11-1-2007

Racial Diversity and Aggregate Productivity in U.S. Industries: 1980-2000

Chad Sparber
Colgate University, csparber@colgate.edu

Racial Diversity and Aggregate Productivity in U.S. Industries: 1980-2000*

Chad Sparber†

Forthcoming, *Southern Economic Journal*, 2008

Abstract

This paper employs industry-level U.S. Census data from 1980-2000 to assess the aggregate effects of racial diversity. While most international accounts find that diversity reduces productivity, I argue that the U.S. experience is more nuanced. Unqualified statements about the costs and merits of diversity are unwarranted, as racial heterogeneity increases productivity within many, but not all, industries. Sectors employing a large number of workers responsible for creative decision-making and customer service experience gains from diversity, while industries characterized by high levels of group effort suffer losses. The results thus reconcile two competing literatures by suggesting that diversity improves decision-making and problem solving, but also encumbers common action and public goods provision.

Key Words: Racial Diversity, Productivity

JEL Classification Codes: O40, J24, O51

*I would like to acknowledge the guidance and advice from the faculty and graduate students of the University of California, Davis. I especially thank Giovanni Peri, Hilary W. Hoynes, Alan M. Taylor, Florence Bouvet, and Ahmed S. Rahman. Three anonymous referees, Rukmani Gounder, Margriet Caswell, and the attendees of the Western Economics Association International Conference in San Francisco provided helpful comments and suggestions, as did Takao Kato, Robert W Turner, and faculty members of Colgate University and San Diego State University. A fellowship from the University of California Office of the President provided support for this research.

†Author Contact Information: Email: csparber@mail.colgate.edu; Address: Colgate University, 13 Oak Drive, Hamilton, NY 13346.

1 Introduction

The United States grew increasingly diverse between 1980 and 2000 as Whites declined from 83% to 73% of the labor force. Many states are reconsidering their affirmative action policies while the federal government reevaluates immigration and citizenship statutes. Unfortunately, we do not yet fully understand the aggregate effects of diversity. International empirical analyses often argue that racial (or ethnic) heterogeneity hinders macroeconomic growth and development by fostering conflict, corruption, political instability, and competitive rent-seeking. Many top U.S. business executives reject these views, however, and claim that individuals from varied backgrounds bring unique approaches to problem solving and complement each other in production. Several academic analyses echo these sentiments. This paper adds to the debate on the costs and benefits of diversity by building upon the traditional economic growth and development framework to measure racial diversity's net effect on productivity within sectors of the United States.

My methodology advances the existing literature by addressing two limitations of prior studies. First, although cross-country evidence for the costs of diversity is clear, the United States may be especially effective in managing diversity. This paper will determine whether the U.S. experience differs from international accounts. Second, typical international growth analyses measure aggregate outcomes for countries but fail to address differences in marginal effects that could exist across sectors of the economy. I depart from this strategy by analyzing industrial performance across U.S. states. To begin, I measure racial diversity by using decennial Census data to construct fractionalization indices for state-industry cells between 1980 and 2000. I then employ two stage weighted least squares regressions with fixed effects to estimate diversity's industry-specific influence on average wages paid to workers – a proxy for labor productivity in the absence of observed output per worker measures.

Unqualified statements regarding the costs and merits of diversity are unwarranted, as racial heterogeneity increases productivity within most, but not all, U.S. industries. These results are robust to controls for individual-level heterogeneity, state-specific trends, industry-specific trends, and supply factors. Prior academic research, anecdotes from political events, and the sector-specific coefficients in this paper provide intuition about the industrial characteristics that determine whether diversity leads to economic gains or losses. I use the U.S. Department of Labor's O*NET Consortium data to show that industries heavily reliant upon creative decision-making, problem solving, and customer service benefit from diversity. In contrast, sectors characterized by high levels of group or team work experience losses, which suggests that heterogeneity encumbers common action. The results therefore reconcile competing literatures by recognizing both the costs and benefits of diversity.

2 Motivation

Do racially diverse environments generate net economic gains or losses? International growth economists typically find evidence for the latter. Easterly and Levine's (1997) seminal investigation argues that "A movement from complete heterogeneity to complete homogeneity is associated with an income increase of 3.8 times." To them, diversity increases polarization, facilitates competitive rent-seeking between groups, and promotes growth-reducing political policies.[1] A number of social scientists have argued that diversity also conflicts with common action, and that heterogeneous societies tend to oppose wealth redistribution and public goods provision.[2] Others find evidence that diversity escalates corruption and political instability.[3]

Many top U.S. executives, in contrast, instead argue that a racially diverse workforce bolsters productivity. In the late 1990s, the University of Michigan began defending the constitutionality of its affirmative action admissions policies in a series of well-publicized court cases. In October 2000, executives from thirty *Fortune* 500 companies united to file

[1] Recent work by Alesina et al. (2003) provides a more detailed examination of diversity with similar results.

[2] See Poterba (1997), Alesina et al. (1999), Gilens (1999), Luttmer (2001), Alesina and Glaeser (2004), and Alesina and La Ferrara (2005). Caselli and Coleman (2002) provides an important theoretical model to descibe greed-motivated conflict between ethnic groups.

[3] See Shleifer and Vishny (1993), Mauro (1995), Knack and Keefer (1997), and Fukuyama (1999).

an *amicus* brief supporting the school. The brief contends "A diverse group of individuals educated in a cross-cultural environment has the ability to facilitate unique and creative approaches to problem-solving arising from the integration of different perspectives," and that "such individuals are better able to develop products and services that appeal to a variety of consumers and to market offerings in ways that appeal to these consumers." These anecdotal claims do find support in academic literature. In psychology, Campbell (1960) argues "persons who have been thoroughly exposed to two or more cultures seem to have an advantage in the range of hypotheses they are apt to consider, and through this means, in the frequency of creative innovation." Similarly, Simonton (1999) maintains, "creativity is favored by an intellect that has been enriched with diverse experiences and perspectives... It is as if the mere exposure to different lifestyles and divergent values enables individuals to expand the range and originality of their ideational variations."[4]

Empirical, qualitative, and anecdotal evidence suggests that an ideal theoretical model of diversity and productivity should reconcile the costs and benefits of diversity within a single production function. Assume an economy is composed of L_i workers from groups $i = 1...N$. The number of groups (N) is fixed, and group i comprises a fraction of the labor force equal to θ_i. A generic function $C(L_1, L_2, ..., L_N)$ measures the fraction of output remaining after firms (and society) suffer costs associated with employing individuals from disparate groups. These costs might reflect communication and conflict costs, diversion from productive activity, and/or disputes over proper public goods provision.[5] $C(L_1, L_2, ..., L_N)$ is homogenous of degree zero so that only the relative size of groups (that is, $\theta_1, \theta_2, ..., \theta_N$) will affect output. Costs are maximized when each group maintains an identical share of the labor force; costs decrease as the labor force share of one group approaches unity.

Suppose, however, that workers of different types also supply unique skills so that a diverse workforce provides a greater range of perspectives, ideas, and problem-solving tech-

[4]See Cox, Lobel, and McLeod (1991), O'Reilly, Williams, and Barsade (1998), Florida (2002), Hong and Page (2004), and Page (2007) for further evidence on the gains from diversity.

[5]See Lazear (1999), Alesina and La Ferrara (2005), Alesina, Spolaore, and Wacziarg (2005), Ottaviano and Peri (2005), and Sparber (2007b).

niques. Workers from varied groups will complement each other in production. A "love of labor variety" term can account for these complementarities, as can a nested CES specification that allows for imperfect substitutability across labor groups.[6] Let $G(L_1, L_2, ..., L_N)$ represent a homogenous of degree one function identifying potential gains. An equal labor force share across groups will maximize production complementarities; gains are negligible if a single group comprises nearly all of the labor force.

Equation (1) combines these cost and complementarity terms into a single production function where output (Y) experiences constant returns to capital (K) and labor $(L_1, L_2, ..., L_N)$, augmented by an exogenous technology parameter (A). Equation (2) represents average labor productivity (output per worker, y) as a function of capital per worker (k) and each group's proportion of the labor force $(\theta_1, \theta_2, ..., \theta_N)$. An increase in diversity refers to a mix of labor force shares that comes closer to an equal representation of each group in production.

$$Y = F(K, L_1, L_2, ..., L_N) = A \cdot K^{1-\alpha} \cdot C(L_1, L_2, ..., L_N) \cdot G(L_1, L_2, ..., L_N)^{\alpha} \tag{1}$$

$$y = A \cdot k^{1-\alpha} \cdot C(\theta_1, \theta_2, ..., \theta_N) \cdot G(\theta_1, \theta_2, ..., \theta_N)^{\alpha} \tag{2}$$

Clearly, if diversity causes both costs and labor complementarities to increase, one must impose further assumptions on the nature of $C(\cdot)$ and $G(\cdot)$ to predict the theoretical effect on net productivity. Rather than explore these theoretical implications, I advocate empirical analysis to improve understanding. While the majority of prior work has assessed ethnic diversity and economic performance across countries, this paper prefers analysis of racial diversity in the United States as an alternative strategy.

