5. Alternate Specifications for Customer Compatibility and Firm Performance

To test the robustness of the effects documented in Study 3, in the tables below, we replicate the primary analyses with alternate specifications, since in some cases, the same institution had multiple branches in the same zip code in particular years. In the first table below, we replicate the main results presented from Study 3 with robust standard errors clustered at the zip code, rather than branch level. In the second, in cases where the same institution had multiple branches in a single zip code in the same year, we randomly selected a single branch, and re-ran the primary analysis on this subsample of data. In both cases, we see that the results are highly consistent with those presented in the primary specifications.

	(1) ΔLn(Deposits)	(2) ΔLn(Deposits)	(3) ΔLn(Deposits)	(4) ΔLn(Deposits)	(5) ΔLn(Deposits)	(6) ΔLn(Deposits)
Customer divergence	-0.006***	-0.006***	-0.006***	-0.002***		-0.013***
	(0.000)	(0.001)	(0.000)	(0.000)		(0.002)
New branch indicator			-0.011***			
			(0.001)			
Acquired branch indicator			-0.031***			
			(0.001)			
Customer dispersion					-0.110***	
					(0.004)	
Customer dispersion^2					0.022***	
					(0.001)	
Divergence x dispersion						0.003***
						(0.001)
Ln(1+Lagged branch deposits)	-0.106***	-0.123***	-0.108***	-0.051***	-0.108***	-0.123***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
Branch count	0.000***	-0.000***	0.000***	0.000***	0.000**	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Total institution deposits)	0.000	0.073***	-0.000	-0.000	0.005***	0.072***
	(0.000)	(0.005)	(0.000)	(0.000)	(0.000)	(0.005)
Population density (per sq. mile)	0.000	-0.000	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Per capita income	0.000**	-0.000	0.000***	-0.000***	0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Median age	0.001***	0.001	0.001***	0.002***	0.001**	0.001
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)
Average household size	0.023***	0.006	0.024***	0.029***	0.019***	0.005
	(0.005)	(0.008)	(0.005)	(0.004)	(0.005)	(0.008)
Proportion of population aged 18-34	0.025	0.056	0.027	0.036	0.014	0.049
	(0.029)	(0.047)	(0.029)	(0.027)	(0.029)	(0.047)
Proportion of population owning home:	0.024	0.010	0.025	-0.007	0.021	0.010
	(0.015)	(0.026)	(0.015)	(0.014)	(0.015)	(0.026)
Constant	1.056***	0.058	1.083***	0.487***	1.117***	0.096
	(0.030)	(0.090)	(0.031)	(0.029)	(0.031)	(0.091)
Observations	941,890	477,384	941,890	464,860	941,890	477,384
Sample	All branches	Above-median branch count branches	All branches	Non-acquired, non-new branches	All branches	Above-median branch count branches
Institution fixed effects	No	Yes	No	No	No	Yes
Adjusted R-squared	0.128	0.166	0.130	0.051	0.130	0.166

Table A20: Customer divergence, customer dispersion, and financial performance with Robust Standard Errors Clustered by Zip Code (Study 3). All models include zip code-level fixed effects, and (2) and (6) include institution-level fixed effects. Year indicator variables are included in all models, but are not displayed. Robust standard errors, clustered by zip code are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

