7-1998

Race, Redlining, and Automobile Insurance Prices

Scott E. Harrington

Gregory R. Niehaus

University of South Carolina - Columbia, gregn@moore.sc.edu

Publication Info

Published in *Journal of Business*, Volume 71, Issue 3, 1998, pages 439-469.

Scott E. Harrington
Greg Niehaus
University of South Carolina

Race, Redlining, and Automobile Insurance Prices*

Following Becker's (1993) suggestion that tests for discrimination should attempt to infer whether profits differ for products sold to minorities and nonminorities, this article tests the hypothesis that racial discrimination affects market prices of auto insurance in Missouri. Compared with tests for discrimination in lending markets, our results are less susceptible to bias from omitted variables. Controlling for available demographic and coverage-related factors, we do not find that loss ratios at the zip-code level are negatively related to percent minority population. This finding is inconsistent with the hypothesis that racial discrimination increases premiums relative to expected claim costs for minorities.

I. Introduction

In recent years, substantial attention and controversy has focused on tests for racial discrimination in markets for financial services, especially in residential mortgage lending (e.g., Berkovec et al. 1994; Holmes and Horvitz 1994; Ferguson and Peters 1995; Munnel et al. 1996; Harrison 1998).[1] Insurance companies also are often alleged to engage in racial discrimination in the form of "redlining" that raises prices and restricts availability of coverage, and these allegations have received considerable public-policy attention.[2] Despite this controversy, little is known about the extent to which premiums are higher

* We thank the National Association of Independent Insurers for purchasing the data, an anonymous reviewer, Doug Diamond, Anil Shivdasani, and workshop participants at the University of Michigan, the University of South Carolina, and the 1996 meetings of the American Risk and Insurance Association and the Risk Theory Society for helpful comments.

1. Also see the compendiums of essays edited by Yezer, Lindsey, and Isaac (1995) and by Stegman and Goering (1996). Much of the early empirical literature on discrimination focuses on labor markets (e.g., Ashenfelter and Rees 1973).

2. See, e.g., Squires, DeWolfe, and DeWolfe (1979) and Squires, Velez, and Taeuber (1991). Klein (1995) provides a brief survey of the literature. Congress held extensive hearings regarding alleged insurance redlining in 1993 (*Insurance Redlining Practices* 1993), and a number of state insurance departments have studied the issue and feel pressure to adopt regulations (e.g., Harrington and Niehaus 1992).

(*Journal of Business*, 1998, vol. 71, no. 3)

0021-9398/98/7103-0005$02.50

in areas with large minority populations relative to the costs of providing coverage in these areas.[3]

Racial discrimination that had a substantive effect on the supply of insurance would imply unexploited profitable investment opportunities and thus would require pervasive prejudice and significant entry barriers for potential nonprejudiced entrants. The existence of large numbers of insurers that are licensed to write auto insurance in most states with relatively low market concentration and entry costs (e.g., Klein 1989; Cummins and Tennyson 1992) militates against substantive economic effects from racial discrimination. The alternative hypothesis is that racial prejudice is pervasive enough in conjunction with entry costs or other market imperfections to allow discriminatory practices to harm minorities significantly.

Following Becker's (1993) suggestion that tests for discrimination should attempt to infer whether profits differ for products or services sold to minorities and nonminorities, this article tests the hypothesis that widespread racial discrimination affects the pricing of auto insurance at the market level. We first present a framework for analyzing discrimination in pricing or risk selection. The key testable implication is that if insurers discriminate, either by charging higher markups in areas with high minority populations or, more opaquely, by applying more restrictive underwriting standards to minorities, then the ratio of expected claim costs to premiums (the expected-loss ratio) will be lower, ceteris paribus, in areas with a higher percentage of minorities, consistent with higher expected profit margins in these areas. We test the prediction that discrimination implies lower loss ratios (ratios of average claim costs to average premiums) for minorities using insurance market and demographic data by zip code from Missouri, one of the few states that has collected premiums and claim cost information by zip code.[4] Missouri's population exhibits diversity in terms of race and other demographic characteristics. Because its automobile-insurance market essentially is free of prior approval rate and underwriting regulation, the data can be viewed as the result of market conduct rather than regulation.

Tests of discrimination in mortgage lending using data on default rates and rejection rates have been controversial, in large part because of the possibility of omitted variables related to differences in credit

3. Klein (1995) finds that homeowners insurance average premiums for zip codes in 33 metropolitan statistical areas in 20 states are positively related to percent minority population after controlling for a variety of demographic variables. His models include a variable measuring average claim costs by rating territory, but this variable does not control for possible differences in claim costs across zip codes within a territory that could be correlated with percent minority population.

4. The loss ratio is a commonly used measure of (inverse) price or profitability in insurance markets.

risk for minorities and nonminorities (e.g., Berkovec et al. 1994; Ferguson and Peters 1995; Shaffer 1996; Stegman and Goering 1996). The possibility of omitted-variable bias would be reduced significantly if mortgage-lending studies could examine a loan profitability measure that compared expected default losses with lending rates because any omitted variable affecting credit risk would influence expected default losses and lending rates in the same direction if lending rates vary with credit risk. Since measures of insurance profitability (loss ratios) are available and premiums vary with expected claim costs, our predictions and tests for discrimination should be affected substantially less, if at all, by omitted variables related to possible differences in claim-cost distributions between minorities and nonminorities.

Controlling for a variety of demographic and insurance coverage variables that might affect loss ratios at the zip-code level, our key finding is that loss ratios are not significantly lower in zip codes with larger minority populations. Thus, our loss-ratio analysis is inconsistent with the argument that widespread racial discrimination increases the price of automobile insurance in Missouri. An immediate implication is that comparatively high auto-insurance premiums in urban areas with larger minority populations are attributable to high claim costs in these areas—not to discrimination. Consistent with this implication, univariate comparisons indicate a strong, positive relation between average claim costs and percent minority population by zip code in Missouri. We also provide evidence that average claim costs remain positively related to percent minority population after controlling for a variety of demographic and insurance-coverage variables that could influence costs, which suggests that percent minority population is correlated with omitted variables that increase costs.