First, the costs and benefits of diversity could vary across national borders. Though major intercultural conflict or warfare is certainly a common occurrence for some countries, others could conceivably exploit complementarities in creativity and problem solving. Collier (2000)

[6] Each option has historical parallels to employing a complementarity term for the variety of machines in production. See Ethier (1982) or Romer (1987), for example.

and Collier (2001) address this potential for cross-country variation in attitudes toward diversity by arguing that diversity is only detrimental to non-democracies, since democracies establish better systems to manage cultural conflict. Alesina and La Ferrara (2005) prefer instead to focus on stages of economic development. To them, diversity negatively affects GDP per capita growth in poor countries but not in their developed counterparts. A third alternative for avoiding the variation in attitudes toward diversity is to simply focus on a single country (e.g., the United States) since cultural sentiments, ethnic strife, and racial tolerance are likely to be more consistent across states or cities within a country than they are across national boundaries. This method also has the obvious advantage of providing insight into a specific country's experience with diversity.[7]

Second, ambiguous ethnicity definitions confound analysis. Ethnicity classifies people according to cultural, linguistic, religious, or national identities. Daniel Posner (2004) criticizes the inherent "grouping problem" associated with cross-country measurement of ethnic demography in that analysts artificially categorize people into groups that do not accurately reflect their identity. Similarly, Fearon (2003) argues that one "must make all manner of borderline-arbitrary decisions, and that in many cases there simply is no single right answer to the question 'What are the ethnic groups in this country?'" Emphasis on the U.S. experience helps reduce this problem, but individuals may still be unable to correctly identify their own ethnicity. Race, by contrast, is more straightforward and generally classifies people according to the continent from which their ancestors descended – European (White), African (Black), Asian, Native American, and so on. Not only can individuals easily identify their own race, but they can often identify the race of others as well. Caselli and Coleman (2002) argue that conflict and reduced labor productivity arise when groups are readily recognizable. On the other hand, ease of group identification also suggests that a person's race will affect his or her experiences. If experiences shape a person's world view, thought processes,

[7]For example, Ottaviano and Peri (2006) find that heterogeneity (based upon immigrants' countries of origin) complements production and boosts native-born wages and productivity in US cities. Similarly, Sparber (2007a) compares racial diversity and gross state product per worker across US states.

etc., people from varied races may have differing ideas, innovative strategies, and approaches to problem solving. Firms wishing to employ a labor force with diverse perspectives may be able to increase productivity by hiring a racially diverse workers.[8]

Finally, U.S. politicians consider many race-based initiatives, including affirmative action policies aimed at promoting racial minorities, on a regular basis. Executives filed the aforementioned *amicus* brief to argue for the merits of racial diversity in the United States. Not only will an empirical exploration of the aggregate effects of increased racial diversity in the U.S. control for variation in attitudes toward diversity and eliminate the grouping problem, but it will provide evidence that can contribute to current U.S. policy debates as well.

3 Methodology and Data

The regressions in this paper will assess the net productivity effect of racial diversity across state-industry cells in 1980, 1990, and 2000. The decennial Census provides pertinent data for an individual's race, wage, industry of employment, and various control variables, which I aggregate to the state-industry level for each Census year. First, I must manipulate the data to make it consistent over time since Census industry codes vary across years. Fortunately, the Integrated Public Use Microdata series (IPUMS) has restored comparability by translating current-year codes into their 1950 equivalents. For 1970-2000 data, this results in creating approximately 150 finely-sorted industries under several broad categorical listings. The detail in most of the 150 industry groups remains overly specific for this paper, as many of the cells would be formed on only a few observations. Instead, I reclassify industries into larger aggregates. In most cases, I simply employ the IPUMS categorical headings. I only abandon the IPUMS scheme when 1950 equivalents to current year industries do not exist. This is relevant for "Computer Manufacturing" and "Computer Software & Design," which

[8] A person's cultural background will also shape his or her experiences in a similar manner. A series of papers by Ottaviano and Peri more thoroughly address the effects of cultural diversity. This paper does not dismiss the importance of culture, but rather assesses whether gains are present when ethnic identities are aggregated to the level of race.

IPUMS lists as subcomponents of "Office and Store Machines and Devices" and "Miscellaneous Business Services," respectively. Specific information regarding the aggregation and refinement of the IPUMS data is available on request. Ultimately, this paper considers the 42 industries in Table 1.

Next, I turn to an appropriate measure of labor productivity. The population Census data does not provide a direct measure of output per worker. Instead, assume that, in the aggregate, firms pay each factor its marginal product. Equation (2) and Euler's theorem clearly demonstrate that the weighted average wage paid to labor is directly proportional to average labor productivity. That is, $w = \frac{\sum_{i=1}^{N} w_i \cdot L_i}{\sum_{i=1}^{N} L_i} = \alpha \cdot y$. The Census does provide individual wage data, which I use to calculate average wages paid to workers for each state-industry cell, my proxy for labor productivity.[9]

To construct diversity indices, one must first determine an appropriate racial classification scheme. I classify individuals into five groups: Asians, Blacks, Hispanics, Whites, and Others.[10] Note that although "Hispanic" most accurately reflects an ethnic description, half of the Hispanic population chose "some other race" or "two or more" races on 2000 Census forms. Moreover, 97% of Census respondents marking "some other race" in 2000 were of Hispanic origin. These anomalies persuade the National Research Council (2004) to argue that research on race in the U.S. should include a separate category for Hispanics despite definitional contravention.

Though Equation (2) expresses production as a non-linear function of each racial group's

[9] Wages averaged across workers of all races could manifest discrimination factors against minorities, and thereby lead to biased estimates of the effects of diversity on society as a whole. An earlier version of this paper adopted the average wage paid to White workers as the proxy for productivity to avoid this problem. However, Sparber (2007b) demonstrates that it is possible for diversity to generate wage gains for Whites while reducing aggregate labor productivity. The model in Equation (2) suggests that average wages paid to all workers is the more appropriate proxy, so this paper adopts the measure.

[10] In 2000, the "Other" category represented less than 3% of the labor force. In this year only, the category includes non-Hispanics of mixed heritage (1.6% of the labor force). Inclusion of this group does not affect qualitative results. I perceive the emergence of this group simply as an evolution in the US understanding of diversity, which will not affect the interpretation of the results. Thus, I include them in this paper's analysis.

Table 1: Racial Fractionalization and Labor Force Share of Industries, 2000.

	Racial Diversity		% of Labor		Racial Diversity		% of Labor
Industry	**Average**	**S.D.**	**Force**	**Industry**	**Average**	**S.D.**	**Force**
Advertising	0.217	0.130	0.42	Insurance	0.248	0.138	1.88
Agriculture	0.300	0.175	2.24	Legal Services	0.183	0.116	1.13
Aircraft & Parts	0.243	0.148	0.51	Mining	0.169	0.150	0.39
Apparel Retail	0.355	0.176	0.64	Motor Vehicles & Engines	0.339	0.168	1.24
Bars & Restaurants	0.401	0.161	5.24	Other Business Services	0.339	0.146	4.96
Colleges & Universities	0.331	0.126	2.11	Other Durables	0.332	0.162	1.60
Computer & Software Design	0.285	0.123	1.16	Other Machinery	0.270	0.159	1.42
Computer Manufacturing	0.346	0.181	0.33	Other Non-Durables	0.381	0.169	0.89
Construction	0.274	0.156	7.00	Other Professional Services	0.314	0.148	2.77
Drugs & Chemicals	0.320	0.150	0.81	Other Retail	0.245	0.145	6.02
Education, Non-College	0.280	0.152	7.68	Paper Products	0.276	0.145	1.82
Electrical Machinery	0.354	0.152	1.39	Personal Services	0.369	0.191	1.76
Engineering & Architecture	0.188	0.121	0.93	Public Administration	0.328	0.155	5.33
Entertainment & Recreation	0.323	0.145	1.97	Raw Durables	0.313	0.176	1.40
Fabricated Metals	0.322	0.156	1.41	Real Estate	0.251	0.135	1.60
Finance	0.279	0.148	3.13	Repair Shops	0.282	0.161	1.71
Food & Beverage Manuf	0.445	0.172	1.17	Telecommunications	0.332	0.158	1.03
Food, Drug, & Alcohol Retail	0.313	0.165	3.09	Textiles	0.405	0.183	0.95
General Retail	0.342	0.169	2.34	Transportation	0.319	0.176	3.83
Health Services	0.328	0.166	8.92	Utilities	0.274	0.152	1.17
Hotels	0.442	0.181	1.06	Wholesale Trade	0.271	0.155	3.54

labor force share, I prefer instead to employ a single aggregate measure of diversity in the empirical analysis. Ideally, such an index would provide a description of the relative size and variety of racial backgrounds present in a society. The most common measure – racial fractionalization (RF) – ranges from zero to one and represents the probability that two people in the labor force, drawn at random, will be of different racial groups.[11] High RF implies the existence of many groups and/or a large minority share of the labor force, and hence captures the variety and size of racial groups in a given industry within a state.[12] For

[11] Mauro (1995), Easterly and Levine (1997), Knack and Keefer (1997), Alesina et al. (2003), and Ottaviano and Peri (2006) employ fractionalization indices. For comparison, Alesina et al. (2003) also uses an index of polarization as a proxy for diversity. (See Lian and Oneal (1997) and Fearon (2003) for more comment). Polarization indices further distinguish whether an observation has high diversity due to the existence of a few equally-represented groups, or instead is characterized by the existence of one large group and several small minority groups. Thus, differences between polarization and fractionalization become more pronounced when the number of groups used in the construction of the indices varies widely across observations. Since fractionalization indices are more common in economics, more easily interpretable, and less subject to criticism when each observation has roughly the same number of racial groups, I use it as my sole measure of diversity.