	(1) ΔLn(Deposits)	(2) ΔLn(Deposits)	(3) ΔLn(Deposits)	(4) ΔLn(Deposits)	(5) ΔLn(Deposits)	(6) ΔLn(Deposits)
Customer divergence	-0.008***	-0.007***	-0.008***	-0.003***		-0.016***
	(0.000)	(0.001)	(0.000)	(0.000)		(0.002)
New branch indicator			-0.011***			
			(0.001)			
Acquired branch indicator			-0.031***			
			(0.001)			
Customer dispersion					-0.124***	
					(0.003)	
Customer dispersion^2					0.023***	
					(0.001)	
Divergence x dispersion						0.004***
						(0.001)
Ln(1+Lagged branch deposits)	-0.111***	-0.130***	-0.112***	-0.054***	-0.113***	-0.130***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Branch count	0.000***	-0.000***	0.000***	0.000***	0.000*	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Total institution deposits)	-0.001*	0.077***	-0.001***	-0.001**	0.007***	0.075***
	(0.000)	(0.005)	(0.000)	(0.000)	(0.000)	(0.005)
Population density (per sq. mile)	0.000**	-0.000	0.000**	0.000	0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Per capita income	0.000**	-0.000	0.000**	-0.000***	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Median age	0.001***	0.001***	0.001***	0.002***	0.001***	0.001**
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)
Average household size	0.024***	0.009	0.025***	0.026***	0.020***	0.008
	(0.005)	(0.008)	(0.005)	(0.004)	(0.005)	(0.008)
Proportion of population aged 18-34	0.042	0.075*	0.044	0.031	0.031	0.069
	(0.028)	(0.042)	(0.028)	(0.027)	(0.028)	(0.042)
Proportion of population owning homes	0.021	0.002	0.022	-0.001	0.015	0.002
	(0.015)	(0.025)	(0.015)	(0.013)	(0.015)	(0.025)
Constant	1.100***	0.027	1.131***	0.536***	1.164***	0.073
	(0.030)	(0.091)	(0.030)	(0.028)	(0.030)	(0.092)
Observations	757,083	359,287	757,083	378,560	757,083	359,287
Sample	All branches	Above-median branch count branches	All branches	Non-acquired, non-new branches	All branches	Above-median branch count branches
Institution fixed effects	No	Yes	No	No	No	Yes
Adjusted R-squared	0.132	0.170	0.133	0.050	0.135	0.171

Table A21: Customer divergence, customer dispersion, and financial performance, on subsample of data presenting one randomly-selected branch of a particular institution (Study 3). All models include zip code-level fixed effects, and (2) and (6) include institution-level fixed effects. Year indicator variables are included in all models, but are not displayed. Robust standard errors, clustered by branch are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

A6. Customer Compatibility and Firm Performance as Measured by Growth in Net Loans

Study 3 presents the effects of customer divergence and customer dispersion on deposit growth. However, another important component of a bank's financial performance is growth of its loan portfolio. In the United States, banks are required to regularly report net loan balances at the institution level to the FDIC. Since data on net loans are not available at the branch level, we are not able to fully replicate the analysis from Study 3.

However, we are able to test whether divergence between the capabilities of an operation and the needs of the customers it serves is negatively associated with net loan growth (which is consistent with Hypothesis 4) and whether dispersion among the needs of customers served by an operation is negatively associated with net loan growth, but at a diminishing rate (which is consistent with Hypothesis 5).

	(1)	(2)
	ΔLn(Net Loans)	ΔLn(Net Loans)
Customer divergence	-0.005***	
	(0.001)	
Customer dispersion		-0.024***
		(0.005)
Customer dispersion²		0.007***
		(0.002)
2008 indicator	0.001	0.001
	(0.007)	(0.007)
2009 indicator	-0.049***	-0.049***
	(0.006)	(0.006)
2010 indicator	-0.103***	-0.103***
	(0.006)	(0.006)
2011 indicator	-0.142***	-0.142***
	(0.006)	(0.006)
2012 indicator	-0.112***	-0.112***
	(0.006)	(0.006)
2013 indicator	-0.105***	-0.105***
	(0.007)	(0.007)
2014 indicator	-0.070***	-0.070***
	(0.006)	(0.006)
2015 indicator	-0.057***	-0.057***
	(0.006)	(0.006)
2016 indicator	-0.057***	-0.057***
	(0.005)	(0.005)
2017 indicator	-0.064***	-0.064***
	(0.006)	(0.006)
Institution branch count	-0.000**	-0.000**
	(0.000)	(0.000)
Ln(1+Total institution deposits)	0.010***	0.010***
	(0.002)	(0.002)
Constant	0.021	0.025
	(0.020)	(0.023)
Observations	74,186	74,186
Adjusted R-squared	0.020	0.020

Table A22. Customer divergence, customer dispersion, and growth in net loans. Year indicator variables are included in all models. Robust standard errors, clustered by institution are shown in parentheses. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.

