DIVERSITY, DISCRIMINATION, AND PERFORMANCE

Jonathan S. Leonard* and David I. Levine
University of California, Berkeley

November 2006

Abstract: Employee diversity may affect business performance both as a result of customer discrimination and as a result of how members of a group work with each other in teams. We test for both channels with data from more than 800 retail stores employing over 70,000 individuals matched to Census data on the demographics of the community. We find little payoff to matching employee demographics to those of potential customers except when the customers do not speak English. Although age diversity does predict lower sales, diversity of race or gender within the workplace does not significantly or substantially affect sales or sales growth.

JEL codes: J7, K31, L23
Keywords: Discrimination, customer discrimination, diversity, organizational performance, labor law

* email: Leonard@haas.berkeley.edu

Acknowledgements: Laura Giuliano provided outstanding research assistance. We thank the employer for providing the data, BOLD, Thomas Kochan, and the U.C. Berkeley Institute of Industrial Relations for financial support. Comments from Florian Zettelmeyer and seminar participants at U.C. Berkeley and UCLA were helpful.

Decades after employment discrimination was outlawed by the Civil Rights Act of 1964, the CEO of Shoney's Inc. personally investigated one of its restaurants which suffered from lagging sales. Noticing many black employees in visible positions and many white customers, he sent a memo to the restaurant manager directing him to employ more whites up front so as to increase sales. In 1993, this attempt to accommodate the CEO's perception of customers' discriminatory preferences was part of a case that resulted in a $132 million settlement (Watkins, 1997).

In contrast to Shoney's CEO, proponents of workplace diversity have frequently claimed that demographic diversity is good for business (Cox, 1993; Bantel and Jackson, 1989). But ironically they often share the view of Shoney's CEO that customers prefer to deal with employees of the same race or sex. The difference between these two sets of advocates of accommodating customer discrimination is that Shoney's CEO saw his potential customers as white, while diversity proponents assume the customer base is more demographically diverse. If customers are diverse and many customers prefer to deal with a demographically similar salesperson, then employee diversity can increase sales.

Employee diversity may also affect sales through team effects, a second mechanism distinct from customer discrimination. Differences among employees can directly affect group performance. Proponents and opponents of diversity also differ in their view of how employees' similarity with each other affects group performance. When creativity and the presence of diverse information sources are important, diversity can improve workgroup performance (e.g., Bantel and Jackson 1989; Jehn, Northcraft, and Neale, 1999; Watson, Kumar, and Michaelson 1993). In contrast, other theories (reviewed below) emphasize how workforce diversity can reduce cohesiveness and communication among employees.

The impact of diversity on workplace performance is an open empirical question. We use longitudinal evidence from more than 800 similar business establishments within a single large employer to examine the two diversity channels. First, we examine how the demographic match between customers and employees affects performance. Second, we examine how employees' racial, ethnic, gender and age diversity affect work team performance. Our measure of workplace performance is an objective one of central importance to business: sales. We examine how changes in workplace demographics affect sales within a store over time.

One significant barrier faced by studies of diversity (or of any other dimension of human resource policy and practice) across establishments is that establishments are likely to differ in a number of unmeasured dimensions. If we could run a controlled experiment on diversity, we would want to replicate the same workplace, experimentally varying only employee demographics. Although demographics have not been randomized, the workplaces we examine are members of national chains that by design attempt to hold fixed many confounding factors that might affect process or sales. In pursuit of their brand strategy, the chains have attempted to replicate these workplaces in every significant U.S. market.

To examine employee-customer matching, we use Census data on the demographics of the community (that is, potential customers). Because we can follow stores over time and often have multiple workplaces in one community, we are also able to control for the fixed features of a community. We separately analyze Hispanics[1] and Asians who speak English versus those who do not, as employee-customer similarity can be more important when language is a potential barrier.

We identify diversity effects within teams (measured as one minus a Herfindahl index of employment shares) as a nonlinear effect of employee demographic shares. Because we examine workplaces with both female and male majorities as well as stores with both white and nonwhite majorities, we can identify diversity effects distinct from the linear effects of demographics.

In standard theory, profit-maximizing managers select demographics to have no effect on profits at the margin. In practice, we do observe that sales are affected by some measures of employee demographics, and that employee demographics change without a change in the customer base. We follow stores over time to difference out unchanging omitted store and neighborhood characteristics. Store demographics change much more rapidly and frequently than that of their surrounding neighborhoods. We can also compare otherwise similar stores in the same zip-code or even the same mall: stores with different employee demographics despite the similar customer base that close proximity drives.

We study two distinct effects of employment diversity on sales, the first reflecting customer preferences, the second a direct team output effect irrespective of customer demographic preferences. Our findings can be briefly summarized. With one exception, sales are insensitive to the demographic match between a store's customers and its employees. The exception is a benign one: sales are higher if employees speak the language of customers,

particularly Asians, who do not speak English. Team theories suggest that diversity of gender or race might reduce sales due to worse communication and cooperation among workers or raise sales due to pooling information, sparking creativity, and understanding diverse customers. Our results support neither set of hypotheses. While diversity of age does consistently predict lower sales, racial and gender diversity are generally not correlated with sales.

I. Previous Work

Theories of the impact of employee diversity on business performance fall into two broad classes. We first discuss theories that examine whether sales of a service business depend on employee demographics because customers care about the demographics of those who serve them. We then turn to theories on how diversity may affect productivity by affecting the internal dynamics of the workgroup.

A. Customer Preferences

Customers might prefer demographically similar employees because of discrimination, as in Becker's classic 1957 theory of segregation. Similarity might improve communication (Lang, 1986; Tannen, 1990, Jackson and Alvarez, 1992; Cox, 1993). In settings such as the one we study, employees can also attract customers using connections within the community (Cox, 1993; Ibarra, 1992, 1995)[2].

Communication costs grow when a large number of potential customers do not speak English well. Although most immigrants learn English rapidly (Friedman and DiTomaso, 1996), large immigrant enclaves in many cities contain a substantial number of people who cannot or prefer not to speak English. These motivations can all lead profit-maximizing employers to desire a workforce that is demographically similar to its customers. Costly search for customers leads to the hypothesis that sales are higher when workforce demographics are similar to customer demographics, notwithstanding the legal risk incurred by discriminating in employment.

1. An Extension

The standard economic model of discrimination (Becker, 1957) does not distinguish between liking whites and disliking blacks: preferences are relative and the effects of similarity should be broadly proportional to the match of customers and employees. We extend this standard model to propose a method of distinguishing positive from negative discrimination. With negative discrimination customers of one race avoid stores with employees of other races

(no matter how few). If negative discrimination against blacks is prevalent, employing even a small number of blacks would reduce sales, and stores would tend to segregate.

In contrast, with "positive discrimination" customers are attracted to stores with at least a few employees of their own race (no matter how many). A customer who speaks only Spanish would prefer at least one employee in the store who speaks Spanish. There are diminishing returns to having multiple Spanish-speaking salespeople. When customers have positive discrimination, stores maximize profits by having a few employees of every race. If these cases are common, we should see sales increasing as a concave function as each nonwhite race's share rises above zero.

2. Evidence that Customers Prefer Similar Employees

Diversity proponents are ironically united with segregationists in promoting race conscious policies based on the premise that customers prefer dealing with like employees, so they propose that employers should judge employees on the color of their skin rather than the content of their character. While a number of studies find employers trying to match employee to customer demographics, the evidence is far more limited that customers themselves have strong preferences. Newly hired low-wage workers who have direct contact with customers are more likely to match the demographics of those customers than are new hires who have no customer contact (Holzer and Ihlanfeldt, 1988). Moreover, about 20 percent of urban low-wage employers feel their customers dislike black service providers, and such employers are much less likely to hire black men (controlling for the racial mix of applicants and of customers (Moss and Tilly 2001: 146-7; Holzer 1999). In contrast, Raphael, Stoll and Holzer (2000) find that the probability that blacks experience hiring discrimination is not greater in the (whiter) suburbs than in central-cities. More broadly, employers as different as federal agencies (Borjas, 1982) and restaurants (Neumark, 1996) have been shown to hire workforces that approximate that of their clients.

The record of litigation is replete with cases charging, and often finding, employment discrimination in the retail sector.[3] This record makes it difficult to argue that—because of routinized work procedures, limited employee discretion, pervasive brand image, or comfort with women or minorities in service roles—neither employers nor customers exhibit any preferences for the race or sex of retail employees.

Although the lawsuits make clear that some employers act as if customers discriminate, it is rare that the academic literature directly investigates the impact of customer discrimination on sales. One particularly compelling exception, related to ours in concept, studies customer discrimination in professional sports. White basketball players attract more fans than do black players of similar quality. (Kahn and Sherer, 1988). In a study of tipping by taxi customers, Ayres et. al. (2005) finds that within the same locale, black taxi-drivers are paid less in tips than are white drivers. It is not clear how far these results generalize. A few small-scale studies in the marketing literature offer a mixture of results with no clear pattern that sales are higher when customer and employee demographics are similar (e.g., contrast Churchill, Collins, and Strang [1975] with Dwyer, Richard, and Shepherd [1998]).