[12] Since this paper is concerned with long-run aggregate output per worker, I base diversity indices on the labor force. These figures are similar to those based upon employed workers since unemployed individuals must continue to self-identify with a particular industry to be included in the labor force dataset.

a state-industry cell "s, i,"

$$\begin{aligned} RF_{s,i} &= 1 - \sum_{r=1}^{R} \left(Labor\ Force\ Share_{r,s,i}\right)^2, \qquad (3) \\ \text{or, } RF_{s,i} &= 1 - \sum_{r=1}^{R} \left(\frac{LF_{r,s,i}}{T_{s,i}}\right)^2 \\ \text{where } LF_{r,s,i} &= \text{Labor force participants of race } r \text{ in state } s \text{ and industry } i. \\ r &= \{\text{Asians, Blacks, Hispanics, Whites, Others}\}. \\ \text{and } T_{s,i} &= \text{Total labor force in state } s \text{ and industry } i. \end{aligned}$$

I calculate RF indices across the 48 contiguous states and 42 industries (for 48*42=2016 state-industry cells) in 1980, 1990, and 2000.[13] The 2000 Louisiana computer manufacturing industry is the most diverse, with RF=0.729. Several state-industry cells were completely homogeneous. The unweighted mean and standard deviation of racial fractionalization over the period are 0.251 and 0.161, respectively. The weighted mean and standard deviation equal 0.341 and 0.164.

4 Industry-Specific Effects of Diversity

Weighted Least Squares Regression

Most analyses of diversity and productivity assume the effects of diversity are equal across industries. However, if the importance of problem solving, product development, innovation, marketing, and customer service varies across industries, the effects of diversity may vary as well. Thus, sectoral analysis may be more informative than traditional aggregate regressions.

The regression specification in Equation (4) will identify industry-specific diversity coefficients by employing a decennial panel dataset covering 1980 to 2000 with state-industry cells representing the unit of observation. Since small state-industry cells may be associated

[13] Although data from 1970 is available, I use it to construct instrumental variables later in the analysis. The methodology necessitates omission of 1970 data from the empirical model.

with more error, and these cells should not be allowed to drive results, regressions will weight observations by the weighted sum of Census respondents within each cell.[14]

$$\ln(Wage_{s,i,t}) = \alpha_{s,i} + \sum_{i=1}^{42} \beta_i * RF_{s,i,t} + \gamma_1 * Ed_{s,i,t} + \gamma_2 * EdFrac_{s,i,t} + \sum_{t=1990}^{2000} \delta_t * Decade_{s,i,t} + \epsilon_{s,i,t} \quad (4)$$

Where s = 48 contiguous states, $i =$ 42 industries, $t = 3$ decades.

$Wage$ = Average wage earnings of individuals.

$\alpha_{s,i}$ = Unobserved state*industry fixed effects.

RF = Racial fractionalization (diversity variable).

Ed = Average years of schooling.

$EdFrac$ = Educational fractionalization.

$Decade$ = Decade indicator variables for 1990 and 2000.

The model is parsimonious, and includes only a few controls. First, education is a clear determinant of aggregate wages, so productivity regressions must control for it accordingly. Furthermore, since educational attainment is correlated with race, racial diversity and educational diversity will be correlated as well. Failure to control for educational diversity might generate spurious correlation between wages and racial diversity. Thus, I include an educational fractionalization variable ($EdFrac$) to measure the probability that two people, drawn at random, will be of different educational groups.[15] Regressions will also account for wage trends over time by including decade indicator variables, and they exploit the advantages

[14]The typical observation contains aggregate data representing approximately 2900 Census respondents. However, the number of respondents within a cell is highly skewed to the right. More than 400 cells are based upon fewer than 100 respondents, while 86 represent more than 20,000 respondents.

[15]The educational groups include people that have dropped out of school, those that have a high school degree, individuals with some college experience, and a final group of workers who have obtained a bachelors degree or more schooling. Alternative results for regressions allowing $EdFrac$ to enter as a quadratic and/or cubic term, as well as those permitting industry-specific values, are available upon request. These variants have little effect on the qualitative racial diversity results. Columns 7 & 8 of Table 5 also display results for specifications with more comprehensive $EdFrac$ controls.

of the panel dataset by including fixed-effects for state-industry cells over the three decade panel to control for the existence of unobserved time invariant factors specific to states and industries.[16]

The weighted least square results reported in Table 2 cluster on state-industry cells to control for time correlation in standard error calculations. The regression produces 42 industry-specific coefficients, which the table lists in descending order of magnitude, conditional upon being significant at the 10% confidence level. Diversity is positively related to productivity for 26 industries, and negatively related for seven sectors. Positive correlations mark the Legal, Health, and Finance industries – services in which the ability to communicate well with clients is especially important. At the same time, diversity is negatively associated with wages in the more traditional sectors of the economy such as Raw Durables, Fabricated Metals, and Transportation.

Though these results are informative about the industries that might experience gains or losses from diversity, they fail to establish causality. Moreover, omitted variables may be generating spurious correlations. Therefore, more careful analysis of endogeneity and omitted variables is required before fully addressing the magnitude and interpretation of diversity coefficients.

Endogeneity and Omitted Variables

The large number of positive coefficients in Table 2 may indicate that diversity augments productivity, or instead they could demonstrate that productive states and industries simply attract a diverse labor force. To establish the direction of causation, I develop instrumental variables according to a three-step shift-share methodology similar to Card (2001) and Ottaviano and Peri (2006).

In the first step, I begin by recording the labor force demography of the United States in

[16]Suppose that discrimination exists so that minorities earn only a fraction (λ) of their marginal product of labor. Then simple algebra can illustrate that estimation of (4) will disproportionally weight the effects of diversity on the productivity of the majority group, which may be positive or negative.

Table 2: Industry-Specific Marginal Effects of Diversity.

Dependent Variable: ln(Average Wage)			
Instrumental Variables	No		
Time Fixed Effects	d		
Panel Fixed Effects	s*i		
Unit of Observation	s*i*d cells		
Legal Services	2.704	Telecommunications	0.198
	(0.207)***		(0.078)**
Finance	2.102	Education, Non-College	0.237
	(0.230)***		(0.184)
Hotels	1.688	Mining	0.218
	(0.190)***		(0.249)
Health Services	1.685	Utilities	0.128
	(0.269)***		(0.154)
Agriculture	1.267	Food & Beverage Manuf	0.125
	(0.179)***		(0.090)
Other Business Services	1.125	Other Machinery	0.079
	(0.094)***		(0.120)
Electrical Machinery	1.019	Food, Drug, & Alc Retail	0.063
	(0.250)***		(0.083)
Bars & Restaurants	0.983	Aircraft & Parts	0.031
	(0.086)***		(0.133)
Textiles	0.902	Construction	-0.059
	(0.171)***		(0.111)
Insurance	0.893	Colleges & Universities	-0.197
	(0.153)***		(0.139)
Other Durables	0.842	Other Professional Srvcs	-0.210
	(0.136)***		(0.126)*
Entertainment & Recreation	0.809	Public Administration	-0.345
	(0.071)***		(0.201)*
Real Estate	0.755	Fabricated Metals	-0.370
	(0.160)***		(0.220)*
Comp & Software Design	0.662	Raw Durables	-0.389
	(0.142)***		(0.217)*
Advertising	0.634	Motor Vehicles & Engines	-0.396
	(0.119)***		(0.192)**
Other Retail	0.611	Other Non-Durables	-0.511
	(0.074)***		(0.094)***
Repair Shops	0.587	Transportation	-1.168
	(0.085)***		(0.140)***
Personal Services	0.584	Years of Schooling	0.141
	(0.145)***		(0.011)***
Computer Manufacturing	0.562	Education Fractionalization	-0.886
	(0.208)***		(0.117)***
Engineering & Architecture	0.455	Observations	6048
	(0.120)***	R-Squared	0.91
General Retail	0.426		
	(0.114)***		
Drugs & Chemicals	0.354		
	(0.129)***		
Paper Products	0.310		
	(0.141)**		
Wholesale Trade	0.300		
	(0.075)***		
Apparel Retail	0.272		
	(0.093)***		

Cluster-robust standard errors in parenthesis.
Fixed Effects: d = decade, s = state,
i = industry.
*** Coefficient significant at 1%
** Coefficient significant at 5%
* Coefficient significant at 10%
Constant suppressed.

1970. That is, I count the number of labor force participants by race for each state-industry cell of that year. Traditional shift-share instruments simply assume past economic success did not influence the demographic composition of cells in the base year. Indeed, the U.S. map in Figure 1 suggests that this may be true. States with racial fractionalization indices greater than 0.35 are shaded black. Those with RF indices between 0.30 and 0.35 are gray. It appears that two factors determine a state's racial diversity in 1970 – the historical reality of slavery in the U.S. South, and immigration trends bringing large Hispanic populations to the Southwest.