To test these hypotheses, we collected data on 8,637 federally-insured banking institutions operating in the United States from 2006-2017 – the same period during which we document growth in net deposits in Study 3. **Table A22** above provides our results. Column (1) shows that controlling for a bank's scale, (branches and total deposits), institutions with higher average branch-level customer divergence exhibit slower growth in net loans (ν=-0.005, p<0.01), which is consistent with the results in Study 3, and with Hypothesis 4. Increasing average branch-level divergence by a standard deviation reduces the institution-level annual growth in net loans by 0.92% – a 13.3% decline relative to baseline growth rates of 6.9% per year. Column (2) shows that institutions with more dispersed customer bases exhibit slower growth in net loans (ν=-0.024, p<0.01), but at a diminishing rate (ν=0.007, p<0.01), which is also consistent with the results presented in Study 3, and Hypothesis 5. Increasing an institution's level of dispersion from the 1st percentile (a very focused firm) to the 50th percentile (a less-focused firm with a median level of customer dispersion) reduces net annual loan growth by 2.1% (from 7.4% to 5.3%). Since we do not have access to branch-level loans data, and hence, do not have variation in customer divergence within particular institutions, we are not able to directly test Hypothesis 6 with these data. However, the results presented here provide corroborating evidence that customer compatibility can have a meaningful impact on a firm's financial performance.

Moreover, building on the analyses presented in A.4.5, piecewise linear regressions demonstrate that in the bottom three quartiles for customer dispersion (among more focused firms), deposit growth falls in dispersion, but in the top quartile, deposit growth increases in dispersion. Net loan growth does not exhibit a similar increase in dispersion among the least focused firms (**Table A23**).

	(1) ΔLn(Deposits)	(2) ΔLn(Deposits)	(3) ΔLn(Net Loans)	(4) ΔLn(Net Loans)
Customer dispersion	-0.038***	0.043***	-0.015***	0.001
	(0.002)	(0.005)	(0.003)	(0.008)
Ln(1+Lagged branch deposits)	-0.110***	-0.126***		
	(0.002)	(0.003)		
2008 indicator	-0.028***	-0.019***	0.003	-0.004
	(0.002)	(0.003)	(0.008)	(0.013)
2009 indicator	0.023***	0.037***	-0.044***	-0.065***
	(0.002)	(0.003)	(0.007)	(0.015)
2010 indicator	-0.027***	-0.044***	-0.098***	-0.120***
	(0.002)	(0.003)	(0.007)	(0.016)
2011 indicator	-0.045***	-0.012***	-0.141***	-0.148***
	(0.002)	(0.003)	(0.006)	(0.017)
2012 indicator	-0.030***	0.009***	-0.118***	-0.096***
	(0.002)	(0.003)	(0.007)	(0.013)
2013 indicator	-0.034***	0.028***	-0.114***	-0.082***
	(0.002)	(0.003)	(0.008)	(0.013)
2014 indicator	-0.024***	0.020***	-0.081***	-0.042***
	(0.002)	(0.003)	(0.007)	(0.013)
2015 indicator	-0.010***	0.048***	-0.067***	-0.035***
	(0.002)	(0.003)	(0.007)	(0.013)
2016 indicator	-0.001	0.052***	-0.066***	-0.036***
	(0.002)	(0.003)	(0.006)	(0.013)
2017 indicator	0.011***	0.061***	-0.069***	-0.056***
	(0.002)	(0.004)	(0.006)	(0.015)
Branch count	-0.000***	-0.000***	-0.000***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Ln(1+Total institution deposits)	0.011***	0.023***	0.015***	0.005**
	(0.001)	(0.001)	(0.004)	(0.002)
Population density (per sq. mile)	0.000**	0.000		
	(0.000)	(0.000)		
Per capita income	0.000***	-0.000*		
	(0.000)	(0.000)		
Median age	0.001*	-0.001		
	(0.000)	(0.001)		
Average household size	0.015***	0.004		
	(0.005)	(0.010)		
Proportion of population aged 18-34	0.025	-0.092		
	(0.031)	(0.064)		
Proportion of population owning hom	0.039**	-0.034		
	(0.017)	(0.038)		
Constant	1.005***	0.928***	-0.034	0.067*
	(0.033)	(0.073)	(0.041)	(0.039)
Observations	702,905	238,985	55,559	18,627
Dispersion range	Bottom 75%	Top 25%	Bottom 75%	Top 25%
Adjusted R-squared	0.133	0.163	0.020	0.019

Table A23: Piecewise linear regression analyses of the relationships among customer dispersion and deposit and net loan growth (Study 3). Columns (1-2), which are conducted at the branch level, include zip code fixed effects and robust standard errors clustered by branch. Columns (3-4), which are conducted at the institution level, do not include zip code fixed effects and are estimated with robust standard errors clustered at the institution level. *, **, and ***, denote significance at the 10%, 5%, and 1% levels, respectively.