We also provide evidence of the relation between percent minority population and the market share of so-called nonstandard insurers, which specialize in insuring applicants who do not qualify for coverage with insurers that specialize in providing coverage at lower premium rates to "standard" or "preferred" risks. Thus, nonstandard insurers are more likely to insure drivers with high expected claim costs and drivers with a greater likelihood of nonpayment of premiums or nonrenewal, which increases the premium needed to recover up-front underwriting costs over the expected duration of the contractual relationship. While our theoretical framework implies that discrimination via more restrictive underwriting will cause the market share of nonstandard insurers to be higher in areas with a higher percentage of minorities after controlling for all relevant cost and demand factors, empirical analysis of nonstandard market share, like analyses of mortgage rejection rates, is highly susceptible to omitted-variable bias. Indeed, because our average claim-cost results imply that percent minority population is correlated with omitted variables that increase claim costs, these results predict a positive relation between the market share of nonstandard (higher-cost/higher-premium) insurers without discrimination after controlling for our demographic and coverage variables, which is exactly what we find. Moreover, our loss-ratio findings, which are inconsistent with discrimination, imply that a positive relation between nonstandard market share and percent minority population would reflect omitted variables. Thus, although our analysis of nonstandard market share by itself cannot distinguish whether a positive relationship between nonstandard share and percent minority population reflects discrimination or omitted variables, our loss-ratio and average claim-cost results strongly support the omitted-variable explanation.

A limitation of our analysis is that the insurance market data are aggregated at the zip-code level and are not available separately for minority and nonminority drivers. We also do not have data at the zip-code level on nonloss costs associated with the production of insurance, including costs that could be related to differences in quality. As a result, we rely on reduced form equations that include percent minority population and other demographic variables for the total population rather than the insured population. While we can test for a relationship between loss ratios and percent minority population, we cannot directly test or control for possible racial discrimination through lower quality that might cause profits to vary by race even if loss ratios do not. Our finding that loss ratios are not negatively related to percent minority population, which is inconsistent with racial discrimination in pricing holding quality fixed, is nonetheless important in view of the policy debate and dearth of empirical analysis in this area.

The next section develops the main hypotheses and the methodology for our empirical tests. The data and variables analyzed are described in Section III. Results for the loss-ratio models are presented in Section IV. Our analysis of the relationship between percent minority population, average claim costs, and nonstandard market share is presented in Section V. Our conclusions are summarized in the last section.

II. Hypotheses and Empirical Framework

A. *Risk Selection, Loss Ratios, and Nonstandard Market Share*

Auto insurers file with state regulators premium rates for a large number of "driver classes" and "territories." About half the states (but not Missouri) subject these rates to prior regulatory approval. In Missouri and most other states, rating plans (driver classes, territories, and associated rates) can vary across insurers. Since racial discrimination is illegal, rating plans cannot depend overtly on race, and a maintained assumption throughout our analysis is that race is not an underlying determinant of expected costs.

If persistent discrimination at the market level is feasible, there are two main ways that racial discrimination could influence prices charged insurance buyers.[5] First, insurers could file rates that exhibit a greater markup over expected costs for rating territories with a higher proportion of minorities, thus causing loss ratios to be negatively related to percent minority population. Second, insurers could discriminate by applying more stringent underwriting standards to insure minorities than nonminorities. Such discrimination through risk selection would likely be more opaque than increasing markups in areas with large minority populations.

Insurers can decline coverage to applicants in Missouri and most other states. They use information other than driver characteristics included in their filed rating plan to establish criteria for accepting or rejecting applicants. The acceptance criteria (''underwriting standards''), often are proprietary and can vary across insurers, including across affiliated insurers under common ownership. Stricter underwriting standards are associated with lower rates. If insurers apply stricter underwriting standards to minorities, then minorities would be rejected more often than is justified by unbiased forecasts of expected costs. As a consequence, minorities would be pushed to insurers with higher rates (and less stringent standards) and therefore have to pay a higher markup over unbiased expected costs, thus compensating these insurers for their aversion to serving minorities. This form of discrimination would cause loss ratios in both the overall market and the nonstandard market to be lower in areas with large minority populations, and it would increase nonstandard market share.

B. Model of Discrimination through Risk Selection

We present a simple model of racial discrimination through risk selection that yields predictions about the relationship between race, loss ratios, and market shares of insurers that specialize in terms of underwriting criteria. Although insurer specialization and nonstandard insurers are commonly discussed in the professional literature, there has been little formal analysis of insurer specialization with respect to underwriting criteria.[6] The model also helps explain why differences in

5. Evidence on price variation in relation to expected claim costs in insurance markets is not readily reconciled with models of rationing (e.g., Stiglitz and Weiss 1981; Ferguson and Peters 1997). Given our data and focus, we also abstract from issues of asymmetric information and associated consumer sorting à la Rothschild and Stiglitz (1976). D'Arcy and Doherty (1990) and Dionne and Doherty (1994) develop models in which ''low-ball'' or ''high-ball'' pricing by insurers in response to asymmetric information could cause variation in loss ratios across insurers and over time in conjunction with differences in insurer growth rates. As noted below, we control for exposure growth in some of our tests.

6. Smallwood (1975) considers a model in which direct-writing insurers apply more selective underwriting standards than insurers that use independent agents. Our model is analogous to models of the effect of discrimination in mortgage lending on default rates

the distributions of expected claim costs between racial groups are much less likely to bias the analysis of insurance loss ratios than the analysis of nonstandard market shares.

Assume for simplicity that each insurer uses the same factors in its filed risk-classification system with N distinct risk (driver type and geographic territory) classes, $j = 1, 2, \ldots, N$. The premium charged by insurer k to all drivers in risk class j equals P_j^k. An insurer's premiums will differ across risk classes in relation to expected claim costs (e.g., young males will be charged higher premiums). The focus of this analysis, however, is on how premiums for a given risk class will differ across insurers in relation to their underwriting criteria (e.g., State Farm or Allstate may charge lower premiums to young males than a nonstandard insurer). Since the conceptual analysis focuses on one risk class, further reference to class j is omitted with the understanding that subsequently defined variables can vary across risk classes.

Each consumer has a set of characteristics that place him or her in a particular risk class, as well as a characteristic that is observable to insurers, denoted Z, that influences expected claim costs, but which is not part of the risk-classification system. Without loss of generality, we define Z in terms of dollars of additional expected claim costs above the base expected claim cost for the risk class, which is denoted A. Thus, a consumer is assumed to have expected claim cost conditional on information that can be obtained at sufficiently low cost equal to $C = A + Z$, where A is constant for all consumers in the risk class and Z varies across individuals within the risk class.

Underwriting for insurer k is implemented by setting a maximum value for Z such that it accepts black applicants if their Z is less than M_b^k and accepts white applicants if their Z is less than M_w^k. If M_b^k is less than M_w^k, then insurer k requires, at the margin, a lower expected claim cost to insure blacks than whites (stricter standards are applied to blacks). The average expected claim cost for blacks and whites who are offered coverage by firm k is

$$E_b^k \equiv E(C|Z < M_b^k) = A + E(Z|Z < M_b^k)$$

and

$$E_w^k \equiv E(C|Z < M_w^k) = A + E(Z|Z < M_w^k),$$

where the expectation is with respect to the distributions of Z for black and white applicants for firm k.