Rather than assume that the 1970 demography is exogenous, I compute the number of Asian, Black, Hispanic, White, and Other workers in each cell that are predicted by

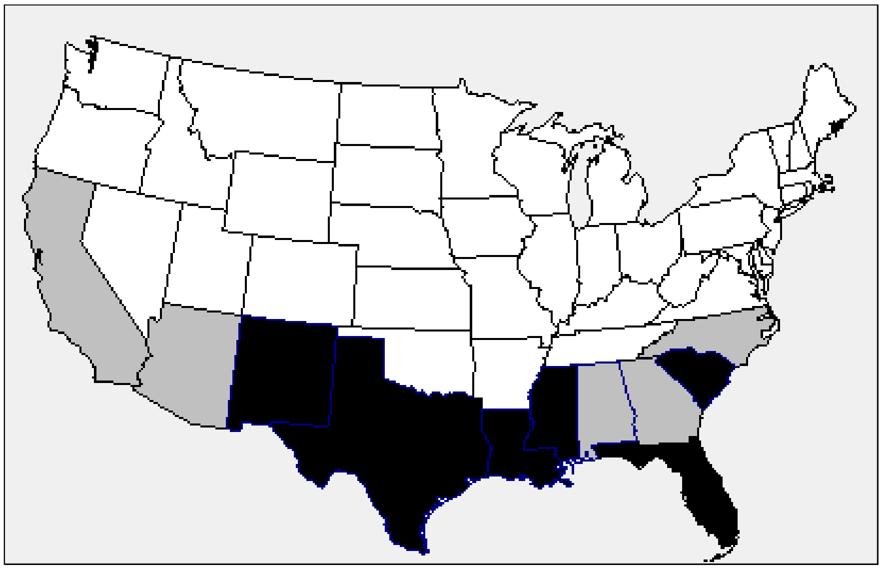

Figure 1: Racial Fractionalization in 1970. States with racial fractionalization indices above 0.35 are shaded black. Those between 0.30 and 0.35 are shaded gray.

exogenous geographic forces. First, I calculate the natural log of one plus the number of workers in each cell, by racial group, in 1970. Then I regress this figure on indicator variables for former Confederate states, Mexican border states, and Pacific Coast states. I also include the natural log of the labor force to control for cell size. The results of these regressions are in the Appendix. The coefficients of these exogenous features then predict the number of workers, by race, that comprised state-industry cells in 1970.

The second step in this modified shift-share methodology requires calculation of the *national* labor force growth rate of each racial group from 1970 to 1980, 1970 to 1990, and 1970 to 2000. I then multiply these national growth rates by the exogenous 1970 state-industry demography constructed in the first step in order to predict state-industry racial composition in subsequent decades.

In the final step, I use the predicted demography to calculate new RF indices, which I employ as instruments for observed values. Altogether, this methodology describes what the U.S. would look like if geography and cell-size determined the 1970 racial composition of state-industry cells, and these populations subsequently grew at their national rates.[17]

Column 1 of Table 3 revisits the regression of the previous subsection. I maintain the assumption that unobserved time invariant factors specific to state-industry cells exist, and continue to control for time trends, education, and educational diversity. Now, however, I adopt a two stage weighted least squares approach, employing predicted racial fractionalization ($RFIV$) values as instruments for their observed counterparts (RF).[18] As before, the coefficients are listed in descending order of magnitude, conditional upon being significant at the 10% confidence level.

The instrumental variable results provide causal evidence for the associations uncovered in the previous section. A one standard deviation increase of racial fractionalization (0.161) in the Legal, Finance, or Health industries – roughly the same increase in diversity experienced by moving from Michigan to Arizona, or from South Dakota to Tennessee[19] – would generate more than a 30% increase in wages. The large coefficient on Advertising (0.661) indicates that a similar shock would increase wages by 10.6%, thus providing evidence that a diverse workforce improves marketing capabilities. High-tech firms also benefit greatly, as the results for computer manufacturing and software demonstrate. In sum, the results lend credence – at least superficially – to the argument that diversity facilitates the creation of new products, aids marketing, and improves customer service.

While we have evidence that 15 industries experience gains from diversity, four traditional sectors (Fabricated Metals, Non-Durables, Raw Durables, and Transportation) see declines. Lack of product differentiation minimizes the necessity of diverse customer service and marketing skills, while creative decision making and innovation are not likely to be vital within these industries. Positive coefficients should not be expected. That the marginal effects are

[17] The standard shift-share methodology requires base-year data to generate instruments for subsequent years. Since no instruments exist for the base year, observations from that year must be dropped in all two stage least squares regressions. I therefore choose to omit 1970 data from all specifications.

[18] Eighteen state-industry cells did not exist in 1970. Instrumental variables are not available for these cells in subsequent years, thus reducing the total number of observations in each regression by 54.

[19] This approximation is based upon the state-level racial diversity of the labor force in 2000. Michigan ($RF = 0.316$) and Tennessee (0.313) were the median states. Arizona (0.476) was the tenth most diverse state, while South Dakota (0.152) was the seventh least diverse.

Table 3: Industry-Specific Marginal Effects of Diversity.

	1	2		1	2
Dependent Variable	ln(Average Wage)	ln(Wage)			
Instrumental Variables	Yes	Yes			
Time Fixed Effects	d	d*s, d*i			
Panel Fixed Effects	s*i	s*i*o			
Unit of Observation	s*i*d cells	Individuals			
Legal Services	2.593	2.647	**Paper Products**	0.016	1.286
	(0.474)***	(0.437)***		(0.486)	(0.656)*
Finance	2.246	1.709	**Food, Drug, & Alc Retail**	-0.081	-0.155
	(0.342)***	(0.342)***		(0.314)	(0.495)
Health Services	2.045	2.467	**Education, Non-College**	-0.101	1.749
	(0.636)***	(0.322)***		(0.672)	(0.697)**
Hotels	1.976	1.410	**Other Professional Srvcs**	-0.252	1.256
	(0.443)***	(0.647)**		(0.400)	(0.235)***
Electrical Machinery	1.314	1.776	**Colleges & Universities**	-0.462	0.364
	(0.483)***	(0.477)***		(0.346)	(0.549)
Other Business Services	1.121	1.024	**Construction**	-0.564	-0.018
	(0.263)***	(0.157)***		(0.473)	(0.619)
Computer Manufacturing	1.011	1.573	**Public Administration**	-0.709	1.329
	(0.238)***	(0.635)**		(0.507)	(0.241)***
Insurance	0.925	1.381	**Other Machinery**	-0.757	0.239
	(0.366)**	(0.247)***		(0.716)	(1.218)
Bars & Restaurants	0.869	0.591	**Motor Vehicles & Engines**	-1.280	0.813
	(0.347)**	(0.622)		(0.927)	(2.100)
Other Durables	0.845	1.390	**Mining**	-1.904	-1.050
	(0.464)*	(0.751)*		(1.470)	(1.731)
Comp & Software Design	0.794	1.707	**Other Non-Durables**	-0.810	-0.315
	(0.210)***	(1.059)		(0.359)**	(0.392)
Entertainment & Recreation	0.749	0.816	**Raw Durables**	-1.170	-0.017
	(0.262)***	(0.197)***		(0.669)*	(1.417)
Advertising	0.661	0.882	**Transportation**	-1.566	-0.510
	(0.256)***	(0.348)**		(0.375)***	(0.396)
Real Estate	0.646	1.323	**Fabricated Metals**	-2.517	-0.731
	(0.312)**	(0.197)***		(1.201)**	(2.182)
Other Retail	0.490	0.599	**Years of Schooling**	0.119	0.058
	(0.292)*	(0.409)		(0.021)***	(0.001)***
Agriculture	2.394	4.662	**Education Fractionalization**	-0.882	-0.549
	(1.621)	(6.151)		(0.237)***	(1.102)
Textiles	1.774	1.091	**Age**		0.011
	(1.421)	(9.144)			(0.000)***
Personal Services	0.574	2.770	**Reside in Metroarea Indicator**		0.113
	(0.528)	(1.461)*			(0.002)***
Drugs & Chemicals	0.459	1.587	**Female Indicator**		-0.334
	(0.332)	(0.295)***			(0.003)***
Repair Shops	0.429	0.336	**Foreign-Born Indicator**		0.032
	(0.374)	(0.568)			(0.005)***
General Retail	0.358	0.248	**Asian Indicator**		-0.064
	(0.329)	(0.489)			(0.007)***
Engineering & Architecture	0.310	1.164	**Black Indicator**		-0.045
	(0.389)	(0.553)**			(0.004)***
Telecommunications	0.199	0.617	**Hispanic Indicator**		-0.023
	(0.267)	(0.616)			(0.005)***
Apparel Retail	0.189	0.173	**Other Minority Race Indicator**		-0.047
	(0.234)	(0.342)			(0.007)***
Wholesale Trade	0.124	0.587	**Observations**	5994	1514920
	(0.339)	(0.462)	**R-Squared**	0.90	0.13
Food & Beverage Manuf	0.096	0.505			
	(0.437)	(1.156)			
Utilities	0.077	1.496			
	(0.554)	(0.687)**			
Aircraft & Parts	0.059	1.411			
	(0.310)	(0.311)***			

Cluster-robust standard errors in parenthesis.
Fixed Effects: d = decade, s = state, i = industry, o = occupation.
*** Coefficient significant at 1%
** Coefficient significant at 5%
* Coefficient significant at 10%
Constant suppressed.

negative, however, indicates that diversity causes a diversion from productive activity. This outcome resembles the costs envisioned by the international literature.