These cases make two important points. First, in the retail and service sectors there is mixed evidence of customer discrimination. Second, employers often act as if customers have this preference.

B. Team Effects

Even if diversity does not affect business performance through customer preferences, it may directly affect productivity by affecting how employees work with each other in teams. Both the theory and evidence on how employees' similarity with each other affects performance in this second channel show mixed results.

Theories of diversity emphasize that diversity can have both positive and negative effects depending on how demographic diversity affects collaboration within workgroups, the firm's "connective capital" (Ichniowski, Shaw and Gant, 2003). Diverse teams may help performance because they are more likely to have the information needed to solve any given problem (Lazear, 1998), come up with more creative solutions than do homogeneous groups (Thomas and Ely, 1996; Nemeth, 1985), and have employees with insights into the needs of customers (Thomas and Ely, 1996). In contrast, diversity may increase the costs of communication within the workforce (Lang, 1986; Zenger and Lawrence, 1989), lower group cohesiveness (Pfeffer, 1983), increase employee turnover (O'Reilly et al., 1989; Jackson et al., 1991), and reduce incentives for cooperation (Greif, 1993, Kandel and Lazear, 1992).

The few empirical studies are also mixed. Hellerstein and Nemark (2004) find that the segregation of Hispanic employees appears to be a result of crowding, not due to productivity-enhancing segregation by language abilities. In contrast, a recent study of garment

manufacturing similar in spirit to ours (Hamilton, Nickerson and Owan, 2004) reports some evidence that all-Hispanic teams are more productive than integrated teams, although they specifically caution the results are "non-robust." In a different setting and century, Costa and Kahn (2002) find less shirking during the Civil War among military units that were more homogeneous in ethnicity and age. Distinct from these studies of diversity itself, a set of empirical papers has documented systematic sex and race differences in productivity (Bayard, Hellerstein, Neumark and Troske, 2003; Hellerstein, Neumark and Troske, 1999; Holzer and Neumark, 1999). At a much higher level of analysis, a parallel literature has shown that ethnic fragmentation within a nation predicts low economic growth (Collier and Gunning, 1999; Rodrik, 1998). Given the contradictory theories and the mixed evidence surrounding diversity's effects, it is useful to examine directly how diversity affects establishment performance.

II. Data and Methods

We examine over 800 workplaces and over 70,000 employees of a single large service-sector employer. This figure is roughly the total number of natural work groups in *all* the field studies reviewed by Williams and O'Reilly (1998). Studies of employment are bedeviled by unmeasured differences in policies, practices, and working conditions across different employers. To test the effect of employment demographics on performance, an ideal experiment would randomly vary demographics while holding all other possibly confounding factors fixed. As in all workplace studies to date, the employer did not allow us to randomize employee demographics. This limitation pervades the entire literature on companies. But our design does minimize unmeasured differences across workplaces. We exploit the fact that as a matter of corporate policy, this employer, like many national chains, actively pursues uniformity in product and process across its outlets. Store demographics do change over time within store at a much shorter frequency and to a more dramatic extent than population changes. We focus on the impact of these changes in diversity within store over time.

Retail chains, as a matter of policy, seek to reduce heterogeneity across locations. This employer has purposefully attempted to replicate the same outlet characteristics in every U.S. market of significance, as is common among national chains that promote a brand image. In most field studies, demographics are correlated with other features of the workplace or job. The workplaces in our study, however, exhibit little of this variation. Each workplace has minimal local discretion, as each must implement the detailed human resource policies disseminated from

corporate headquarters. Wages, occupational structure, internal hierarchy, fringe benefits, and job content are for the most part centrally set and uniformly implemented. Wages and prices do not vary meaningfully with the demographics of the workforce or community. Corporate uniformity extends well beyond HR policy. Advertising, product selection, and pricing, are all centrally determined to promote uniformity. The employer's goal is that customers and employees perceive workplaces in different locations as essentially interchangeable. This standardization limits possible confounds between demographics and omitted job, product, or establishment characteristics.

We use specifications designed to capture fixed features, measured or not, of the workplace, labor market, and customers. These location-specific factors may affect both demographics and sales. For example, inner-city establishments may have both low sales and a high percentage of minority employees without any direct causal link. Retailer's entry and exit decisions should limit the impact of any such factors with a predictable effect on profits. We also include a community fixed effect when examining changes in sales to difference out local labor market shocks that might affect both changes in demographics and changes in sales.

A. The Setting

The employer is in an industry characterized by numerous small outlets that sell somewhat differentiated products. Each workplace we study is company owned and operated, and typically employs 15 to 40 part-time employees with several full-time managers and assistant managers. Because employees work scattered shifts through the week, they work with a changing mix of the other employees. Most frontline employees rotate through the several tasks in the store, spending some of their time dealing with customers and other time in support tasks.

Nonmanagerial employees receive minimal training when they are hired. These employees interact with each other to maintain stock and service customers, but these interactions are not complex. The Taylorist production techniques, with highly centralized decision making and limited local discretion, may limit the potential impact of any employee differences on productivity. Managers receive some training in managing a diverse workforce.

Production is not so standardized that diversity could not matter. Empirically, total factor productivity (that is, sales adjusted for employees, size in square feet, and the many observable characteristics of the workplace and community listed in Table 1) varies substantially across

workplaces. Overall, a fourth of the variation in sales remains even after adjusting for all observable features of the workplace and community. Not all of this variation can be due to variation in location quality (which should be captured in rents). In results not shown we find that when a high-sales manager shifts to a new store, the new store has increased sales. Organizational factors can affect sales even with this company's standardization. Data and variables are discussed in detail in the appendix.

B. Specifications

We assume that the current match between a store and its community (zip code) determines the current level of sales in a store. Equation 1 presents a simple reduced-form empirical specification where the logarithm of sales at store i in community c at time t depends on store demographics ($demog_{ict}$), store characteristics (X_{ict}), community demographics ($demog_c$), community characteristics such as the distribution of household income (Z_c), and monthly dummies (*time*):

1) $$S_{ict} = a + b_0\, time + b_1 X_{ict} + b_2 Z_c + b_3\, demog_{ict} + b_4\, demog_c + b_5\, demog_{ict} \cdot demog_c + e_{ict}$$

We take advantage of the fact that many communities have multiple stores. For the theories of store-community match, the coefficient of interest is b_5, which tells us (for example) if adding more Hispanics to a store is more useful in areas with a high proportion Hispanic. If b_5 is positive, then moving from 3 to 15 percent Hispanic employees in a community that is 20 percent Hispanic will increase sales more than the same shift in employee demographics in a community with 2 percent Hispanics.[4]

The main effect on store demographics b_3 captures worker and neighborhood characteristics correlated with race. The main effects also capture customer discrimination that is shared by all customer demographic groups. Because the main effects on mean age, race and gender conflate several forces, the coefficients on the main effects are open to a variety of interpretations. These are of secondary interest here.

We present a pooled specification and its components, the within and between specifications, along with a formal test of the fixed-effects model.

The residual e_{ict} may be correlated with unobservable features of the store and community. Specifically, assume the residual includes unmeasured store characteristics that are fixed (u_i) and unmeasured community characteristics that are fixed (v_c), as well as a white noise residual ε_{ict}:

(2) $e_{ict} = u_i + v_c + \varepsilon_{ict}$.

If the persistent but unobserved determinants of a store's characteristics v_c are correlated with both sales and employee demographics, then estimates of the employee demographic coefficients in equation (1) will be biased. To the extent that the factors affecting both demographics and sales are fixed, we can first difference equation (1) to eliminate the omitted store and community characteristics (u_i and v_c):

3) $\Delta S_{ict} = b_0' + b_1'\Delta X_{ict} + b_3'\Delta demog_{ict} + b_5'\Delta demog_{ict} \cdot demog_c + \Delta\varepsilon_{ict}$.

The first difference estimator in (3) analyzes the within-store component of variance, throwing out most of the variation in store demographics. We also examine the between-store component that averages each store's sales and characteristics over the sample period, which gives greater weight to the long-term relation between community and store demographics and store sales:

4) $S_{ic} = a'' + b_1''X_{ic} + b_2''Z_c + b_3''demog_{ic} + b_4''demog_c + b_5''demog_{ic} \cdot demog_c + e_{ic}''$

Omitted community factors may change over time, affecting both workplace demographics and sales. A store that is experiencing a positive demand shock may hire from demographic groups that it normally avoids, biasing coefficient estimates. Equation (5) presents the residuals in this case:

5) $e_{ict} = u_i + v_{ct} + \varepsilon_{ict}$.

We also present within-mall, or within zip-code, estimates. The addition of finely detailed location-specific time*place interactions, exploiting the fact that many communities, indeed many ZIP codes, have multiple stores, limits remaining omitted variable bias due to local shocks. This specification corresponds to including a separate intercept for each ZIP code in the first-difference version of a two-period panel:

6) $\Delta S_{ict} = b_0 + b_1\Delta X_{ict} + b_3\Delta demog_{ict} + b_5\Delta demog_{ict} \cdot demog_c + ZIP_c + \Delta e_{ict}$.