If the distribution of Z for applicants that apply to firm k is the same for whites and blacks and if M_b^k is less than M_w^k (stricter standards are applied to blacks), then the average expected claim cost for blacks of-

(e.g., Berkovec et al. 1994), except that these models invariably assume a uniform lending rate regardless of applicant credit risk.

fered coverage is less than that for whites. In addition, given that the premium charged by insurer k, P^k, is the same for all consumers (black and white) in a risk class, the average expected loss ratio for blacks offered coverage in a risk class is less than the average expected loss ratio for whites: $E(LR_b^k) \equiv E_b^k/P^k < E(LR_w^k) \equiv E_w^k/P^k$.

If the distribution of Z for applicants that apply for coverage with company k is not the same for blacks and whites, it is possible that the average expected claim cost and loss ratio for blacks could exceed that for whites even if stricter standards are applied to blacks. This could occur, for example, if there are relatively fewer blacks with low values of Z insured by company k. That is, although a lower value of Z is required for blacks, the average expected cost is nonetheless higher for blacks because there are relatively few low Z blacks compared with whites.[7] A related result has been emphasized in analyses of mortgage lending that consider whether racial discrimination implies lower loan default rates for minorities than for nonminorities (e.g., Berkovec et al. 1994; Ferguson and Peters 1995).[8] However, this possibility has less relevance to the analysis of loss ratios in insurance markets where insurers vary premiums according to expected claim costs and underwriting standards.

To elaborate, if insurance companies specialize according to the maximum values of Z that they will insure and set premium rates accordingly, then companies with stricter underwriting standards (lower maximums for Z) will attract inframarginal applicants from companies with looser underwriting standards. This process reduces the variation in Z for a given insurer. With costless search and free entry, in equilibrium, consumers will each buy coverage from an insurer that has a maximum value of Z equal to the consumer's value of Z, and there will be no inframarginal consumers. The expected claim costs for black and white policyholders with company k will be related as follows:

$$E_b^k = E(C|Z = M_b^k) < E_w^k = E(C|Z = M_w^k).$$

Since the price charged by insurer k is the same for blacks and whites, the expected loss ratio for blacks insured, E_b^k/P^k, is unambiguously less than that for whites insured, E_w^k/P^k, if stricter standards are applied to blacks ($M_b^k < M_w^k$). Given similar behavior across risk classes, aggregate expected loss ratios are lower for blacks than for whites insured

7. This outcome leads to a somewhat perverse notion of discrimination. Specifically, if differences in the distribution of Z caused blacks to have higher average expected claim costs and loss ratios than whites despite $M_b^k < M_w^k$, average expected profits will be lower on policies sold to whites than blacks, in contrast to the standard for economic discrimination enunciated by Becker (1993).

8. Peterson (1981) provides an early treatment of this issue in the context of possible sex discrimination in mortgage lending. Also see the collection of essays on the Berkovec et al. study (Stegman and Goering 1996) and many of the essays in Yezer et al. (1995).

by each company, and market loss ratios are lower for blacks than whites.

Costly search and incomplete sorting of consumers according to the maximum values of Z used by different insurers would lead to some heterogeneity in expected claim costs among insureds of the same race in a given class for a given insurer. Nevertheless, the analysis of insurer specialization with respect to underwriting criteria implies that the greater the degree of specialization, the less likely that any difference in the distribution of Z between whites and blacks could lead to higher loss ratios for blacks despite a higher underwriting standard for blacks, thus reducing the likelihood that loss ratio comparisons will be biased against finding evidence of discrimination (lower loss ratios).[9]

Nonstandard insurers can be viewed within this framework as having looser underwriting standards (higher cutoffs for Z). If the distribution of drivers across risk classes and the distribution of Z within a risk class are the same for blacks and whites, then tighter underwriting standards for blacks (discrimination) imply a higher nonstandard market share for blacks. But nonstandard market share could be higher for blacks without discrimination if relatively more black drivers than white drivers were in higher risk classes or if relatively more blacks than whites had high values of Z. Thus, inability to control for all the factors affecting risk classes or Z could give rise to a positive relation between nonstandard share and race. We return to this issue below.

C. Testing for Discrimination Using Loss Ratios

The theoretical framework predicts that racial discrimination will cause loss ratios to be lower for blacks than for whites, holding non-claim-costs constant. To motivate our tests of this hypothesis with available data, we now describe a simple model of aggregate voluntary insurance market pricing at the zip-code level.[10] Underwriting automobile insurance involves three types of costs: claim costs, transaction costs (administrative, marketing, and claim-processing costs), and the cost of investing capital to bond insurer promises to pay claims. Let EXPCC_i equal the average (per exposure) of expected claim costs for zip-code i, and let λ_i equal the proportional loading factor that reflects the average non-claim costs for zip-code i and incorporates a present value factor. Then the average premium per exposure in zip-code i ($\mathrm{AVGPREM}_i$) can be written:

$$\mathrm{AVGPREM}_i = \mu_i \lambda_i \mathrm{EXPCC}_i,$$

9. Also, the greater the degree of insurer specialization, the lower the likelihood that low-expected-loss consumers in a given risk class will select out and not purchase insurance.

10. The same framework applies to loss ratios for the nonstandard market at the zip-code level.

where μ_i is the average markup factor relative to expected average costs.[11] Racial discrimination in risk selection or by using a higher markup in areas with larger minority populations implies that μ_i will be higher in areas with larger minority populations, thus producing higher premiums for a given value of EXPCC_i.

Average premiums also will clearly be higher in areas with larger minority populations without discrimination if expected claim costs are higher. In a similar manner, nonstandard market share will be higher in these areas without discrimination given that the factors causing higher expected claim costs will also increase nonstandard market share. Thus, inability to control for all factors that increase claim costs and that are correlated with percent minority population would cause both average claim costs and nonstandard market share to be positively related to percent minority population without discrimination.