While the signs on the productivity effects of diversity are intuitively appealing, the magnitudes are uncomfortably large. Individual-level analysis provides an alternative estimation strategy that may prove to be more informative, although such specifications are less consistent with theoretical models of diversity and productivity.

First, it may be that baseline regressions exhibit large omitted variables bias. Individual-level regressions can control for an increased number of factors without overwhelming the data. Second, by accounting for individual-level heterogeneity, regressions should be able to ascertain whether diversity directly complements production or instead simply attracts the most talented workers. Assume that individuals who are identical in their race, sex, age, occupation, native citizenship, choice to live in an urban environment, state of residence, and industry of employment also have homogeneous talents. By using individual-level regressions accounting for these talents, positive diversity coefficients will indicate that complementarities between races exist, and that productivity gains are not due to self-selection of talented individuals into diverse state-industry cells. Third, individual-level regressions permit a richer set of fixed effects. Expanded time-invariant fixed effects for state-industry-occupation indicators can provide a more rigorous control for individual unobservables. More importantly, the increased number of observations associated with individual-level regressions allow decade*state and decade*industry terms to control for state-specific or industry-specific time trends.

The second column of Table 3 presents a regression of the natural log of an individual's annualized wage on the aforementioned controls and the diversity of the state-industry cell in which he/she works. First, I sample 10% of the labor force that earned a wage in the year of observation, dropping all individuals working in state-industry cells in which instruments are unavailable. I then convert wages to reflect a person's yearly wage earnings, assuming a 52-week work year. That is, $Wage = Annual\ Wage\ Earnings \cdot \left(\frac{52}{Weeks\ Worked}\right)$. Schooling and age variables should control for variation in an individual's skills, while gender, race, and foreign-born indicators will reveal a combination of discrimination and ability gaps across groups in the U.S.[20]

This alternative methodology does little to alter qualitative results. Racial diversity generates productivity gains for half of the industries. The largest effects continue to occur in sectors employing creative decision makers. A one standard deviation increase in diversity causes wages to rise by more than 40% within Legal Services and Health services, while the same increase in diversity would lead to wage gains in excess of 25% for Computer Manufacturing and Finance. Unlike aggregate-level specifications, however, individual-level

[20] See Neal and Johnson (1996), Neumark (1998), Hellerstein, Neumark, and Troske (1999), and Hellerstein and Neumark (2004) for extensive analysis of discrimination and ability gaps across racial groups.

results fail to identify any industry that experiences significant losses from diversity.

5 Industry Characteristics and the Effects of Diversity

The results of the previous section, prior academic research, and the anecdotes expressed in the *amicus* brief provide guidance for better understanding of the characteristics of industries that might complement or hinder a diverse workforce. *A priori*, past evidence suggests that diversity may increase productivity in sectors that require creative decision-making, problem solving, and customer service, but decrease it in industries requiring common action or group effort. The difficulty in assessing these possibilities is one of measurement.

The U.S. Department of Labor's O*NET database has become an invaluable source of information about occupational characteristics that economists can use to analyze workforce issues. O*NET administers surveys to occupational analysts, experts, and incumbents that ask hundreds of questions about workforce characteristics, knowledge, and activities. Four questions are appropriate for this study. First, O*NET's *Work Activities* survey asks, "How important is making decisions and solving problems to the performance of your current job?" and "How important is thinking creatively to the performance of your current job?" These questions seem broadly consistent with notions of creative decision-making. Similarly, O*NET's *Knowledge* survey question "How important is customer and personal service knowledge to the performance of your current job?" is a reasonable proxy for customer service intensity. For a measure of common effort, I turn to *Work Context* survey question, "How important are interactions that require you to work with or contribute to a work group or team to perform your current job?"

O*NET records the average response to each of these questions, on a scale of one to five, for more than 800 occupations that roughly correspond to the 2000 Census occupation classification codes. I rescale the data into percentile values so that the median worker earned a value of 0.5 for each of these categories in 2000.[21] Table 10 in the Appendix

[21] Autor, Levy, and Murnane (2003) and Peri and Sparber (2007) perform a similar rescaling for O*NET's

provides examples of occupations receiving the highest, median, and lowest scores in 2000.

Though Census occupation codes vary across years, the IPUMS variable occ1950 facilitates comparability over time. I assign each occ1950 occupation an O*NET percentile value based upon the a weighted average of the year 2000 occupation scores that comprise these categories. I then use these values to construct industry-specific average O*NET values for each year. Industries employing many workers involved in decision-making, creative-thinking, customer service, or group-effort intensive occupations earn high scores for the respective variables. Table 4 provides (unweighted) summary statistics, and Table 11 in the Appendix lists the industry-specific values for 2000.

Table 4: Summary of Industry Characteristic Data, 1980-2000.

Importance of:	*Mean*	*Std Deviation*	*Minimum*	*Maximum*
Making Decisions and Solving Problems	0.499	0.094	0.252 (1980 Hotels)	0.765 (2000 Computer & Software Design)
Thinking Creatively	0.503	0.103	0.288 (1980 Bars & Restaurants)	0.782 (2000 Computer & Software Design)
Customer and Personal Service	0.433	0.129	0.213 (1980 Textiles)	0.695 (2000 Health Services)
Work With Work Group or Team	0.440	0.083	0.143 (1980 Education, Non-College)	0.631 (1990 Engineering & Architecture)

Rather than estimate the industry-specific marginal effects of diversity, the regressions in Table 5 analyze the relationship between industry characteristics and the effects of diversity.[22] Columns 1 and 2 are analogous to the results in Table 2, and present WLS results with state-industry cells representing the unit of observation. Columns 3 and 4 adopt the same unit of observation, but instead perform two-stage WLS regressions. Columns 5 and 6 use individual-level regressions similar to those of Column 2 in Table 3. The final two columns repeat the individual-level specifications, but also allow the effects of educational fractionalization to vary across industries. (The output suppresses these 42 industry-specific coefficients).

Though the magnitudes of the coefficients vary across specifications, the qualitative re-

predecessor, the *Dictionary of Occupational Titles*.

[22]Note that while diversity continues to be measured for state*industry cells across time, the O*NET variables only vary across industries and time.

Table 5: Industry Characteristics and the Marginal Effects of Diversity.

	1	2	3	4	5	6	7	8
Dependent Variable:	***ln(Average Wage)***		***ln(Average Wage)***		***ln(Wage)***		***ln(Wage)***	
Instrumental Variables	No		Yes		Yes		Yes	
Time Fixed Effects	d		d		d*s, d*i		d*s, d*i	
Panel Fixed Effects	s*i		s*i		s*i*o		s*i*o	
Unit of Observation	s*i*d cells		s*i*d cells		Individuals		Individuals	
Racial Fractionalization	-1.254	-0.823	-1.181	-0.616	-0.750	-0.641	0.554	0.277
	(0.212)***	(0.296)***	(0.696)*	(0.712)	(0.374)**	(0.345)*	(0.685)	(0.504)
RF * Making Decisions	2.434		2.600		3.227		2.669	
	(0.421)***		(0.555)***		(0.454)***		(0.703)***	
RF * Thinking Creatively		1.404		1.291		2.002		2.283
		(0.384)***		(0.501)**		(0.341)***		(0.385)***
RF * Customer Service	3.038	2.681	3.723	3.133	3.428	2.841	1.856	1.764
	(0.313)***	(0.318)***	(0.500)***	(0.509)***	(0.414)***	(0.399)***	(0.649)***	(0.539)***
RF * Work with Group	-1.688	-1.129	-2.236	-1.236	-3.397	-1.903	-4.099	-3.258
	(0.551)***	(0.403)***	(0.577)***	(0.517)**	(0.492)***	(0.427)***	(0.545)***	(0.569)***
Years of Schooling	0.148	0.165	0.160	0.181	0.058	0.058	0.058	0.058
	(0.013)***	(0.011)***	(0.021)***	(0.021)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***
Education Fractionalization	-0.618	-0.851	-0.640	-0.907	-0.524	-0.738	[Industry-Specific]	
	(0.109)***	(0.108)***	(0.140)***	(0.129)***	(0.096)***	(0.075)***		
Age					0.011	0.011	0.011	0.011
					(0.000)***	(0.000)***	(0.000)***	(0.000)***
Reside in Metroarea Indicator					0.113	0.113	0.114	0.114
					(0.002)***	(0.002)***	(0.002)***	(0.002)***
Female Indicator					-0.334	-0.334	-0.334	-0.334
					(0.003)***	(0.003)***	(0.003)***	(0.003)***
Foreign-Born Indicator					0.032	0.032	0.031	0.031
					(0.005)***	(0.005)***	(0.005)***	(0.005)***
Asian Indicator					-0.063	-0.062	-0.063	-0.062
					(0.006)***	(0.006)***	(0.006)***	(0.006)***
Black Indicator					-0.045	-0.045	-0.044	-0.044
					(0.004)***	(0.004)***	(0.004)***	(0.004)***
Hispanic Indicator					-0.022	-0.023	-0.023	-0.023
					(0.005)***	(0.005)***	(0.005)***	(0.005)***
Other Minority Race Indicator					-0.047	-0.047	-0.047	-0.046
					(0.006)***	(0.006)***	(0.006)***	(0.006)***
Observations	6048	6048	5994	5994	1514920	1514920	1514920	1514920
R-Squared	0.88	0.87	0.88	0.88	0.13	0.13	0.13	0.13

Cluster-robust standard errors in parenthesis.
Fixed Effects: d = decade, s = state, i = industry, o = occupation.
*** Coefficient significant at 1%
** Coefficient significant at 5%
* Coefficient significant at 10%
Columns 7 and 8 allow for 42 industry-specific coefficients on Education Fractionalization, which have been suppressed for brevity.
Constant suppressed.

sults are remarkably robust to the different methodologies. Creative decision-making and customer service complement diversity, but diversity conflicts with group effort. Thus, the results reconcile competing literatures by recognizing both the costs and benefits of diversity.