The resulting estimates of the interaction term b_5 can be thought of as answering the following question: Consider increasing the proportion Hispanic in one store in a community but not in a nearby store. Will that addition increase relative sales of the increasingly Hispanic store more if it takes place in a highly Hispanic region of the Southwest than if it takes place in a low-Hispanic portion of the Great Plains? This specification differences out both fixed-store

characteristics and community-level shocks that might affect both store demographics and sales,[5] but at the cost of reduced precision.

To study team effects, we add measures of the level (equations 1 and 4) or change (equations 3 and 6) of workplace diversity – measured as one minus a Herfindahl index (see appendix).

Our key results will come from the specifications using time-series variation within store (equations 3 and 6). What sources of variation remain after all of this differencing? These workplaces hire roughly three entire workforces a year, as is standard in entry-level jobs in this sector. Fluctuations in who walks in the door will provide substantial variation in employment that is reasonably exogenous to sales. (In related research we examine in more depth how the race of managers affects the hiring and retention of workers of different races.) Perfect matching of employees to discriminating customers predicts a zero impact on marginal profits. But this is of limited relevance in our within-store specifications because store demographics change much more frequently, rapidly, and substantially than do community demographics. Given the relatively slow changes in population, any claim that one month's store demographics perfectly matches its community is immediately undercut by the next month's change in store demographics. In addition, some of these monthly employment shifts reflect exogenous supply shifts: they follow the seasonal pattern of school vacations.

Finally, because of the strong advantage that may arise from speaking a foreign language when customers do not speak English, we test whether the presence of Hispanic (Asian) employees predicts higher sales when many nearby residents do not speak English, We examine the impact of the share of Hispanic employees interacted with the share of nearby residents who speak Spanish but not English. We also examine the share of Asian employees interacted with the share of residents who speak Asian-Pacific languages but not English. Our estimates will understate the benefits of employees who speak the language of linguistically isolated customers to the extent employees who self-identify as Hispanic do not speak Spanish. Similarly, Asian employees who speak an Asian language may not share a common language with all non-English-speaking immigrants from Asia in the community.

III. Results

A. Summary Statistics

Table 1 presents summary statistics. The employer hires a diverse workforce. This employment pattern arises partly because the employer has a reputation for gender and race diversity in its marketing and employment. In addition, in our interviews, managers noted that they hire many employees from among the ranks of customers. A diverse customer base contributes to, but does not fully determine, a diverse workplace.

The gender diversity index in our sample spans the entire possible range from zero (all female) to one-half (an even mix of men and women), with a mean of .34. An increase in gender diversity is not necessarily the same as an increasing proportion of women. As the proportion female rises above half, gender diversity falls. The proportion of women in the stores ranges from 6 percent to 100 percent with a mean of 75 percent. The racial diversity index ranges from zero to .79, with a mean of .39. These are entry-level jobs; the stores are more black, more Hispanic, more Asian, more female, and younger than their communities. The mean age of employees in our data is only 24 years. As this is not a sector or a firm in which most employees stay to build a career, most employees fall within a fairly narrow range of ages. The mean of the within-store standard deviation of the logarithm of ages is only 27 percent. We next consider two mechanisms through which diversity may affect performance: employee-customer matching and team effects. In both cases we present within-store tests (to which we give more weight) (Table 3) as well as tests using pooled and across-store data (Table 4). Table 2 provides an overview of our results.

B. Employee-Customer Match

Matching employees to the surrounding community's race does not increase sales. Between-store regressions (to be presented in Table 4) are subject to omitted variable bias from unmeasured locational factors that may affect both sales and demographics. Although we control for income, unemployment, population density, retail density, and other community factors, a Hausman test strongly supports the importance of store fixed effects.[6] Within-store estimates that difference out the remaining omitted unchanging factors, as in equation (3), are preferable even though there is less testable variation in the time-series. Table 3 shows the within-store results. Column 1 presents estimates using year-on-year changes in the logarithm of real sales. [7]

We also present a within–mall or zip-code estimate that differences out omitted local shocks that change over time even within a very small geographic area. Consider two stores in the same neighborhood. Whatever omitted forces affect product demand or labor supply in one such store are likely to affect the other store as well. We difference out demand or supply shocks common to such "brother" stores. In Table 3, columns 3 and 4, we add controls for community fixed effects based on ZIP codes, as in equation (6). Two levels of differencing are applied: differencing within stores across time and comparing across stores sharing a ZIP code. When one store in a community changes employment demographics to better match community demographics, do its sales increase in comparison to a neighboring store that does not similarly adjust demographics? The location fixed effects capture any local change in community income, taste, or demographics. The cost of this more rigorous procedure is that it reduces the number of stores and ignores all variation in sales that is persistent across malls or communities. A Hausman test strongly supports the importance of the ZIP code fixed effects in the regression of the rate of change of sales.

Our first key result here is how little diversity seems to matter. In communities with few blacks, increasing the store's black share has a modest but insignificant positive effect on sales. Surprisingly, raising a store's percent black reduces sales slightly in highly black communities, but only when controlling for the ZIP code fixed effects (Table 3, column 3). This result suggests that the patterns we observe are not simply due to potential white customers discriminating against blacks. Matching is similarly insignificant for women and Hispanics. Increasing the percent Asian has no effect on sales in most communities, but the effect is negative in highly Asian communities (column 1). This effect loses statistical significance with ZIP code fixed effects (column 3). As the next section shows, employee-customer matching generally remains unimportant even when the cross-section component of variation is included.

1. Results in Pooled Time-Series Cross-Section Data

Table 4 present results in pooled cross-section time-series data (eq. 1,2,3) and in the cross-section (eq. 4,5). This data is rich enough to detect demographic effects on sales, even in what some might consider a sector insensitive to individual ability. We can observe that a store's race and age distributions do help predict sales. Sales are significantly lower in stores with greater proportions of black employees. Since wages differ only negligibly by race, the standard first-order condition for profit maximization relating relative marginal sales to relative

wages is violated (ignoring potential liability costs). Under depressed economic conditions, white men tend to bump down into this sector, which works against finding negative effects for both female and minority employees. A 10 percentage point increase in black employment share (at the expense of the baseline group of whites) is associated with .8 percent lower sales. The same increase in Asian employment share is associated with .6 percent greater sales. The Hispanic employment share does not significantly predict which stores have high sales. The workforce's average age predicts slightly higher sales, a result consistent with spillover effects of general human capital. Many of these results are sensitive to the alternative specifications discussed below.

Estimates in the pooled cross-section time-series data again show the general insignificance of employee-customer matching effects previously shown in within-store estimates. The specification in table 4, column 2, corresponds to equation (1).[8] The effects of the interaction of store and community race are mixed, providing no consistent support for theories of customer preference. Specifically, the coefficient on (Store %Asian)*(Community %Asian) is a small negative number (contrary to theory), while the corresponding interactions for black and Hispanic are small and positive; none are statistically significant. The signs of these interactions are not stable across specifications.

Unlike race, the proportion female is similar in almost every community in the United States. To avoid multicollinearity, we use the gap between store and community percent female (instead of their interaction) and contrast stores in the top and bottom quartile of this distribution with those in the middle. Stores in the bottom quartile of "store percent female minus community percent female" have 1.2 percent higher sales than stores in the middle two quartiles. Working against the importance of this result is that stores with the top quartile of "store percent female minus community percent female" have 0.3 percent higher sales than stores in the middle two quartiles. The cross-section results in Table 4, column 4, correspond with equation (4). Matching a community's race or gender composition has no statistically significant effect on sales.

We see neither significant nor substantial evidence that matching employment shares to population shares in the surrounding community matters for sales.

Using the pooled data we can also ask the baseline question of how community demographics affect sales. Sales do depend on the community's racial and gender composition,

even after controlling for the community's income, unemployment, and population density (Table 4 column 1). Sales are significantly higher in communities with a greater female and a lower black population share. Female population share varies little; so this coefficient has limited economic significance. These results implicitly condition on the firm's decisions of how to market and where to open stores. (Few stores closed in our sample period.) Either the company has not completely succeeded in marketing to a diverse customer base, or its choice of locations has not equalized sales on the margin across stores. The impact on profits depends on the extent to which these sales differences are capitalized in store rents.

2. Positive or Negative Customer Discrimination

To test positive versus negative customer discrimination, Table 4, column 2, adds quadratic terms in each race's store employment share. Results differ across the racial groups. Employing Hispanics is useful in the relevant range but at a declining rate. In other words, a store's sales are higher if it employs at least a few Hispanics, but this effect quickly tails off, consistent with our theory of positive discrimination.

The reception of blacks is strikingly different. The first-order term on the proportion black in the store is insignificantly negative while the squared term is significantly negative. This suggests that the first few blacks in a store have little effect on sales, but that beyond that low threshold, sales decline with the proportion black. Omitted productivity characteristics (for example, that blacks attend worse schools), could account for a linear effect. But a priori it would seem unlikely that the quality distribution could be so thin at this end of the labor market. The accelerating decline in sales as black employment share increases suggests negative customer discrimination: many customers avoid stores with many blacks.