Given the preceding expression for AVGPREM_i, the expected loss ratio equals

$$\text{EXPCC}_i/\text{AVGPREM}_i = (\mu_i\lambda_i)^{-1}.$$

While the expected loss ratio is inversely related to both μ_i and λ_i, there is less reason to expect that it will depend on expected claim costs and thus on unobservable factors that could cause expected claim costs to differ between races. As a result, the analysis of loss ratios will be much less susceptible to bias from the omission of variables that are correlated with claim costs and race.[12]

Expected claim costs and thus expected loss ratios are not observable. Let AVGCC_i equal realized average claim costs and assume that $\log(\text{AVGCC}_i)$ equals $\log(\text{EXPCC}_i)$ plus a mean-zero random disturbance, ϵ_i. Taking the log of AVGPREM_i, substituting for $\log(\text{EXPCC}_i)$, and rearranging terms gives

$$\log(\text{LOSSRAT}_i) = -\log(\mu_i) - \log(\lambda_i) + \epsilon_i,$$

where $\text{LOSSRAT}_i = \text{AVGCC}_i/\text{AVGPREM}_i$.[13] As discussed above, racial discrimination in risk selection or by using a higher markup in areas with larger minority populations implies that μ_i will be higher

11. The distinct parameters for nonclaim costs (λ_i) and the average markup (μ_i) would not be identifiable using data on average premiums and average claim costs without a model of how these parameters vary with other factors. We introduce such a model below.

12. If there are fixed costs in selling policies and the effects of discounting and all other nonclaim costs are strictly proportional to expected claim costs, then the expected loss ratio will be negatively related to expected claim costs. However, the magnitude of the resulting bias from omitted variables related to expected claim costs should nonetheless be small given that fixed costs are likely to represent a small proportion of premiums. (See appendix A for further discussion.) Also see n. 26 below.

13. An advantage of using the log of the loss ratio in our analysis and logs in our average cost equations (discussed below) is that it reduces positive skewness in the disturbances. The use of log models for average claim costs is common (e.g., Butler 1994).

and thus that expected and realized loss ratios will be lower in these areas.

Because we do not have data on the racial composition of the insured population by zip code, we cannot estimate a structural model of loss ratios, perhaps controlling for self-selection bias that could occur if buyers that face the lowest expected loss ratios (highest prices) are less likely to buy coverage. Instead, we estimate the following reduced form regression equation:

$$\log(\text{LOSSRAT}_i) = \gamma^{\text{LOSSRAT}}\text{PCTBLACK}_i + \beta^{\text{LOSSRAT}}X_i + \epsilon_i^{\text{LOSSRAT}},$$

where PCTBLACK_i is the percentage of the total population in zip-code i that is black, X_i is a vector of demographic and other factors that could affect λ_i, the loading factor that reflects transactions/capital costs and discounting to present value. The estimate of γ^{LOSSRAT} measures the sensitivity of $[-\log(\mu_i) - \log(\lambda_i)]$ to PCTBLACK. If X_i does not omit any variable that affects λ_i and which is correlated with PCTBLACK, then a nonnegative estimate of γ^{LOSSRAT} is inconsistent with racial discrimination in pricing. A negative estimate would be consistent with discrimination

The theory of competitive insurance prices (e.g., Myers and Cohn 1986) suggests that λ_i will depend on (1) the timing of claim payments, (2) the risk-adjusted discount rate, (3) the amount of capital held to bond insurer promises to pay claims (due to tax and agency costs of capital), and (4) the expense "loading" for insurer underwriting, distribution, claim-settlement, and policy-service costs. With the possible exception of the risk-adjusted discount rate, these factors could vary across zip codes due to variation in demand for quality, in market penetration by direct-writing (exclusive agent) insurers with lower distribution costs, in renewal rates on existing contracts, and in the prevalence of multicar policies and opportunities for cross-selling different types of coverage to a given policyholder. Data are not available by zip code to measure any of these factors directly, but variation in these factors should be related to demographic and coverage related factors that we can measure, at least approximately. We introduce the specific control variables in the next section and describe their possible relation to λ_i in the appendix.

If PCTBLACK is positively correlated with an omitted variable that is negatively related to λ_i and positively related to the loss ratio, then the coefficient estimate for PCTBLACK in the loss-ratio equation would be biased upward and against finding evidence of discrimination. To argue that profits are higher for blacks than for whites without a significantly negative coefficient for PCTBLACK would thus require evidence that non-claim costs (λ_i) are materially lower for blacks due, for example, to (1) lower demand for quality by blacks that is not captured by our control variables or (2) discriminatory reductions in certain types of product quality.[14] Finally, any correlation between PCTBLACK and omitted characteristics related to income or wealth that increase λ_i, by reducing renewal rates, opportunities for cross-selling, and so on, would bias the coefficient for PCTBLACK downward and toward finding evidence of discrimination.

III. Data and Variables

A. *The Missouri Data*

We analyze Missouri auto-insurance data for the period 1988–92 for collision, comprehensive (coverage for theft and physical damage other than collision, such as weather-related claims), and liability coverage (the data combine property-damage and bodily-injury liability). These data, which were obtained from the Missouri Department of Insurance, reflect reported underwriting results of all insurers with more than 500 exposures in Missouri. Since the insurance market data are analyzed in combination with 1990 census data, we aggregate these data over time and estimate cross-sectional models.

For each line of business, the data include the number of exposure units (car years), the number of claims, losses paid (exlusive of claim-settlement expenses), and premiums written by zip code for five policy (insurer) types (e.g., standard, nonstandard) and five exposure classes based on car value ranges or liability limits (see the appendix). We exclude policies issued by the Missouri Joint Underwriting Association (JUA) because the proportion of consumers that are insured in the JUA is small in all zip codes and because the state insurance commissioner regulates JUA rates.[15] We aggregate the voluntary market data (non-

14. Some evidence exists regarding product quality differences between minorities and nonminorities in homeowners' insurance, but it generally relates to quality dimensions that are unlikely to bias our tests against finding discrimination, even if it is robust and generalizes to auto insurance. For example, Squires et al. (1991) provide evidence that insurers have fewer marketing outlets (agencies) located in areas with high minority populations. If associated with lower quality, this difference could reflect higher costs of providing marketing services in areas with minority populations, which would bias our results toward finding evidence of discrimination. In a similar manner, Chan (1997) provides evidence that complaints against homeowners' insurers are greater in areas with high minority populations, which might suggest lower-quality service to minorities. However, if a greater likelihood of claims fraud and associated efficient insurer responses in the form of greater expenditures on investigating claims were correlated with percent minority population due to a relation with unmeasured factors, then expected loss ratios (exclusive of claim settlement costs, as is the case with our data) would be negatively related to percent minority population absent discrimination, thus biasing loss ratio comparisons toward finding evidence of discrimination.

15. The JUA share of insured exposures for liability coverage in our sample of urban zip codes ranged from 0%–3.7% with a mean of .2%. The percentages are smaller for collision and comprehensive coverage.