Table 6: Average Effects of a One Standard Deviation Increase in Diversity.

	1	2	3	4	5	6	7	8
Dependent Variable:	***ln(Average Wage)***		***ln(Average Wage)***		***ln(Wage)***		***ln(Wage)***	
Instrumental Variables	No		Yes		Yes		Yes	
Time Fixed Effects	d		d		d*s, d*i		d*s, d*i	
Panel Fixed Effects	s*i		s*i		s*i*o		s*i*o	
Unit of Observation	s*i*d cells		s*i*d cells		Individuals		Individuals	
Estimated Effect of a One Standard Deviation (0.161) Increase in Racial Diversity								
Effect Assuming Average Industry Characteristics	0.086	0.088	0.120	0.136	0.137	0.122	0.143	0.122
Additional Effect of a One Standard Deviation Increase in Given Industry Characteristics								
Making Decisions	0.037		0.039		0.049		0.040	
Thinking Creatively		0.023		0.021		0.033		0.038
Customer Service	0.063	0.056	0.077	0.065	0.071	0.059	0.039	0.037
Work with Group	-0.023	-0.015	-0.030	-0.017	-0.045	-0.025	-0.055	-0.044
Observations	6048	6048	5994	5994	1514920	1514920	1514920	1514920
R-Squared	0.88	0.87	0.88	0.88	0.13	0.13	0.13	0.13

Table 6 provides more intuitive interpretation for the size of the effects estimated in Table 5. The first row indicates that a one standard deviation increase in racial diversity

(0.161) would increase wages by roughly 9% to 14% for an industry characterized by average decision-making, creativity, customer service, and team work intensity. The next four rows demonstrate the additional effect of a one standard deviation increase in a particular industry characteristic for the same diversity shock. For example, a one standard deviation (0.094) increase in the importance of making decisions and solving problems – the difference between Other Machinery (the median industry) and Health Services in 2000[23] – would facilitate an additional 4% wage gain from the diversity shock. If instead creative decision-making is measured by the importance of thinking creatively, a one standard deviation (0.103) rise would cause the diversity shock to generate an additional 2-3% wage increase. Results for customer service are larger. The effects of the diversity shock improve 6-7% for a 0.129 rise in service importance – the difference between Education and Personal Services – though the estimates are smaller if the coefficient on educational diversity is allowed to vary across industries. In contrast to these results, however, group effort conflicts with diversity and serves to reduce productivity. The magnitude of these losses is comparable to the gains generated by complements between diversity and creative thinking.[24]

6 Robustness – Labor Supply and Demand

Equation (2) and diminishing marginal returns to labor imply that reduced labor supply will cause wages to rise. This paper has implicitly assumed that labor supply is fixed. However, many analysts recognize the possibility that diversity could alter labor supply if White workers respond to diversity by leaving the labor force, though empirics usually fail to uncover evidence that such an effect truly exists.[25] To ensure wage effects from aggregate regressions do not solely reflect supply changes, I now consider estimation of labor supply as a

[23] For further examples of industries separated by one standard deviation in a given skill, see the values in Table 11 of the Appendix.

[24] A previous version of this paper (available upon request) measured creative-decision making, customer service, and common action intensity with the share of workers with a college degree, working in customer service, and belonging to a union, respectively. In that version, diversity complemented creativity and conflicted with unionization.

[25] See Becker (1971), Marsden (1988), and Ottaviano and Peri (2006).

share of the total labor force in the U.S. according to Equation (5). Except for the dependent variable, the labor force specification is identical to the wage regression in Equation (4).

$$\frac{100 * LF_{s,i,t}}{\sum_{s=1}^{48}\sum_{i=1}^{42} LF_{s,i,t}} = \alpha^l_{s,i} + \sum_{i=1}^{42} \beta^l_i * RF_{s,i,t} + \gamma^l_1 * Ed_{s,i,t} + \gamma^l_2 * EdFrac_{s,i,t} + \sum_{t=1990}^{2000} \delta^l_t * Decade_{s,i,t} + \epsilon^l_{s,i,t} \tag{5}$$

Where s = 48 contiguous states, i = 42 industries, $t = 3$ decades.

$Wage$ = Average wage earnings of individuals.

LF = Labor force.

$\alpha^l_{s,i}$ = Unobserved fixed effects correlated with the regressors.

RF = Racial fractionalization (diversity variable).

Ed = Average years of schooling.

$EdFrac$ = Educational fractionalization.

$Decade$ = Decade indicator variables for 1990 and 2000.

If diversity bolsters productivity, the labor demand curve will shift out, increasing both the equilibrium quantity of labor supplied and wages. If utility also depends upon diversity, however, the labor supply curve will shift as well. Empirically, the change in wages and labor force captures the net effect of these shifts, making it impossible to ascertain the magnitude of productivity gains or losses. By estimating both the regression in (4) and (5) with instruments outlined in Section 4, however, I may be able to identify the sign of the productivity shift for some industries. Assuming that labor demand is not perfectly inelastic, we can be sure that diversity increases productivity if the sign on racial fractionalization in either the wage or labor supply regression is positive and the sign in the other is non-negative. The converse is also true – if the sign in one regression is negative and the other is non-positive, diversity causes productivity to decrease. Ambiguity exists only when diversity has opposite effects on wages and the size of the labor force. In these cases, the productivity consequences could be positive, negative, or nonexistent.

Table 7 displays the marginal effects of racial diversity for only those industries in which productivity consequences can be ascertained. Racial diversity increases productivity in at least 14 industries. These continue to be the sectors in which creative decision-making plays a dominant role. Seven industries continue to exhibit adverse consequences. As in prior specifications, losses occur in the traditional sectors of the economy – Fabricated Metals, Raw Durables Manufacturing, and Mining, for example. Creative decision making may not play an important role within these industries, and instead conflict, rent-seeking behavior, and costs of common action dominate.

Table 7: Industry-Specific Marginal Effects of Diversity – Industries in which Diversity Positively or Negatively Affects Productivity.

Industry-Specific Marginal Effects of Diversity
Dependent Variables:
ln(Average Wage)
100*($LF_{s,i,t}$)/(Total LF_t)

IV	Yes	Yes
Decade Fixed Effects	Yes	Yes
State*Industry Fixed Effects	Yes	Yes

Positively affected by diversity.	Wage	LF Shr
Advertising	0.661	-0.256
	(0.256)***	(0.292)
Bars & Restaurants	0.869	-0.226
	(0.347)**	(0.379)
Comp & Software Design	0.794	0.239
	(0.210)***	(0.246)
Computer Manufacturing	1.011	-0.257
	(0.238)***	(0.202)
Entertainment & Rec Srvc	0.749	-0.111
	(0.262)***	(0.318)
Finance	2.246	-0.184
	(0.342)***	(0.309)
Health Services	2.045	0.064
	(0.636)***	(0.504)
Hotels	1.976	-0.379
	(0.443)***	(0.400)
Insurance	0.925	-0.486
	(0.366)**	(0.396)
Legal Services	2.593	-0.306
	(0.474)***	(0.376)
Other Business Services	1.121	0.528
	(0.263)***	(0.471)
Other Durables	0.845	-0.515
	(0.464)*	(0.502)
Other Retail	0.490	-0.108
	(0.292)*	(0.346)
Real Estate	0.646	-0.376
	(0.312)**	(0.315)

Negatively affected by diversity.	Wage	LF Shr
Aircraft & Parts	0.059	-0.867
	(0.310)	(0.377)**
Fabricated Metals	-2.517	-2.823
	(1.201)**	(1.331)**
Other Machinery	-0.757	-1.649
	(0.716)	(0.771)**
Other Non-Durables	-0.810	-0.351
	(0.359)**	(0.398)
Paper Products	0.016	-0.949
	(0.486)	(0.534)*
Raw Durables	-1.170	-1.085
	(0.669)*	(0.763)
Transportation	-1.566	-0.307
	(0.375)***	(0.337)
Years of Schooling	0.119	-0.040
	(0.021)***	(0.023)*
Education Fractionalization	-0.882	0.048
	(0.237)***	(0.212)
Observations	5994	5994

Panel covers 48 contiguous US states and 42 industries in 1980, 1990, and 2000.
Unit of observation: state-industry cells.
Diversity measured as racial fractionalization (RF).
Cluster-robust standard errors in parenthesis.
*** Coefficient significant at 1%
** Coefficient significant at 5%
* Coefficient significant at 10%
Constant and indicator variables suppressed.