3. Immigrant Enclaves

Table 5 presents tests to see if additional Hispanic or Asian employees are particularly valuable in communities with nearby enclaves of Hispanic or Asian immigrants who do not speak English. The effects for Hispanics are not statistically significant. Matching is important for stores in communities with many non-English speaking Asians.

Column 1 presents the pooled time series, cross-section results (with random effects for stores). Stores with more Asian employees have higher sales if the community has many Asian immigrants who do not speak English. Because we necessarily group together Asian employees

of varying languages and fluency, the effect of hiring an employee who speaks the language of the enclave is presumably larger than the estimate reported here.

To understand the magnitude of the coefficient of 7.1 on the interaction of the share of the store's percent Asian and the community's percent speaking an Asian-Pacific language but not English, consider two communities that differ by ten percentage points on the share of linguistically isolated Asians. This coefficient implies that a store with a 10 percent point greater Asian employee share has 7.1 percent higher sales in the community with more linguistically isolated Asians than in the community with fewer. This effect is both economically and statistically significant across specifications. When we look between stores (column 2), the interaction for Asians rises in size. Examining a complementary cut of the data, when we look within stores (column 3), the point estimates on having a rising proportion of the store's workforce who share the background of the linguistically isolated remain statistically significant. Finally, in the within-store regression with ZIP code fixed effects (column 4), the coefficient on the interaction for Asians drops in size but remains statistically significant.

C. Team Effects: Diversity within Store

Team effects are the second channel through which diversity may affect sales. Our second major result is the small magnitude of most of these team effects. Diversity is identified by a non-linear effect of changing demographic employment shares. Even where the effect of team diversity on sales is statistically significant, it is modest in magnitude. Age diversity is the exception.

The within-store estimates in Table 3 indicate that growing age diversity predicts significantly lower sales growth. One standard deviation above-average dispersion of log age (almost 5 percent, so that two workers picked at random are about a year further apart in age) reduces sales growth by slightly less than 0.5 percent in Table 3, column 1, and slightly more than 0.5 percent in column 3 (with ZIP code fixed effects).

While we can statistically identify diversity as a nonlinear effect distinct from the main effects, at least two of the racial shares must change to change racial diversity. The total effect will depend on the corresponding shifts in employment shares. The partial effect of rising racial diversity is statistically significant and negative, but for moderate changes in store demographics the total effect on sales is insignificant. A move from an all-white store to roughly the retail

chain average (70 percent white and 10 percent each other group) would reduce sales by 1.3 percent (statistically significant at the 1 percent level). If we continue to increase diversity from a point close to the retail store average to a highly diverse store (40 percent white, 20 percent each other group) sales are unchanged. The point estimate is a tiny and insignificant -0.3 percent. Because of the positive main effect on percent Asian and the negative main effect on percent black, this result varies depending on the precise mix of workers that changes to create any given shift in overall diversity.[9] In contrast to race, changes in gender diversity do not predict changes in sales.

1. Results in Pooled Time-Series Cross-Section Data

In the pooled data **(**Table 4, column 3), age diversity again hurts sales, gender diversity is again insignificant, and racial diversity is now weakly positive. Given that most stores have a white majority, increasing racial diversity implies increasing the share of Asians, blacks, and Hispanics. Including the negative main effects on sales of each nonwhite race (whose shares increase with diversity) to calculate the total derivative, we find that over most of the relevant range, the total effect of increasing diversity is small, negative, and not statistically significantly.

In contrast, the estimated effect of age diversity is important; increasing our measure of age diversity by one standard deviation lowers sales by 15 percent. Hamilton, Nickerson, and Owan (2003) report a similar negative impact of age diversity on productivity in a manufacturing plant. This result recurs in all specifications, and will be discussed below. [10]

2. Between-Store Results

Most of the main results from the pooled analyses reappear when we ignore time-series variation and look solely at between-store averages. The results in Table 4, column 4, correspond with equation (4). In column 5 gender diversity has no statistically significant effect on sales. As in the pooled specification, age diversity again predicts lower sales; the effect is even larger in the cross-section. Over part of its range, the direct effect of racial diversity is to help sales, an effect that is both stronger and more significant in the cross-section than in the pooled specification.

The total effect of changing the racial composition of a store from an all-white store to one with a mixture close to the national average (70 percent white, 10 percent each of black, Hispanic, and Asian) would raise predicted sales by 4.2 percent. This is statistically insignificant even at the 10 percent level. Moving from that medium level to a highly diverse store (40

percent white, 20 percent each of black, Hispanic, and Asian) would decrease sales by 2.5 percent - again insignificant. Over most of the range of variation in store demographics, team diversity has an insignificant and small impact on sales both across and within stores.

Our most robust result concerning diversity within the store is the cost of age diversity. We replace mean age and the standard deviation of log age with the shares of employees who are teenagers, 20-22, 23-26, 27-33, and over 33. Compared to those 20-22, teens are less productive, while the older employees are slightly more productive, with the precise pattern depending on whether we use variation between stores or look at changes over time. However, when we control for both age diversity (the standard deviation of the log of age) and the proportion of the store under 20 or the proportion over 33, the age diversity measure remains strongly negative and statistically significant, while the age shares are small and statistically insignificant. This result suggests that the negative effects of age diversity that we find results from something more than the lower productivity of teens or of employees who remain in this sector longer than most.

D. Robustness Checks

We have run a large number of robustness checks. In all cases, results are consistent with the results presented above, with most store-community interactions small and insignificant other than results concerning linguistically isolated customers.

We considered a number of factors that might bias our measures of store-community similarity. Some stores are in neighborhoods that attract many shoppers who are not from the community. We use several means to identify such stores and rerun the analyses after dropping stores likely to serve a broader customer base. Results remain unchanged. The potential customer base may be difficult to measure with precision. Measurement error will attenuate the estimated coefficients on store-community demographic interactions. To test for sensitivity and reduce the effect of such error, we reran the regressions examining only communities where the racial mix was fairly similar at both a 5- and 10-mile radius. (Specifically, we took the half of the observations where the share of whites varied the least between the two measures.) Results showed no consistent increase in the effects of hiring employees who matched the community, giving us some confidence that the absence of a community match effect is not an artifact of the particular implementation of community used here. There was a suggestion in some specifications that Hispanic employees might be most useful in Hispanic communities, but when we reran the analysis on communities where the Hispanic share was most similar at the 5- and

10-mile radius, those results were no longer present. The absence of an effect on sales from matching employee and community demographics does not appear to be an artifact due to measurement error on community demographics.

Store reputation might lag changes in employment demographics. As a check, in the pooled and within-store regressions, we use store demographics that are lagged a month or that are the average of the last year. In case reputations take a long time to change, we look at two-year changes in sales as a function of two-year changes in store demographics and their interaction with community demographics. In case reputations are less important in stores with unstable demographics, we check if matching the community matters more in stores with stable demographics. The store-community interactions neither increase in size nor gain statistical significance.

Year-on-year changes in monthly store demographics may amplify the importance of transitory fluctuations in demographics. We average sales and demographics over 3-month periods and analyze year-on-year changes in quarterly store demographics and sales. Results are similar to those reported in the text.

To test whether the functional forms chosen might be driving the results, we perform a simple nonparametric test, looking at how store sales grow when the proportion black at the store rises as a function of the proportion black in the community. The results show no interaction. We repeat this exercise for the other racial and ethnic groups with similar lack of results.

We also replace the interactions of store and community race shares with the absolute value of the gap in store and community demographics. Results remain similar. Because the stores are typically less white than their communities, and because the absolute value of the gap is more sensitive to mismeasurement of demographics, we stress the specifications with the store-community interactions.

While racial discourse in the US is dominated by the categories of black and white, the spectrum of race relations is more complex. We examine the impact on sales of each of the cross-group interactions. We replace the interaction of the percent black in the store times percent black in the community with the three interactions of percent black in the store with percent white, Asian, and Hispanic in the community. We perform similar substitutions for the other groups (percent Asian in the store times percent black in the community and so forth). Overall, results are rarely precisely estimated and show no strong patterns.

For the regressions analyzing linguistically isolated potential customers, we examine the effect of Asian and of Hispanic employees in communities with at least 1 percent linguistically isolated Asian or Spanish speakers and then again in communities with at least 5 percent. Results are consistent with the interactions presented in Table 5 in that minority employees are particularly useful in the communities where customers are most likely to need the employees' language skills.

We test if within-store racial diversity is most useful in racially diverse communities. This interaction is neither large nor statistically significant.

We were interested in whether manager-community similarity increased sales. The hypotheses here are identical to those for worker-community similarity. The results were similarly unsupportive overall, with one exception. The single result supportive of manager-community similarity increasing sales is that, when comparing across stores, stores with black managers had higher sales when in highly black communities than in other communities. At the same time, using the more convincing longitudinal variation, stores that gained a black manager had *slower* growth when the store was in a highly black community than in a less black community. Similarly, when controlling for ZIP code fixed effects, when a store switches to a Hispanic manager, sales decline in highly Hispanic communities. The other manager race interactions are negative but not significant.