JUA policies) across policy types and across exposure classes.[16] We limit the analysis to zip codes with at least 1,000 voluntary market liability exposure units. Random variation in claim costs will likely be quite large and highly skewed in zip codes with few exposures. For each zip code and line of business, we calculate the following variables:[17]

LOSSRAT = losses paid per premiums written,
AVGPREM = premiums written per number of exposure units,
AVGCC = losses paid per number of exposure units,
FREQ = number of claims per number of exposure units,
SEV = losses paid per number of claims, and
NSSHARE = number of nonstandard exposure units per number of exposure units.

As discussed above, our key test for discrimination is whether LOSSRAT is negatively related to PCTBLACK.

We report results only for the 270 zip codes (in 17 counties) that fall within Missouri's six metropolitan statistical areas (MSAs): Columbia, Joplin, Kansas City, St. Joseph, St. Louis, and Springfield. Most of the controversy about race and insurance markets centers on urban areas, and the MSA criterion eliminates zip codes with generally small numbers of exposures and relatively few minorities. The average number of exposures and average percentage of the population that is black are 9,097 and 1.5%, respectively, in the 613 nonurban zip codes versus 36,204 and 9.3% in urban zip codes. We also estimated models using all zip codes with 1,000 or more liability exposures during the 5-year period with little change in the results.

16. The regression analysis includes control variables for the percentage of the market in different exposure classes. As discussed below, we also conduct separate analysis for the nonstandard market and for the voluntary market excluding nonstandard policies, as well as estimating models of nonstandard market share.

17. The loss data are for claims paid during a given calendar year. Some of these claims will be for accidents that occurred in prior years. The premium data are for policies sold in the calendar year, which should reflect the present value of expected claim costs for accidents during the coverage period. Depending on the growth rates in exposures and in average expected claim costs per exposure over time and on the interest rate used in discounting, this mismatching of claims and premiums (which should only be material for liability coverage) could cause the LOSSRAT variable to differ from the expected loss ratio for policies sold in a given year (see Harrington 1992). If growth rates in insured exposures or expected costs per exposure differ across zip codes, this mismatch could lead to a bias in our results if growth rates were correlated with any of the exogenous variables. The use of 5-year averages should mitigate the mismatch problem. In addition, we included growth in exposures as an explanatory variable and obtained results similar to those reported in the article.

B. Control Variables

The five exposure classes allow us to control for differences in average insured-vehicle value and liability-coverage limits that will affect expected claim costs and possibly loss ratios. These variables also will likely be positively related to differences in wealth and income that will not be captured by our income measures. For collision and comprehensive coverage, we include log of average car value (log AVGCARV). For liability coverage, we include the percentage of exposure units in coverage-limit class 2 (MEDLIMITS) and the percentage of exposure units in exposure classes 3, 4, and 5 (HIGHLIMITS). The omitted category is exposure class 1, which includes policies with the lowest coverage limits.[18]

Information on demographic characteristics of the population by zip code was obtained from 1990 census data.[19] The demographic variable of primary interest is the percentage of the population that is black (PCTBLACK). We control for a variety of other demographic variables in the models that reflect: (1) population density, (2) age distribution and marital status, and (3) employment, income, and education. The demographic and coverage variables are summarized in table 1, which also shows how each variable (other than PCTBLACK) could be related to loss ratios (see the appendix for further discussion). The predicted effect of the variables on loss ratios is usually ambiguous.

We estimated models with and without county dummy variables to allow for fixed county effects. Including fixed county effects reduces the possibility of bias from any omitted factors that have a similar effect on loss ratios for zip codes within a county, but we have no strong prior concerning material county effects. Because including fixed county effects generally produced results with the same implications for PCTBLACK and led to instability in the coefficients for some control variables, we primarily report results for equations without fixed effects. Given that the employment, income, and education variables generally are highly correlated and correlated with average car value and HIGHLIMITS, we focus on equations that only include PCTUNEMP from this category. We also report estimated coeffi-

18. We treat these coverage variables as exogenous. It might be argued that they will depend on prices and thus loss ratios. These variables also could reflect self-selection associated with the application and approval process, which might create some bias of unknown magnitude and direction (see Rachlis and Yezer 1993). Our judgment is that including these variables is likely to increase efficiency and power to reject the null hypothesis of no relation between PCTBLACK and the loss ratio. We obtained results with similar implications for PCTBLACK when these variables were omitted.

19. The census files were prepared from U.S. Bureau of the Census data by CIESIN (Consortium on International Earth Science Network) with support from the National Aeronautics and Space Administration. Zip codes change over time and consequently there is not a perfect match between the census data and insurance data.

TABLE 1 **Definitions of Control Variables and Predicted Relationships with Loss Ratios**

Underlying Factor and Variable Name	Variable Definition	Predicted Effect on Loss Ratios due to Relationship with Major Factors That Influence Transaction/Capital Costs (see Appendix)			
		Product Quality	Direct-Writer Share	Persistency-Repeat Business	Multicar Policies and Cross-Selling
Percent minority: PCTBLACK	Percentage of total population that is black	. . .	. . .	. . .	. . .
Population density: log TOTPOP	Natural logarithm of total population	?	?	?	?
Age distribution and marital status:					
PCT1824	Percentage of adult population (18 and older) between the ages of 18 and 24	?	−	−	−
PCT55UP	Percentage of adult population older than 54	?	+	+	+
MARRIED	Percentage of households (excluding single person households over the age of 65) with a married couple	?	+	+	+
Employment, income, and education:					
PCTUNEMP	Percentage of the working-age population that is unemployed	+	−	−	−
PCTPOOR	Percentage of population in households with income below the poverty level	+	−	−	−
LMEDHINC	Natural logarithm of median household income	−	+	+	+
PCTHIGHS	Percentage of adult population with at least a high school education	−	+	+	+
Coverage:					
log AVGCARV	Natural log of weighted-average car value using midpoints of car value ranges	−	+	+	+
MEDLIMITS and HIGHLIMITS	Percentage of exposures in liability limit classes 2 and 3, 4, 5, respectively	−	+	+	+

SOURCES.—Demographic variables are from the 1990 census, and the insurance-coverage variables are from the Missouri Department of Insurance.

cients for PCTBLACK for alternative specifications of the loss ratio model involving PCTPOOR, LMEDHINC, and PCTHIGHS.