7 Conclusions

The *amicus* brief filed by thirty *Fortune* 500 companies endorsing the University of Michigan's affirmative action admissions policies claims that a diverse workforce bolsters creative decision making, product development, customer service, and marketing. Many social scientists concur with these beliefs, yet international growth economists remain skeptical and argue that diversity exacerbates rent-seeking activity and the costs of common action. This paper maintains that unqualified statements regarding the costs and merits of diversity are inappropriate for the U.S. economy, as racial heterogeneity increases productivity in many, but not all, industries.

Two-stage weighted least squares regressions at the aggregate-level imply that a one standard deviation increase in diversity would raise wages by 42% in Legal Services, 16% in Computer Manufacturing, 13% in Computer Software, and 11% in Advertising. Though large, individual-level regressions controlling for selection bias, state and industry specific time trends, and individual characteristics reinforce baseline results. The gains from diversity are sizeable and economically relevant. Losses envisioned by the international growth literature occur in only seven industries and are concentrated in traditional sectors of the economy including Mining, Raw Durables, Fabricated Metals, and Transportation.

O*NET data on occupational characteristics identifies the features that determine whether industries benefit or suffer from diversity. Regressions show that diversity bolsters productivity and wages in industries employing many creative decision makers. This is true whether regressions employ the O*NET variable measuring the necessity of making decisions and solving problems, or instead uses a variable that accounts for the importance of thinking creatively. These regressions also show that diversity complements customer service. Heterogeneity reduces productivity in sectors requiring high levels of group effort and common action, however, just as the international literature would predict.

This paper creates numerous possibilities for continued research across several economic fields. Most importantly, aggregate data cannot describe the level of diversity existing in a person's workplace. Instead, behavioral economists should assess how diversity changes group behavior, and how these changes could influence economic outcomes. Similarly, microeconomists could employ plant-level data to analyze how diversity affects the productivity of firms.

International economists will ask whether these results are unique to the United States, or if they may be applicable to other countries as well. Recent events in France and Australia suggest that even developed democracies have difficulties managing diversity, but empirics are required to measure the net effects. If the U.S. is different, why is this the case? What institutions are in place to help the United States reap the benefits of diversity? Are meritocratic norms and rewards to innovation and creativity better established in the U.S., or are other factors more essential?

Diversity complements creativity in generating productivity gains, so policy makers may want to increase the level of diversity in creativity-intensive sectors. This paper does not, however, take a stance on the efficacy of affirmative action or various immigration policies in achieving this goal. Public economists certainly have great interest in such issues.

Contrasting literatures exist highlighting the costs and benefits of racial diversity. This paper reconciles these views by illustrating that diversity generates both gains and losses for the United States. Racial heterogeneity bolsters creative decision making and customer service, but exacerbates losses associated with common action. I encourage further exploration of these issues and their effect on the U.S. and world economy.

References

[1] Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public Goods and Ethnic Divisions," Policy Research Working Paper Series 2108, The World Bank.

[2] Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg. 2003. "Fractionalization," *Journal of Economic Growth*, Vol. 8(2): 155-194.

[3] Alesina, Alberto and Edward L. Glaeser. 2004. *Fighting Poverty in the U.S. and Europe*. New York: Oxford University Press.

[4] Alesina, Alberto and Eliana La Ferrara. 2005. "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, Vol. 43(3): 762-800.

[5] Autor David H., Frank Levy, and Richard Murnane. 2003. "The Skill Content of Recent Technological Change: an Empirical Exploration," *The Quarterly Journal of Economics*, Vol. 118 (4): 1279-1333.

[6] Becker, Gary S. 1971. *The Economics of Discrimination*. Chicago: University of Chicago Press.

[7] Brief of *Amici Curiae* Steelcase, Inc. et al., Gratz v. Bollinger, 135 F.Supp.2d 790. 2001. No. 97-75231.

[8] Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, Vol. 19: 22-61.

[9] Campbell, Donald. 1960. "Blind Variation and Selective Retention in Creative Thought as in Other Knowledge Processes," *Psychological Review*, Vol. 67: 380-400.

[10] Caselli, Francesco and John Coleman. 2002. "On the Theory of Ethnic Conflict," Working Paper, Harvard University.

[11] Collier, Paul. 2000. "Ethnicity, Politics, and Economic Performance," *Economics and Politics*, Vol. 12(3): 225-246.

[12] Collier, Paul. 2001. "Implications of Ethnic Diversity," *Economic Policy: A European Forum*, Vol. 0(32): 127-155.

[13] Cox, Taylor H., Sharon A. Lobel, and Poppy Lauretta McLeod. 1991. "Effects of Ethnic Group Cultural Differences on Cooperative and Competitive Behavior on a Group Task," *Academy of Management Journal*, Vol. 34(4): 827-847.

[14] Easterly, William and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions," *The Quarterly Journal of Economics*, Vol. 112(4): 1203-1250.

[15] Ethier, Wilfred J. 1982. "National and International Returns to Scale in the Modern Theory of International Trade," *American Economic Review*, Vol. 72(3): 389-405.

[16] Fearon, James D. 2003. "Ethnic and Cultural Diversity by Country," *Journal of Economic Growth*, Vol. 8(2): 195-222.

[17] Florida, Richard. 2002. *The Rise of the Creative Class*. New York: Basic Books.

[18] Fukuyama, Francis. 1999. "Social Capital and Civil Society," Prepared for the IMF Conference on Second Generation Reforms.

[19] Gilens, Martin. 1999. *Why Americans Hate Welfare*. Chicago: The University of Chicago Press.

[20] Hellerstein, Judith and David Neumark. 2004. "Production Function and Wage Equation Estimation with Heterogeneous Labor: Evidence from a New Matched Employer-Employee Data Set," NBER Working Paper No. 10325.

[21] Hong, Lu and Scott E. Page. 2004. "Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers," *Proceedings of the National Academy*, Vol. 101(46): 16385-16389.

[22] Knack, Stephen and Philip Keefer. 1997. "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," *The Quarterly Journal of Economics*, Vol. 112(4): 1251-1273.

[23] Lian, Brad and John R. Oneal. 1997. "Cultural Diversity and Economic Development: A Cross-National Study of 98 Countries, 1960-1985," *Economic Development and Cultural Change*, Vol. 46(1): 61-77.

[24] Luttmer, Erzo F.P. 2001. "Group Loyalty and the Taste for Redistribution," *Journal of Political Economy* 109(3): 500-528.

[25] Marsden, Peter. 1988. "Homogeneity in Confiding Relations," *Social Networks*, Vol. 10: 57 – 76.

[26] Mauro, Paolo. 1995. "Corruption and Growth," *The Quarterly Journal of Economics*, Vol. 110(3): 681-712.

[27] National Research Council. 2004. *Measuring Racial Discrimination*. Panel on Methods for Assessing Discrimination, Rebecca M. Blank, Marilyn Dabady, and Constance F. Citro, Editors. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

[28] Neal, Derek and William Johnson. 1996. "The Role of Premarket Factors in Black-White Wage Differences," *Journal of Political Economy* 104(5): 869-895.

[29] Neumark, David. 1998. "Labor Market Information," NBER Working Paper No. 6573.

[30] O*NET 11.0 Database [Database]. U.S. Department of Labor, National O*NET Consortium. Available at: http://www.onetcenter.org/database.html.

[31] O'Reilly, Charles A. III, Katherine Y. Williams, and Sigal Barsade. 1998. "Group Demography and Innovation: Does Diversity Help?," *Research on Managing Groups and Teams*, Vol. 1: 183-297.

[32] Ottaviano, Gianmarco I.P., and Giovanni Peri. 2005. "Cities and Cultures," *Journal of Urban Economics*, Vol. 58(2): 304-337.

[33] Ottaviano, Gianmarco I.P., and Giovanni Peri. 2006. "The Economic Value of Cultural Diversity: Evidence from U.S. Cities," *Journal of Economic Geography*, Vol. 6(1): 9-44.

[34] Page, Scott E. 2007. *The Difference – How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies.* Princeton, NJ: Princeton University Press.

[35] Peri, Giovanni and Chad Sparber. 2007. "Task Specialization, Comparative Advantages, and the Effects of Immigration on Wages," NBER Working Paper No. 13389.

[36] Posner, Daniel N. 2004. "Measuring Ethnic Fractionalization in Africa," *American Journal of Political Science*, Vol. 48(4): 849-863.

[37] Poterba, James M. 1997. "Demographic Structure and the Political Economy of Public Education," *Journal of Policy Analysis and Management*, Vol. 16(1): 48-66.

[38] Romer, Paul M. 1987. "Growth Based on Increasing Returns Due to Specialization," *American Economic Review*, Vol. 77(2): 56-62.

[39] Ruggles, Steven, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander. *Integrated Public Use Microdata Series: Version 3.0* [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2004. http://www.ipums.org.

[40] Simonton, Dean Keith. 1999. *Origins of Genius, Darwinian Perspectives on Creativity.* New York: Oxford University Press.

[41] Shleifer, Andrei and Robert W. Vishny. 1993. "Corruption," *The Quarterly Journal of Economics*, Vol. 108(3): 599-617.