These various tests show no consistent evidence that having workers or managers who resemble their community affects sales.

E. The Locus of Discrimination

Opinion surveys have long attempted to measure the extent and locus of discriminatory attitudes in the US. In recent decades, few will admit to holding such beliefs. While this is encouraging, one wonders whether actions match the stated attitudes. The stores we study are so pervasive that we can use them as a probe of discrimination – a survey with bricks. Rather than ask about professed attitudes, we examine actions, using stores as a fairly uniform test instrument. Are sales affected differently by employee demographics under different circumstances? We compare stores in large and small cities classified by population density within 2 miles of each store; large and small stores classified by square feet; stores in rich and poor communities classified by median household income, and stores in communities with older and younger populations. In each case we compare the demographic effects on sales among

stores in the first quartile of each distribution to the effects found among stores in the last quartile of each distribution. We also compare effects in the Northern states to those in the South. These results are based on cross-section specifications and, because of limited testable variation, are rarely significant in the time-series dimension.

The comparisons across city size test differences in search costs between thin and thick markets. Densely populated communities offer greater choice among retail establishments. In areas with high population density, this employer often has multiple workplaces in nearby shopping districts and may face incentives to segregate its workforce so that each workplace specializes in a single demographic group (Becker 1957). Diversity across establishments—each one of which might be perfectly segregated—can substitute for diversity within an establishment to ensure employees match customers at each store.[11] By pooling the customer sub-communities within a locale, our matching measure would report poor employee-customer match in all of the stores.

At the other extreme, consider the isolated store in a small town. We compare small and large communities to test whether diversity within a store is more important within smaller communities with less retail choice. There is some evidence to suggest it is. To save space, we do not present tables. In cross-section estimates of our standard specifications, racial diversity has a significantly more positive effect on store sales in small than in large communities. The thicker markets in larger cities allow for more specialized stores, including those with more homogeneous staffs, to find sufficient customers. Customers with a preference for staff of a particular race can find them by searching across rather than within stores.

The corollary of more racial segregation across stores in big cities than in small is, however, not strongly borne out in the data. The test is not straightforward, since it depends on non-robust case-control methods that search for small cities with the population diversity found in big cities, and in big cities selects smaller stores that mirror store size in smaller cities. While the prediction of more segregation in bigger cities may seem a paradoxical result to those who think of bigger cities as more sophisticated, perhaps less discriminatory, and inherently more diverse, the result follows directly from classic economic models in which bigger markets allow greater specialization. Our result parallels a similar finding for radio stations (Sieglman and Waldfogel, 2001).

The negative impact of blacks on sales is found in large cities, not in small, and the difference is significant. In addition to the theories examined above, this result is also consistent with suburban blacks differing from urban blacks in ways that whites are more comfortable with—a result distinct from that of Raphael, Stoll and Holzer. The adverse impact of females on sales is also worse in big cities suggesting there are other forces at work. In contrast, Asians have a more positive impact in big cities.

Racial diversity improves sales in small stores but not in big. Since in this company, big stores are found in big cities, this result may reflect the same model at work. A distinct theory for different effects between small and large stores is statistical. These workforces turn over about 3 times a year. If customers are looking for demographic matches, past store demographics are a noisier measure of current demographics at small than at large stores because of the weak law of large numbers. Instead, we see that the negative effect of blacks on sales is greater at small stores.

A third stratification is between rich and poor communities. Because we measure both population and median incomes within two-mile circles, and because population density and incomes are positively correlated, this may again partially reflect city size effects. The adverse impacts of females and blacks on sales are significantly less in rich than in poor communities. Perhaps the rich are more tolerant concerning those who serve them.

Comparing young to old communities captures both life-cycle and historical changes. Communities of older people appear more tolerant of females, blacks, and age diversity. Interpreted as a historical effect, this is not promising because it suggests more recent cohorts discriminate more. However, we cannot empirically distinguish this from the more optimistic interpretation that discrimination fades with age and experience.

Despite the perception left by the Civil War and Reconstruction, the South has had a longer experience of confronting racial division. We find that blacks have a negative impact on sales only in Northern states. In the South, the effect is insignificant.

F. Limitations

These results may not generalize to other employers or to other sectors of the economy. Because the retail and restaurant industries employ roughly one sixth of the U.S. workforce and is often the sector of first employment, results limited to this sector are still important.

Diversity may matter less in this sector than elsewhere. This employer has a strong national brand. It is an open question whether potential customers react more to the brand than to the demographics of current employees. These workplaces demand relatively little employee-customer interaction. The low status of these jobs implies that customers may care less about the race of those that serve them. Diversity may also matter less because frontline workers have so little discretion. In workplaces with more decision-making power, diversity may be helpful in spurring creativity and costly in terms of raising communication costs. All of these forces are muted here.

Employees who work in demographically dissimilar communities may be more familiar with the local customers' group than the average person of their race. Employee selection and self-selection make workplaces' customer-employee demographic match matter less than if employees were randomly allocated.[12]

On the other hand, the effects of diversity on sales may be greater in this sector than in others. It is easy for customers in malls and downtown shopping districts to look in the store window, see the demographic match, and choose a store based on similarities. The costs of both information and switching are low in such open commercial agglomerations, so customers may be particularly responsive to demographic differences with potential salespeople.

IV. Conclusion

By using data from a single employer with more than 800 establishments, our study design reduces the problem of omitted employer or occupation effects. Just as a natural scientist would want to replicate all conditions other than the experimental variable, the employer in this case promotes a consistent national brand and strives to hold fixed both human resource practices and the customer's experience across locations. This creates by design an unusual degree of homogeneity across locations.

Diversity studies can confound not only employer differences but also community differences with diversity effects. We add extensive controls for community characteristics that might affect sales. In within-store specifications, we completely control for all unchanging store and community characteristics by examining changes in sales. A community can experience an employment shock that might affect both the demographic mix of workers and demand for this company's products. In a "brothers" specification, we compare the effects of changing

demographics on sales over time within store, holding constant regional shocks to sales or workforce demographics that might also affect a nearby store within the same zip-code.

We find little payoff to matching employee demographics to those of potential customers except when the customers do not speak English. Asian immigrants who do not speak English apparently buy more from those of similar background. Beyond that result, we find no consistent evidence that most customers care whether the salespeople who serve them are of the same race or gender. Additionally, we tested whether employment diversity might still affect performance through a direct effect on teamwork among employees. We find no consistent evidence that the workgroup's performance depends on its racial or gender diversity, identified as a nonlinear effect.

Age diversity, in contrast, does consistently predict lower sales. Age diversity is not one of the categories economists traditionally emphasize in thinking about diversity and discrimination, perhaps because of the potential for offsetting effects over the lifecycle. At the same time, age diversity is perhaps the most consistent correlate of social distance as measured in organizational scholarship on diversity.[13] One interpretation of our results is that the social distance between a white male at 20 and a white male at 30 is larger than the social distance between a man and a woman of the same age and race or between two women of the same age who differ by race.

The effects of diversity vary along several dimensions. The fact that we can detect economically and statistically significant results on age diversity shows the ability of our methods to detect diversity effects in this setting. If gender or race diversity had effects as important as those of age diversity, we would have been able to detect them.

These results do not generally support the claim that employee diversity is important because customers desire to be served by those who physically resemble them. This result is especially important because many employers in this sector appear to hire based on fears of customer discrimination (Moss and Tilly 2001: 146-7). Workgroup diversity's effects for both good and ill may well be stronger in settings where employees have more discretion and autonomy, where workgroups are more stable, and where relations with customers are more complex.

To those concerned with the long and troubled history of discrimination and with its continuing specter in this country, these results should be heartening. After all, one of the

painful paradoxes of customer discrimination is that it could lead employers to discriminate in pursuit of greater profits even if they are themselves indifferent to race and gender. The paradox is heightened by diversity proponents who argue that customers discriminate and should be pandered to. At least at this firm, race and gender diversity do not appear costly. Managers in mostly white communities will not suffer lower sales if they hire black, Hispanic, or Asian employees. Neither the potential customers nor the employees' performance as measured by sales is much affected by the race or gender diversity of the workplace.

Appendix: Data and Variables

We combine employee-level data on demographics, store-level data on sales, and data from the 1990 Census on community characteristics. The employee data are the complete personnel records from February 1996 to October 1998. We analyze data on frontline workplace employees, dropping workplaces with fewer than ten employees. We organize the data into store-month observations.

We complement our quantitative analysis with semistructured interviews of roughly a dozen employees and a half-dozen managers at workplaces scattered across one region of the country. These interviews were neither random nor a representative sample, but they do help flesh out the statistical analyses.

1. Store-Level Variables

The dependent variable is the natural logarithm of real monthly sales. In our first set of specifications, we analyze data pooled across stores over time. We then look only at variation between stores, averaging each store's sales over all available store-months. We next analyze variation within the history of each store, looking at year-on-year differences in monthly sales. Finally, we add ZIP code fixed effects to the regressions on sales growth.