C. Descriptive Statistics and Collinearity Diagnostics

Tables 2 and 3 present descriptive statistics for the insurance market and control variables for the 270 urban zip codes with complete insurance and demographic data and for two subgroups: (1) all urban zip codes in which blacks comprise less than 20% of the population (237 zip codes), with an aggregate percentage black population in these zip codes of 3.6%, and (2) all urban zip codes in which blacks comprise more than 20% of the population (33 zip codes), with an aggregate percentage black population of 61.3%. Table 2 indicates that blacks on average pay higher premiums, have higher average claim costs, and higher nonstandard market share in each line of insurance, but loss ratios are similar for the two groups in each line.[20] Table 3 indicates that zip codes with a relatively high proportion of blacks have greater population density (TOTPOP), poorer economic conditions and education (higher PCTUNEMP and PCTPOOR, lower MEDHINC and PCTHIGHS), and a much lower percentage of households with married couples (MARRIED).

Relatively high correlations exist between PCTBLACK, PCTUNEMP, and MARRIED (see the appendix).[21] Collinearity diagnostics described by Belsley (1991) suggest that the estimates on the average car value and liability coverage variables, MARRIED, and, to a much lesser extent, the age variables could be "degraded" by collinearity. However, our estimation results for these variables (i.e., frequently large *t*-values) often suggest that this collinearity is not likely to be harmful, especially for MARRIED. More important, the diagnostics provide no indication that the estimates for PCTBLACK are likely to be degraded by collinearity.[22] As is discussed below, we also estimated

20. Comparison of the nonstandard market data versus the data for the total voluntary market in the 270 urban zip codes (not reported in the tables) indicates that average premiums and average claim costs are approximately 30%–90% higher in the nonstandard market with the largest increase for collision and the smallest increase for comprehensive. Higher average premiums and claim costs in the nonstandard market are consistent with this portion of the market reflecting risk categorization and insurer specialization. Although the nonstandard market has higher average claim costs and premiums, the aggregate loss ratios for the nonstandard market are smaller than those for the overall market, which suggests that nonstandard insurers have higher nonclaim costs.

21. Regressions of PCTBLACK on the remaining explanatory variables in each model produced R^2s of 69%–70% without fixed effects and 72%–74% with fixed effects.

22. Specifically, two condition indices and associated variance proportions generally suggested "near" linear dependencies involving the intercept (or county dummies), log AVGCARV, and MARRIED for collision and comprehensive coverage, and for the intercept (or county dummies), MARRIED, MEDLIMITS, and HIGHLIMITS for liability coverage. The sum of the variance proportions associated with these two condition indices for PCTBLACK were less than .37, .24, and .12 for collision, comprehensive, and liability, respectively (compared with Belsley's [1991] suggested cutoff of .5 to indicate potentially degraded estimates).

TABLE 2 **Descriptive Statistics for Insurance Market Variables**

	Collision			Comprehensive			Liability		
Variable and Statistic	270 Urban Zip Codes	237 Zip Codes with PCTBLACK <20%	33 Zip Codes with PCTBLACK >20%	270 Urban Zip Codes	237 Zip Codes with PCTBLACK <20%	33 Zip Codes with PCTBLACK >20%	270 Urban Zip Codes	237 Zip Codes with PCTBLACK <20%	33 Zip Codes with PCTBLACK >20%
LOSSRAT (%):									
Mean	61.48	61.57	60.84	59.12	58.98	60.15	66.16	66.43	64.28
Aggregate	58.59	58.30	60.32	51.86	51.02	56.42	66.33	66.58	64.89
SD	8.07	8.20	7.12	21.67	21.88	20.40	15.29	15.79	11.08
AVGPREM ($100):									
Mean	1.86	1.77	2.55	.99	.93	1.45	2.20	2.10	2.95
Aggregate	1.92	1.86	2.33	.98	.94	1.28	2.33	2.26	2.79
SD	.39	.29	.31	.25	.15	.31	.47	.39	.29
AVGCC ($100):									
Mean	1.14	1.08	1.55	.58	.55	.85	1.46	1.40	1.89
Aggregate	1.12	1.09	1.41	.51	.48	.72	1.54	1.51	1.81
SD	.26	.21	.25	.25	.21	.22	.43	.41	.36
FREQ (%):									
Mean	8.52	8.05	11.94	9.41	9.41	9.41	6.33	6.06	8.25
Aggregate	8.83	8.52	11.13	8.91	8.90	9.01	6.89	6.74	7.98
SD	1.91	1.39	1.63	1.99	2.08	1.17	1.40	1.24	.81
SEV ($100):									
Mean	13.50	13.58	12.96	6.15	5.76	8.93	23.03	23.06	22.85
Aggregate	12.74	12.76	12.66	5.71	5.39	8.03	22.41	22.37	22.69
SD	1.93	1.99	1.25	1.96	1.61	1.98	5.15	5.38	3.13
NONSTSH (%):									
Mean	2.06	1.56	5.59	1.95	1.48	5.31	4.27	3.54	9.57
Aggregate	1.83	1.46	4.57	1.74	1.39	4.33	3.89	3.32	8.00
SD	1.75	.75	2.64	1.66	.70	2.52	2.88	1.75	3.79

NOTE.—Descriptive statistics are for average claim costs divided by average premiums (LOSSRAT), average premiums (AVGPREM), average claims costs (AVGCC), number of claims divided by number of exposures (FREQ), average claim amount (SEV), and proportion of exposures sold by nonstandard insurers (NONSTSH) for collision, comprehensive, and liability automobile insurance coverage in Missouri during the1988–92 period in 270 Missouri zip codes within metropolitan statistical areas, where an exposure unit is an insured car for 1 year.

TABLE 3 Descriptive Statistics for Insurance Markets and Demographic Variables in 270 Missouri Zip Codes

Variable and Statistic	270 Urban Zip Codes	237 Zip Codes with PCTBLACK <20%	33 Zip Codes with PCTBLACK >20%
PCTBLACK (%):			
Mean	9.34	2.21	60.60
Aggregate	14.63	3.56	61.31
SD	21.40	3.67	25.77
TOTPOP (1,000s):			
Mean	12.41	11.43	19.47
Aggregate	N.A.	N.A.	N.A.
SD	12.84	12.69	11.83
PCT1824 (%):			
Mean	12.42	12.15	14.42
Aggregate	13.57	13.31	14.70
SD	4.38	4.47	3.04
PCT55UP (%):			
Mean	28.84	28.72	29.67
Aggregate	28.49	28.13	30.03
SD	7.70	7.81	6.96
MARRIED (%):			
Mean	66.91	71.04	37.24
Aggregate	60.89	65.27	41.68
SD	16.79	12.34	14.58
PCTUNEMP (%):			
Mean	5.92	5.05	12.15
Aggregate	6.12	4.83	11.57
SD	3.78	2.47	5.44
PCTPOOR (%):			
Mean	11.09	8.97	26.27
Aggregate	11.09	8.19	23.34
SD	9.33	6.17	13.46
MEDHINC ($1,000):			
Mean	33.83	34.11	31.82
Aggregate	N.A.	N.A.	N.A.
SD	3.73	3.72	3.21
PCTHIGHS (%):			
Mean	75.39	76.58	66.88
Aggregate	77.62	79.78	68.49
SD	11.16	10.35	13.15
AVGCARV ($1,000) collision:			
Mean	10.97	10.93	11.29
Aggregate	11.35	11.35	11.35
SD	1.05	1.07	.86
AVGCARV ($1,000) comprehensive:			
Mean	10.96	10.64	11.04
Aggregate	11.08	11.08	11.10
SD	1.06	1.08	.82