[42] Sparber, Chad. 2007a. "Racial Diversity and Macroeconomic Productivity across U.S. States and Cities," Colgate University Working Paper 2007-01.

[43] Sparber, Chad. 2007b. "A Theory of Racial Diversity, Segregation, and Productivity," *Journal of Development Economics*, In press.

A Appendix

Exogeneity of Instruments

For the instruments in the analysis to be valid, 1970 state-industry demography must be exogenous. To ensure this, I estimate the number of individuals, by race, who would have been working in each cell if 1970 demography had solely been determined by geographic features and the overall size of the cells. To guarantee that predictions will be positive, the dependent variable equals the natural log of one plus the number of people of a given race. The geographic explanatory variables include an indicator for former Confederate states (to account for large Black populations), Mexican border states (to account for Hispanic populations), and Pacific Coast states (to account for Asian populations). Table 8 displays the coefficients, which are then used to predict exogenous demography in 1970, the necessary first step in the creation of the shift-share instruments described in the text.

Table 8: Determinants of Racial Demography in 1970.

Dependent Variable: ln(1 + Number of Workers in Given Racial Group)

	Asians	Blacks	Hispanics	Whites	Others
Border	0.724	-0.119	2.901	-0.163	1.713
	(0.171)***	(0.164)	(0.164)***	(0.016)***	(0.175)***
Former Confederate State	-0.956	2.145	-0.055	-0.103	-1.089
	(0.106)***	(0.101)***	(0.101)	(0.010)***	(0.109)***
West Coast	2.449	-0.192	-0.560	0.034	0.870
	(0.198)***	(0.190)	(0.190)***	(0.018)*	(0.204)***
ln(Labor Force)	0.792	1.593	1.426	0.993	0.763
	(0.028)***	(0.027)***	(0.027)***	(0.003)***	(0.029)***
Constant	-5.793	-10.470	-9.293	-0.011	-5.225
	(0.267)***	(0.256)***	(0.256)***	(0.025)	(0.274)***
Observations	1998	1998	1998	1998	1998
R-squared	0.38	0.69	0.63	0.99	0.33

Standard errors in parentheses.
* significant at 10%; ** significant at 5%; *** significant at 1%

First Stage Coefficients

The two-stage weighted least squares regression in Column 1 of Table 3 has 42 endogenous variables and an equal number of instruments. A table presenting all coefficients for these 42 first-stage regressions would be overwhelming. Instead, Table 9 provides the coefficient of an industry's $RFIV$ in first-stage regressions of its own observed RF value, the standard error of this estimate, and the F-statistic for a test of joint significance of the instrumental variables. Coefficients equalling one imply that predicted and observed values are identical. Empirically, this does not occur, though standard errors indicate that the instruments are strong predictors of observed diversity.

Table 9: First Stage Results of IV Regression.

	Coefficient on Own Industry IV	Standard Error	F(46, 5947)
Advertising	1.267	(0.011)***	326
Agriculture	0.591	(0.019)***	33
Aircraft & Parts	0.797	(0.008)***	234
Apparel	1.804	(0.016)***	284
Bars & Restaurants	1.615	(0.015)***	411
Colleges & Universities	1.044	(0.009)***	441
Comp & Software Design	1.109	(0.011)***	280
Computer Manufacturing	1.021	(0.012)***	172
Construction	1.180	(0.014)***	275
Drugs & Chemicals	1.034	(0.010)***	263
Education, Non-College	0.866	(0.010)***	265
Electrical Machinery	1.129	(0.009)***	342
Engineering & Architecture	0.767	(0.007)***	317
Entertainment & Recreation	1.412	(0.010)***	496
Fabricated Metals	0.541	(0.010)***	73
Finance	1.186	(0.010)***	371
Food & Beverage Manuf	1.151	(0.018)***	102
Food, Drug, & Alc Retail	1.692	(0.015)***	384
General Retail	1.532	(0.013)***	385
Health Services	0.934	(0.010)***	320
Hotels	1.108	(0.013)***	166
Insurance	1.073	(0.009)***	347
Legal Services	1.144	(0.009)***	357
Mining	0.471	(0.010)***	50
Motor Vehicles & Engines	0.489	(0.008)***	83
Other Business Services	1.505	(0.011)***	507
Other Durables	0.864	(0.009)***	232
Other Machinery	0.755	(0.009)***	164
Other Non-Durables	1.458	(0.018)***	165
Other Professional Srvcs	1.069	(0.008)***	436
Other Retail	1.625	(0.014)***	416
Paper Products	0.840	(0.009)***	194
Personal Services	0.925	(0.011)***	175
Public Administration	0.830	(0.010)***	222
Raw Durables	0.684	(0.009)***	150
Real Estate	1.252	(0.007)***	754
Repair Shops	1.119	(0.009)***	355
Telecommunications	1.419	(0.012)***	353
Textiles	0.502	(0.016)***	32
Transportation	1.475	(0.012)***	401
Utilities	0.715	(0.007)***	217
Wholesale Trade	1.442	(0.011)***	448

*** Coefficient significant at 1%

O*NET Values

O*NET provides information about hundreds of characteristics for over 800 occupations in the U.S. economy. Each O*NET variable used in this paper takes ordinal values ranging from 1 to 5, with 5 representing the highest score for a particular characteristic. I rescale these scores into percentile values based upon the year 2000 distribution of characteristics across occupations. Table 10 provides examples of occupations receiving percentiles near the minimum, median, and maximum points of the distribution.

Table 10: Examples of Occupations with High, Median, and Low O*NET Values.

Importance of:	*High Values (1)*	*Median Values (0.5)*	*Low Values (0)*
Making Decisions and Solving Problems	Chief Executives and Legislators	Designers	Drywall Installers, Ceiling Tile Installers, and Tapers
	Podiatrists	Procurement Clerks	Models, Demonstrators, and Product Promoters
Thinking Creatively	Architects	Insurance Underwriters	Insulation Workers
	Artists and Related Workers	Physical Therapists	Office Machine Operators, Except Computer
Customer and Personal Service	Interviewers	Bailiffs, Correctional Officers, and Jailers	Graders and Sorters, Agricultural Products
	Tellers	Telecommunications Line Installers and Repairers	Machine Feeders and Offbearers
Work With Work Group or Team	Engineering Technicians, Except Drafters	Meeting and Convention Planners	Mail Clerks and Mail Machine Operators, Except Postal Service
	Food Service Managers	Market and Survey Researchers	Private Detectives and Investigators

After assigning O*NET percentile values to occupation categories (as described in the text), I create weighted average values for each industry and year. Table 11 displays the industry-specific O*NET values for the year 2000.

Table 11: O*NET Values by Industry, 2000.

Industry	*Making Decisions and Solving Problems*	*Thinking Creatively*	*Customer and Personal Service*	*Work With Work Group or Team*
Advertising	0.533	0.687	0.491	0.397
Agriculture	0.601	0.495	0.316	0.203
Aircraft & Parts	0.645	0.633	0.316	0.442
Apparel Retail	0.436	0.393	0.626	0.448
Bars & Restaurants	0.308	0.345	0.617	0.498
Colleges & Universities	0.570	0.712	0.394	0.430
Computer & Software Design	0.765	0.782	0.288	0.379
Computer Manufacturing	0.676	0.658	0.347	0.464
Construction	0.415	0.541	0.315	0.416
Drugs & Chemicals	0.563	0.561	0.362	0.490
Education, Non-College	0.556	0.728	0.451	0.440
Electrical Machinery	0.596	0.568	0.306	0.472
Engineering & Architecture	0.684	0.748	0.348	0.512
Entertainment & Rec Services	0.510	0.539	0.495	0.475
Fabricated Metals	0.515	0.504	0.268	0.423
Finance	0.590	0.577	0.575	0.572
Food & Beverage Manuf	0.468	0.466	0.304	0.420
Food, Drug, & Alcohol Retail	0.399	0.399	0.530	0.592
General Retail	0.408	0.367	0.582	0.477
Health Services	0.614	0.460	0.695	0.631
Hotels	0.311	0.332	0.453	0.379
Insurance	0.546	0.398	0.629	0.417
Legal Services	0.597	0.418	0.464	0.270
Mining	0.498	0.537	0.314	0.516
Motor Vehicles & Engines	0.542	0.512	0.263	0.440
Other Business Services	0.492	0.505	0.462	0.452
Other Durables	0.554	0.534	0.322	0.456
Other Machinery	0.524	0.502	0.273	0.406
Other Non-Durables	0.528	0.503	0.297	0.453
Other Professional Services	0.554	0.599	0.477	0.485
Other Retail	0.450	0.445	0.569	0.460
Paper Products	0.505	0.560	0.389	0.434
Personal Services	0.295	0.500	0.579	0.223
Public Administration	0.572	0.528	0.478	0.529
Raw Durables	0.483	0.480	0.246	0.418
Real Estate	0.423	0.435	0.616	0.386
Repair Shops	0.426	0.552	0.566	0.472
Telecommunications	0.574	0.546	0.476	0.422
Textiles	0.507	0.466	0.241	0.409
Transportation	0.425	0.428	0.421	0.468
Utilities	0.513	0.529	0.378	0.478
Wholesale Trade	0.454	0.466	0.487	0.444