From the company's human resource database, we construct a store-month dataset of employee demographics, including the proportion female, average age, and the shares of three categories for race or ethnicity (black, Asian, and Hispanic, with white, the small percentage Native American, and unknown ethnicity categories pooled as the baseline). The race and ethnicity codes are the company's coding, and they create a set of mutually exclusive and collectively exhaustive categories that for simplicity we refer to as "race." Educational requirements are minimal, and educational attainment varies little. Few employees have a college degree. Additionally, the employer imposes few hiring prerequisites.

We control for a rich set of store characteristics when we analyze between-store variation; controls include the logarithm of employment, store age and its square, time since the last store remodel and its square, store size (measured in square feet) and its square, and indicator variables for if the store is on the street, a commercial strip, or in a mall.

Sales per store will also depend on the number of nearby competitors. We control for the number of establishments that are in the same county in the same four-digit industry as reported

in the 1998 County Business Patterns. To control for other local factors, some estimates include an extensive set of dummy variables, one for each ZIP code with more than one store.

2. Community Variables

To construct community demographics, we use each store's ZIP code to identify a zone of "nearby" Census tracts, defined as those in its ZIP code or within two miles of the centroid of its ZIP code. We then merge 1990 Census data for this zone to each store.

We construct the proportion black, Hispanic, Asian, and female surrounding each store, as well as the age distribution in the surrounding community using the following data. The 1990 Census asks questions on race (black vs. white, etc.) separately from ethnicity (Hispanic vs. non-Hispanic). Thus, on the Census, respondents can categorize themselves as both black and Hispanic or as both white and Hispanic. In contrast, the employer has mutually exclusive codes of white, black, and Hispanic (as well as Asian). We allow both the Census categories of population and the employer's categories of employment to enter unrestricted in our equations.

We control for several other community characteristics likely to affect product demand. As control variables, we use Census data on the household income distribution (percentages of households in each of ten detailed income categories), the age distribution (percentages of individuals in each of six age categories), total population within two miles, population within two miles categorized into six size groups, and the unemployment rate. Because population is measured within a fixed two-mile radius, it can be thought of as a population-density measure. The income figures are only available for the store's ZIP code, without the two-mile radius of surrounding tracts.

3. Store-Community Interactions

For matching theories, the variables of interest are the interaction between store and community demographics. Such interactions allow us to test, for example, for the effect of having a highly Hispanic workforce near a Hispanic population center. The racial composition of the stores are highly correlated with the composition of the community (for example, the white shares are correlated at 0.70); nevertheless, substantial variation remains across stores. In addition, the racial shares vary substantially over time as well.

We also measure the interaction between the proportion female at the store and in the community. Aside from some areas containing military bases, single-sex colleges, and mining

operations, there is much less variation in gender shares than in race or ethnicity across locations. Thus, we have little testable variation in the proportion of females across communities.

4. Diversity within the Store

We calculate age, gender, and racial diversity within the store as well as the surrounding community. For race and gender, we use a diversity index equal to the odds that two people selected at random from a workplace differ on race or gender. The formula is that the diversity index is one minus the sum of the demographic shares squared:

$$\text{Diversity index on race or gender} = 1 - \Sigma_i S_i^2,$$

where S_i is the share of each gender or racial group i. This diversity index is zero with complete homogeneity and is maximized when each group has an equal share of employment. Economists might naturally think of it as one minus the Herfindahl Index.

Most past researchers have used the coefficient of variation on age or the standard deviation of age to measure age diversity. We prefer to use the standard deviation within the workgroup of the natural logarithm of age. The standard deviation of log(age) implies that proportional gaps in age are what lead to social distance; for example, the age gap between 18 and 22 usually leads to more social difference than does the age gap of 40 to 44, although the two gaps are the same in absolute years. As with the race and gender diversity indices, the standard deviation of log(age) has a simple interpretation: It is approximately the expected percentage gap in the age of two people chosen at random. This relation holds exactly for normally distributed variables.

References

Ian Ayres, Fred Vars and Nasser Zakariya, "To Insure Prejudice: Racial Disparities in Taxicab Tipping" 114 *Yale Law Journal* 1613 (2005).

Bantel, K. A., and S. E. Jackson, "Top management and innovations in banking: Does the composition of the top team make a difference?" *Strategic Management Journal* 10 (1989), 107-124.

Bayard, Kimberly, Judith Hellerstein, David Neumark, and Kenneth Troske, "New Evidence on Sex Segregation and Sex Differences in Wages from Matched Employer-Employee Data," *Journal of Labor Economics* 21:4 (2003), 887-922.

Becker, Gary Stanley, *The Economics of discrimination* (Chicago: University of Chicago Press, 1957).

Blalock, H. M., *Toward a theory of minority group relations* (New York: Wiley, 1967).

Borjas, George J, "Labor Turnover in the U.S. Federal Bureaucracy," *Journal of Public Economics* 19:2 (November 1982), 187-202.

Burdekin, Richard C. K., and Todd L. Idson, "Customer Preferences, Attendance and the Racial Structure of Professional Basketball Teams," *Applied Economics* 23:1B (January 1991), 179-186.

Churchill, G. A., R.H. Collins, and W.A. Strang, "Should retail salespersons be similar to their customers?" *Journal of Retailing* 51:3 (1975), 29-42.

Collier, Paul, and Jan Gunning, "Explaining African Economic Performance," *Journal of Economic Literature* 37 (March 1999), 64-111.

Costa, Dora, and Matthew Kahn, "Cowards and heroes: Group Loyalty in the American Civil War," *Quarterly Journal of Economics* (May 2003), 519-548.

Cox, T., *Cultural diversity in organizations: Theory, research and practice* (San Francisco: Berrett-Koehler Publishers, 1993).

Dwyer, Sean, Orlando Richard, and C. David Shepherd, "An exploratory study of gender and age matching in the salesperson--prospective customer dyad: Testing similarity--performance predictions," *The Journal of Personal Selling & Sales Management* 18:4 (Fall 1998), 55-69.

Friedman, N., and J. DiTomaso, "Myths about diversity: What managers need to know about changes in the US labor force," *California Management Review* 38:4 (1996), 54-57.

Garson, B., *Money makes the world go around* (Penguin Putnam: New York, 2002).

Greif, Avner, "Contract enforceability and economic institutions in early trade: The Maghribi traders' coalition," *American Economic Review* 83:3 (Jun 1993), 525-548.

Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan, "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation," *Journal of Political Economy* 111:3 (June 2003), 465-497.

Hamilton, Barton H., Jack A. Nickerson, and Hideo Owan, "Diversity and Productivity in Production Teams," mimeo, Olin School, Washington University, (2004).

Hellerstein, Judith, and David Neumark, "Ethnicity, Language, and Workplace Segregation: Evidence from a New Matched Employer-Employee Data Set," *Annales d'Economie et de Statistique* (July/December 2004), 19-78.

Hellerstein, Judith K., David Neumark, and Kenneth Troske, "Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations," *Journal of Labor Economics* (1999), 409-446.

Holzer, Harry J., *What Employers Want: Job Prospects for Less-Educated Workers* (New York: Russell Sage Foundation, 1999).

Holzer, Harry J., and Keith R. Ihlanfeldt, "Customer discrimination and employment outcomes for minority workers," *Quarterly Journal of Economics* 113:3 (August 1998), 835-867.

Holzer, Harry, and David Neumark, "Are Affirmative Action Hires Less Qualified? Evidence From Employer-Employee Data on New Hires," *Journal of Labor Economics* (1999), 534-69.

Ibarra, H., "Race, opportunity, and diversity of social circles in managerial networks," *Academy of Management Journal* 38 (1995), 673-703.

Ibarra, H., "Homophily and differential returns: Sex differences in network structure and access in an advertising firm," *Administrative Science Quarterly* 37 (1992), 422-447.

Ichniowski, Casey, Kathryn Shaw, and Jon Gant, "Working Smarter by Working Together: Connective Capital in the Workplace," mimeo, Stanford GSB (2003).

Jackson, S. E., and E.B. Alvarez, "Working through diversity as a strategic imperative" (pp. 13-35), in S. E. Jackson (Ed.), *Diversity in the workplace: Human resources initiatives*, (New York: The Guilford Press, 1992).

Jackson, S. E., J. F. Brett, V. I. Sessa, D. M. Cooper, J. A. Julin, and K. Peyronnin, "Some differences make a difference: Individual dissimilarity and group heterogeneity as correlates of recruitment, promotions, and turnover," *Journal of Applied Psychology* 76 (1991), 675-689.

Jehn, K. A., G. B. Northcraft, and M. A. Neale, "Why differences make a difference: A field study in diversity, conflict, and performance in workgroups," *Administrative Science Quarterly* 44 (1999), 741-763.

Joshi, A., and S. E. Jackson, Does team member distinctiveness predict team performance? Paper presented at the Academy of Management Meetings, Denver, Colorado (2002).

Kahn, Lawrence M., and Peter D. Sherer, "Racial Differences in Professional Basketball Players' Compensation," *Journal of Labor Economics* 6:1 (January 1988), 40-61.