TABLE 3 Continued

Variable and Statistic	270 Urban Zip Codes	237 Zip Codes with PCTBLACK <20%	33 Zip Codes with PCTBLACK >20%
MEDLIMITS:			
Mean	28.57	28.87	26.41
Aggregate	28.82	28.85	28.55
SD	5.49	5.42	5.60
HIGHLIMITS:			
Mean	37.17	38.61	26.85
Aggregate	42.37	43.99	30.65
SD	14.35	14.03	12.36

SOURCES.—Demographic data are from the 1990 U.S. census. The insurance-coverage data are from the Missouri Department of Insurance.

NOTE.—PCTBLACK is the proportion of the population that is black, TOTPOP is the total population, PCT1824 is the percentage of the adult population between the ages of 18 and 24, PCT55UP is the percentage of the adult population that is 55 years or older, MARRIED is the percentage of households with a married couple, PCTUNEMP is the percentage of the population that is unemployed, PCTPOOR is the percentage of households with income below the poverty level, MEDHINC is the median household income, PCTHIGHS is the percentage of the adult population that has finished high school. AVGCARV is the average insured car value based on classifications reported in the appendix, MEDLIMITS is the percentage of liability policies with medium limits (exposure class 2 in the appendix), and HIGHLIMITS is the percentage of liability policies with high limits (exposure classes 3, 4, and 5 in the appendix).

the models excluding some of the regressors as well as including others. The conclusions with regard to the effects of the variable PCTBLACK are largely unaltered.

IV. Are Loss Ratios Related to Race?

Table 4 presents weighted least squares regression (WLS) results for the loss ratio equation for the overall voluntary market. The weight is the number of exposures in the zip code, given the prior expectation that the variance of the random error in average realized claim costs is inversely related to this weight.[23] The key result is that loss ratios are not significantly and negatively related to PCTBLACK in any line. The estimated coefficient is positive for both collision and comprehensive coverage. These results are therefore inconsistent with the hypothesis that, ceteris paribus, blacks are charged higher prices relative to claim costs than whites.

The positive coefficient on PCTBLACK is statistically significant for collision coverage. Reverse discrimination is an unlikely explanation of this result. Other possible explanations include random variation in

23. Similar results were obtained using ordinary least squares with White standard errors.

TABLE 4 Weighted Least Squares Regression Results for the Log of the Voluntary Market Loss Ratio (LOSSRAT)

	Collision	Comprehensive	Liability
INTERCEPT	−.71	1.98	−.91
	(−3.53)	(2.73)	(7.24)
PCTBLACK	.09	.11	−.01
	(1.96)	(.63)	(.21)
log TOTPOP	−.02	−.10	.03
	(3.25)	(4.43)	(2.47)
PCT1824	.23	−.23	.38
	(1.83)	(.50)	(1.91)
PCT55UP	−.05	−.47	−.06
	(.65)	(1.76)	(.55)
MARRIED	.32	−.19	.39
	(5.47)	(.90)	(4.21)
PCTUNEMP	.86	.11	.58
	(2.77)	(.10)	(1.21)
log AVGCARV	−.17	−.87	. . .
	(.23)	(3.26)	
MEDLIMITS	. . .	. . .	.46
			(2.80)
HIGHLIMITS	. . .	. . .	−.06
			(.74)
OLS adjusted R^2	.17	.11	.06

NOTE.—Observations are for 270 Missouri zip codes in metropolitan statistical areas using aggregate insurance market data from 1988 through 1992; weight is the number of exposures; coefficients on percent variables are multiplied by 100; absolute t-statistics are in parentheses below coefficient estimates. PCTBLACK is the proportion of the population that is black, log TOTPOP is the natural logarithm of total population, PCT1824 is the percentage of the adult population between the ages of 18 and 24, PCT55UP is the percentage of the adult population that is 55 years or older, MARRIED is the percentage of households with a married couple, PCTUNEMP is the percentage of the population that is unemployed, log AVGCARV is the log of average insured car value based on classifications reported in the appendix, MEDLIMITS is the percentage of liability policies with medium limits (exposure class 2 in the appendix), and HIGHLIMITS is the percentage of liability policies with high limits (exposure classes 3, 4, and 5 in the appendix). OLS = ordinary least squares.

claim costs and variation in the demand for quality that is not captured by the control variables.[24] If, on average, demand for quality is lower in zip codes with a higher proportion of minorities, then loss ratios will tend to be higher in these areas. This type of quality variation does not imply discrimination, however, as premiums vary with quality.

The explanatory power of the loss ratio equations is relatively low, which may reflect the ambiguous effects of the control variables on loss ratios in addition to sizable random variation in claim costs at the zip-code level. Several of the variables have statistically significant

24. Also note that the probability that random variation would lead to rejection of the null hypothesis of a zero coefficient in at least one of three lines of business is higher than if only one line were being examined.

coefficients, suggesting that transaction/capital costs vary with these variables.[25] The coefficient on MARRIED is positive and significant for both collision and liability coverage, which is consistent with lower average transactions costs in areas with more married couples due to larger renewal rates (persistency), a greater prevalence of multiple car policies, and/or because married couples are more likely to purchase other types of insurance (life and homeowners) with the same insurer. The positive relation between loss ratios and PCTUNEMP could reflect lower demand for quality in economically depressed regions (or unexpectedly high collision losses in these regions during this period).[26]