Kandel, Eugene and Edward P. Lazear, "Peer Pressure and Partnerships," *Journal of Political Economy* 100:4 (August 1992), 801-17.

Lang, Kevin, "A Language Theory of Discrimination," *Quarterly Journal of Economics* 101:2 (May 1986), 363-382.

Lazear, Edward. "Globalization and the Market for Teammates," National Bureau of Economic Research working paper no. W6579 (May 1998).

Lee, Jennifer, "The Racial and Ethnic Meaning behind *Black*: Retailer's Hiring Practices in Inner-City Neighborhoods" (168-188), in John David Skrentny (Ed.), *Color Lines: Affirmative Action, Immigration, and Civil Rights Options for America* (Chicago: University of Chicago, 2001).

Leonard, Jonathan, and David I. Levine, "The Effect of Diversity on Turnover: A Very Large Case Study," *Industrial and Labor Relations Review* (2006).

Moss, Philip I., and Chris Tilly, *Stories Employers Tell: Race, Skill, and Hiring in America* (New York: Russell Sage Foundation, 2001).

Nemeth, C., "Dissent, group process and creativity" (57-75), in E. Lawler (Ed.), *Advances in group processes: Theory and research* (Greenwich: JAI Press, 1985).

Neumark, David, "Sex discrimination in restaurant hiring: An audit study," *Quarterly Journal of Economics* 111:3 (August 1996), 915-941.

O'Reilly, C. A., D. F. Caldwell, and W. P. Barnett, "Work group demography, social integration, and turnover," *Administrative Science Quarterly* 34 (1989), 21-37.

Pfeffer, J., "Organizational Demography," *Research in Organizational Behavior* 5 (1983).

Raphael, Steven, Michael Stoll, and Harry Holzer, "Are Suburban Firms More Likely to Discriminate against African-Americans?" *Journal of Urban Economics* 48:3 (November 2000), 485-508.

Reskin, Barbara F., Debra B, McBrier, and Julie Kmec, "The Determinants and consequences of workplace sex and race composition," *Annual Review of Sociology* 25 (1999), 335-361.

Rodrik, Danni, "Where Did All the Growth Go? External Shocks, Social Conflict, and Growth Collapses," National Bureau of Economic Research working paper no. 6350 (1998).

Siegelman, Peter and Joel Waldfogel, "Race and Radio: Preference Externalities, Minority Ownership, and the Provision of Programming to Minorities," *Advances in Applied Microeconomics* 10 (2001).

Tannen, Deborah. *You just don't understand: Women and men in conversation*, 1st ed. (New York: Morrow, 1990).

Thomas, D. A., and R. J. Ely, "Making differences matter: A new paradigm for managing diversity," *Harvard Business Review* 74 (1996), 79-90.

U.K. Home Office, Lord Leslie Scarman, "Brixton disorders 10-12 April 1981: report of an inquiry by the Rt. Hon. the Lord Scarman, O.B.E". (London: HMSO, 1981)

Watkins, Steve. *The black O: Racism and redemption in an American corporate empire* (Athens: University of Georgia Press, 1997).

Watson, W.E., K. Kumar, and L.K. Michaelson, "Cultural diversity's impact on interaction process and performance: Comparing homogeneous and diverse task groups," *Academy of Management Journal* 36 (1993), 590-602.

Williams, K. Y., and C. A. O'Reilly, "Forty years of diversity research: A review," in M. A. Neale, E. A. Mannix, and D. H. Gruenfeld (Eds.), *Research on managing groups and teams* (Greenwich, CT: JAI Press, 1998).

Zenger, Todd R., and Barbara S. Lawrence, "Organizational Demography: The Differential Effects of Age and Tenure Distributions on Technical Communication," *Academy of Management Journal* 32:2 (June 1989), 353-376

Table 1: Summary Statistics

	Pooled data		One-year changes	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
log real sales	(omitted)	0.658	(omitted)	0.180
log employment (Average employment is about 30 frontline employees per store, mostly part-time)	(omitted)	0.505	0.127	0.237
Store Demographics				
Average age	24.3	2.28	-0.213	1.727
%Female	0.750	0.137	-0.004	0.089
%Black	0.119	0.134	0.013	0.071
% Hispanic	0.100	0.131	0.007	0.060
%Asian	0.070	0.089	0.006	0.054
Average age^2	595.1	115.4	-10.4	87.9
%Female^2	0.582	0.204	-0.006	0.129
%Black^2	0.032	0.070	0.006	0.038
% Hispanic^2	0.027	0.074	0.002	0.032
%Asian^2	0.013	0.037	0.002	0.021
S.D.(Log(age))	0.270	0.062	0.004	0.047
Gender Diversity= 1-[(%female)^2 + (%male)^2]	0.337	0.140	0.005	0.088
Racial Diversity = 1-[(%White)^2+(%Black)^2+(%Hispanic)^2+(%Asian)^2]	0.392	0.207	0.018	0.112
Community Demographics				
%Female	0.512	0.017		
%Black	0.075	0.094		
% Hispanic	0.051	0.069		
%Asian	0.051	0.078		
% Speak only Spanish	0.005	0.011		
% Speak only an Asian language	0.005	0.015		
%Female^2	0.262	0.017		
%Black^2	0.014	0.047		
% Hispanic^2	0.007	0.030		
%Asian^2	0.009	0.047		
Gender Diversity	0.499	0.002		
Racial Diversity	0.318	0.184		
Store-Community Interactions			Changes use levels of community variable and change in %of store	
Store %Female – Community %Female	0.238	0.138	-0.002	0.046
(Store %Black)*(Community %Black)	0.015	0.039	0.002	0.014
(Store % Hispan)*(Community % Hispan-all races)	0.019	0.061	0.001	0.009
(Store %Asian)*(Community %Asian)	0.008	0.036	0.001	0.013
(Store % Hispan)*(Community %speak only	0.001	0.006	0.0001	0.0013
(Store %Asian)*(Community %speak only Asian	0.001	0.004	0.0001	0.0014

The sample contains over 20,000 store-months at over 800 stores. Between-store summary statistics resemble pooled.

Table 2: Overview of Results

Specification	Pooled Time Series Cross Section	Between stores	Year-on-Year Changes Within Stores	Year-on-Year Changes with ZIP code fixed effects
Community Matching				
Black	Insig. +	Insig. +	Insig. +	Significant -
Hispanic	Insig. +	Insig. +	Insig. -	Insig. -
Asian	Insig. -	Insig. +	Significant -	Insig. -
Female	Insig. Mixed	Insig. +	Insig. +	Insig. +
Language-Hispanic	Insig. +	Insig. -	Insig. -	Insig. -
Language-Asian	Significant +	Significant +	Significant +	Significant +
Team Effect				
Age	Significant -	Significant -	Significant -	Significant -
Gender	Insig. +	Insig. -	Insig. -	Insig. -
Race	Insig. -	Insig. +	Significant -	Significant -

Note: Team effect for race refers to sign of total derivative for a change from all white to average demographics.

Table 3: Year-on-Year Changes Within Stores

Dependent Variable = 1 year %change in sales

	Within-Store Estimates Entire sample		**Adding ZIP Code Fixed Effects** Sample contains stores that have at least two stores in the same ZIP code	
	(1) Interactions	**(2) Diversity**	**(3) Interactions**	**(4) Diversity**
Δ Avg. Age in the Store	0.006	0.004**	-0.004	0.005**
	(0.007)	(0.001)	(0.008)	(0.001)
Store Δ %Female	-0.394	-0.014	-0.169	-0.005
	(0.379)	(0.030)	(0.410)	(0.034)
Store Δ %Black	-0.078**	-0.044*	-0.037	-0.041
	(0.029)	(0.022)	(0.034)	(0.027)
Store Δ %Hispanic	-0.041	0.023	-0.047	0.022
	(0.032)	(0.025)	(0.035)	(0.028)
Store Δ %Asian	-0.010	0.084**	0.064	0.089**
	(0.032)	(0.027)	(0.036)	(0.031)
Δ (Avg. Age in the Store ^{2})	-0.000		0.000	
	(0.000)		(0.000)	
Store Δ (%Black ^{2})	-0.014		0.044	
	(0.061)		(0.073)	
Store Δ (%Hispanic^2)	0.143		0.134	
	(0.093)		(0.099)	
Store Δ (%Asian ^{2})	0.373**		0.095	
	(0.128)		(0.154)	
(Store Δ %Female)*(Community %Female)	0.804		0.344	
	(0.742)		(0.804)	
(Store Δ %Black)*(Community %Black)	0.072		-0.447*	
	(0.152)		(0.184)	
(Store Δ %Hispanic*(Comm. % Hispanic-all races)	-0.183		-0.129	
	(0.204)		(0.238)	
(Store Δ %Asian)*(Community %Asian)	-0.671**		-0.477	
	(0.225)		(0.269)	
Store Δ st.dev. ln(age)		-0.071*		-0.112**
		(0.031)		(0.034)
Store Δ Gender Diversity		-0.031		-0.012
diversity = 1-[(%female)^2 + (%male)^2]		(0.029)		(0.032)
Store Δ Racial Diversity		-0.040*		-0.042*
diversity = 1-		(0.016)		(0.019)
[(%white)^2+(%black)^2+(%Hispanic)^2+(%Asian)^2]				
Observations: stores	over 800	over 800	over 600	over 600
store-months	over 20,000	over 20,000	over 10,000	over 10,000
Number of 5-digit ZIP code dummies	0	0	over 300	over 300
R-squared	.239	.240	.338	.338

Notes: Standard errors in parentheses. * significant at 5%; ** significant at 1%. Additional controls included %change in employment, store age and its square, time since last remodel and its square, store size in square feet and its square, store division, store location type (mall, street, etc.; column 1 only), Δ% Native Americans, Δ% other races, and month dummies. Standard errors are adjusted for first-order autocorrelation within stores and for heteroskedasticity across stores.

Table 4: Pooled Time Series Cross Section & Between Stores

	(1) Baseline Pooled	(2) Interactions Pooled	(3) Diversity Pooled	(4) Interactions Between	(5) Diversity Between
Dependent Variable	**Log Real Monthly Sales**	**Log Real Monthly Sales**	**Log Real Monthly Sales**	**Log (Average real sales)**	**Log (Average real sales)**
Store Employees Avg. Age	0.004**	0.023**	0.007**	0.020	0.020**
	(0.001)	(0.009)	(0.001)	(0.042)	(0.005)
Store %Female	-0.024	0.006	-0.002	-0.390**	-0.348*
	(0.016)	(0.023)	(0.033)	(0.141)	(0.156)
Store %Black	-0.078**	-0.003	-0.118**	-0.064	-0.408**
	(0.020)	(0.035)	(0.027)	(0.164)	(0.098)
Store % Hispanic	0.030	0.047	-0.011	0.661**	-0.050
	(0.024)	(0.038)	(0.030)	(0.194)	(0.120)
Store %Asian	0.058*	0.015	0.010	-0.132	-0.456**
	(0.026)	(0.041)	(0.035)	(0.247)	(0.160)
Community %Female	1.123**	1.138*	1.117*	-0.852	-0.798
	(0.434)	(0.449)	(0.457)	(0.552)	(0.547)
Community %Black	-0.455**	-0.526**	-0.475**	-0.329	0.063
	(0.076)	(0.144)	(0.116)	(0.192)	(0.154)
Community % white Hispanics	0.578**	0.756*	0.586**	0.001	0.448*
	(0.124)	(0.321)	(0.142)	(0.450)	(0.199)
Community %Asian	0.133	0.443*	0.121	0.061	0.421**
	(0.084)	(0.220)	(0.101)	(0.317)	(0.161)
(Store Avg. Age) ^{2}		-0.000*		-0.000	
		(0.000)		(0.001)	
(Store %Black) ^{2}		-0.176*		-0.374	
		(0.069)		(0.332)	
(Store %Hispanic) ^{2}		-0.141		-1.398**	
		(0.108)		(0.521)	
(Store %Asian) ^{2}		0.133		-0.361	
		(0.146)		(0.905)	
(Community %Black) ^{2}		0.178		0.233	
		(0.307)		(0.474)	
(Community %Hispanic) ^{2}		-0.475		0.200	
		(0.639)		(1.044)	
(Community %Asian) ^{2}		-0.503		0.054	
		(0.357)		(0.807)	
Top quartile		0.003		0.009	
(Store %Female – Community %Female)		(0.005)		(0.027)	
Bottom quartile		0.012**		-0.028	
(Store %Female – Comm. %Female)		(0.004)		(0.025)	
(Store %black)*(Community %black)		0.012		0.448	
		(0.156)		(0.551)	
(Store % Hispanic)*		0.230		0.881	
(Community % Hispanic)		(0.215)		(0.720)	
(Store %Asian)*(Community %Asian)		-0.038		0.617	
		(0.269)		(1.488)	
Store Age Diversity			-0.157**		-0.821**
= S.D.(log(age))			(0.039)		(0.195)
Store Gender Diversity			0.022		-0.110
= 1-[(%female)^2 + (%male)^2]			(0.034)		(0.163)
Store Racial Diversity			0.046*		0.278**
= 1-[(%W)^2+(%B)^2+(%H)^2+(%A)^2]			(0.022)		(0.094)
R^2	0.78	0.78	0.78	0.86	0.86

Notes: Standard errors in parentheses. * (**) significant at 5% (1%). Additional controls include store age, time since last remodel, store square feet, and their squares, log(employment), store division, store type (mall, street, etc.), store and community %Native Americans and their interaction (col. 3-5), store and community % other races, 9 community income shares (such as % of households with incomes $50-75,000 per year); %unemployed in community, 5 measures of community age shares (such as % ages 30-49), six measures of population density (such as between 80,000 and 320,000 live within 2 miles), the number of competing establishments in this 4-digit SIC in this county, and month dummies (col. 1-3). Columns 3-5 include community racial diversity and gender diversity. Col. 5 includes each store's months in the sample and a count of the number of Decembers. Sample is over 800 stores and over 20,000 store-month observations (column 1-3).

Table 5: Results Concerning the Linguistically Isolated

Specification	Pooled Time Series Cross Section	Between stores	Year-on-Year Changes	Year-on-Year Changes with ZIP code fixed effects
Dependent variable	Log Real Monthly Sales	Log (Average real sales)	One year %change in sales	One year %change in sales
Controls and sample as in:	Table 4, col. 2	Table 4, col. 4	Table 3, col. 1	Table 3, col. 3
(Store % Hispanic)* (Comm. % Hispanic-all races)	0.199 (0.265)	1.001 (0.789)	-0.112 (0.277)	-0.342 (0.326)
(Store %Asian)* (Community %Asian)	-0.574* (0.285)	-0.586 (1.517)	-1.238** (0.246)	-1.007** (0.313)
(Store %Hispanic)* (Community % speaking only Spanish)	0.898 (1.831)	-0.805 (4.157)	-0.955 (1.769)	2.335 (2.503)
(Store %Asian)* (Community % speaking only an Asian-Pacific language)	7.058** (1.264)	15.414** (5.155)	8.654** (1.701)	5.709** (1.885)

Notes: Each column represents a subset of the coefficients from a separate regression specification. Other controls include the percent speaking only an Asian-Pacific language, the percent speaking only Spanish, and the additional variables as indicated at the top of each column. The proportions speaking only Spanish or an Asian language measure people who do not speak English; they may speak other non-English languages. The first-differences specifications (col. 3 and 4) include first differences of store variables, but not community ones.

Endnotes

[1] Due to data limitations described below, we refer to the categories white, black, Asian, and Hispanic as "race," although Hispanic is more accurately described as an ethnicity.

[2] In her study of retail stores in largely black neighborhoods, Lee (2001) identified a variant of communication as a motive for storeowners to hire employees who match customers' demographics. Storeowners in her inner-city sample prefer to have at least one black employee in the store who can resolve a tense situation without reinforcing racial conflict. Urban policing in the U.K. and in the U.S. provides similar examples (U.K. Home Office, The Scarman Report 1981).

[3] A partial listing of such cases in recent years includes Abercrombie and Fitch, Albertsons, Home Depot, Kroger, Lucky Stores, Safeway, Shoney's, Wal Mart and Winn-Dixie. If we were to include customer service jobs outside the retail sector, the list would include a number of financial services firms.

[4] As noted below, results using the absolute value of the gap in store and community demographics resemble those in the interaction specification (1). This absolute value of the gap is more sensitive to mismeasurement of the appropriate community and racial boundaries than the interaction we use.

[5] The estimates that use time series variation will have autocorrelated errors in the history of each store. We correct standard errors for first-order autocorrelation using the Prais-Winsten correction.

[6] The Hausman test examines if the coefficients on store characteristics are stable when we shift from random to fixed effects; the coefficients differ significantly, suggesting that fixed effects is more appropriate.

[7] Similar results are found comparing months, quarters or years one year apart.

[8] In results not shown, we find (as expected) that store racial composition largely reflects the demographics of the community. There remains testable variation in store demographics beyond community demographics because stores do not simply match their communities.

[9] These main effects could be due to customer discrimination or to differences in human capital, among other explanations.

[10] When we combine the store-community interactions with the within-store diversity measures, results remain similar (results available on request).

[11] For example, Garson (2002) describes several ethnically distinct shopping malls in the diverse city-state of Singapore. Each mall serves speakers of a specific language.

[12] Some evidence on turnover, rather than sales, reported in a companion paper with this dataset, does support the importance of similarity attraction among employees. Men, older workers, whites, and

blacks (but not the other groups) have lower turnover when they work around many similar co-workers. Similarly, blacks and Asians (but not whites or Hispanics) turnover less when customers are more likely to share their race.

[13] Age diversity predicts higher turnover (Jackson, et al. 1991; O'Reilly, et al. 1989; Wagner, et al. 1984), lower communication within a project group (Zenger and Lawrence 1989), higher role ambiguity (Tsui Egan, and O'Reilly 1992), and greater shirking (Costa and Kahn, 2003).