Table 5 provides evidence of the robustness of the results for PCTBLACK for five specifications of the loss ratio model that exclude one or more of the demographic variables and for two specifications that include other demographic variables measuring economic conditions (LMEDHINC and PCTPOOR) and education (PCTHIGHS). We also show results obtained when fixed county effects are included in each specification. For collision and comprehensive coverage, only one coefficient for PCTBLACK is negative, and this estimate has a t-ratio near zero. For liability coverage, several of the coefficients without fixed effects and all of the estimates with fixed effects are negative. The negative estimates are significantly different from zero at the .05 level for a one-tailed test for three specifications without fixed effects and for one specification with fixed effects. Each of these specifications excludes MARRIED. There are good conceptual reasons for including MARRIED, as it may be a proxy for transactions costs and persistency. The variable MARRIED is often significantly related to the loss ratio (and average claim costs, see below) for all of the coverages, suggesting that the preferred specifications include this variable. Thus, the overall results of our analysis of voluntary market loss ratios provide little or no evidence that racial discrimination gives rise to higher premiums relative to claim costs in minority areas.[27]

25. The signs are not always consistent across lines, which might reflect random variation in claim costs across lines.

26. It might be argued that loss ratios should be higher in densely populated urban areas with high expected claim costs based on the presumption that underwriting and distributions costs increase less than proportionately with expected claim costs. While the coefficient on log TOTPOP is positive and significant for liability coverage, the negative and significant coefficients on this variable for collision and comprehensive coverage are inconsistent with this hypothesis. The latter results might indicate higher underwriting and distribution costs or lower renewal rates in more densely populated areas. Recall also that all of our zip codes are in urban areas.

27. In order to allow for the possibility that loss ratios could be related to expected claim costs (see n. 12 above), we also estimated loss ratio models that included the log of average claim costs. Given the random error (measurement error) in average claim costs relative to expected claim costs, we employed instrumental variables estimation (compare Harrington 1987), using the log of the average premium by zip code as an instrument for the log of average claim costs. The coefficient on the log of average claim costs was not

TABLE 5 Alternative Specifications of Loss Ratio Model

Specification	Collision (1)	Collision (2)	Comprehensive (3)	Comprehensive (4)	Liability (5)	Liability (6)
With fixed county effects	no	yes	no	yes	no	yes
Model shown in table 4	.09	.07	.11	.42	−.01	−.08
	(1.96)	(1.51)	(.64)	(3.24)	(.21)	(1.03)
Without MARRIED, PCT-UNEMP, PCT1824, and PCT55UP	.06	.09	.18	.22	−.10	−.12
	(2.09)	(2.98)	(1.83)	(2.92)	(2.01)	(1.76)
Without PCTUNEMP, PCT1824, and PCT55UP	.16	.18	.12	.28	.01	−.08
	(4.68)	(5.01)	(1.01)	(3.08)	(.12)	(1.12)
Without MARRIED	−.03	.01	.16	.39	−.12	−.13
	(.07)	(.19)	(1.05)	(3.16)	(1.93)	(1.64)
Without PCTUNEMP	.18	.18	.12	.28	.03	−.07
	(5.05)	(5.02)	(.94)	(3.00)	(.60)	(1.05)
Without PCTUNEMP; with PCTPOOR and LOG MED-HINC	.15	.14	.19	.45	.01	−.08
	(3.57)	(3.15)	(1.24)	(3.96)	(.15)	(1.11)
With PCTPOOR, LOG MED-HINC, and PCTHIGHS	.12	.13	.03	.49	.03	−.03
	(2.62)	(2.59)	(.20)	(3.73)	(.39)	(.30)
With PCTBLACK only	.04	.09	.13	.23	−.08	−.03
	(1.53)	(2.89)	(1.32)	(3.03)	(1.89)	(.67)

NOTE.—Coefficients for PCTBLACK with absolute *t*-statistics are in parentheses below from weighted least squares estimation for 270 urban zip codes; weight is the number of exposures; coefficients are multiplied by 100. PCTBLACK is the proportion of the population that is black, PCT1824 is the percentage of the adult population between the ages of 18 and 24, PCT55UP is the percentage of the adult population that is 55 years or older, MARRIED is the percentage of households with a married couple, PCTUNEMP is the percentage of the population that is unemployed, MEDHINC is the median household income, PCTHIGHS is the percentage of the adult population that has finished high school, PCTPOOR is the percentage of households with income below the poverty level.

As is true for the overall voluntary market, discrimination would imply that nonstandard insurer loss ratios would be lower in zip codes with larger minority populations. A priori, estimation of separate loss ratio models for the nonstandard market is hindered by the comparatively small number of exposures and correspondingly large random variation in average claim costs. We nonetheless estimated the reduced form models reported in table 4 for nonstandard market policies using the 153 zip codes with at least 500 nonstandard liability exposures (the previous models required 1,000 total exposures) and found little evidence indicating that loss ratios were negatively related to PCTBLACK. The coefficients, with *t*-values in parentheses, on PCTBLACK for collision, comprehensive, and liability are .12 (.85), .68 (2.54), and −.06 (−.46), respectively.[28]

significant for any of the three lines, and the coefficients and *t*-values for PCTBLACK changed very little from those reported in table 4.

28. Similar results to those reported in table 4 also were obtained when we estimated the loss ratio models excluding nonstandard policies.

V. Are Average Claim Costs and Nonstandard Market Share Related to Race?

An implication of our loss ratio results is that auto-insurance-affordability problems in urban areas with comparatively large minority populations are not caused by discrimination. Instead, as is suggested by the higher average claim costs in urban zip codes with PCTBLACK less than 20% than in zip codes with PCTBLACK greater than 20% (see table 2), higher premiums reflect higher claim costs in these areas. Higher average claim costs in these areas also would be expected to be associated with higher nonstandard market share without racial discrimination. This section presents additional evidence on the relation between average claim costs, nonstandard market share, and race.

Table 6 presents WLS results of (1) simple regressions of the log of average claim costs on PCTBLACK for each line of coverage and (2) reduced-form models for average claim costs that include PCTBLACK and the control variables included in the loss ratio models. The table also shows predicted values of average claim costs when PCTBLACK equals 0% and 100% when all other explanatory variables are set to their sample means.[29] Table 6 indicates a very strong univariate relationship between average claim costs and PCTBLACK for each line of business. The coefficient on PCTBLACK declines in size in the multivariate equations, but it remains significantly positive for each line. This result suggests that our control variables omit one or more factors that are related to average claim costs and that are correlated with PCTBLACK.[30] An implication is that these omitted cost factors will also be likely to produce a positive coefficient in reduced form models for nonstandard market share without discrimination.

Table 7 presents the results of simple regressions of nonstandard

29. The predicted values equal $\exp(y)(\times 100)$, where y is the predicted value of the log of average claim costs (with average claim costs in \$100s). While $\exp(y)$ generally will not be an unbiased predictor, the bias should be small and should not affect the implications of these comparisons or similar comparisons for nonstandard market share (see below).

30. Estimation of separate reduced-form models of claim frequency and severity indicated a positive and significant relation between PCTBLACK and collision coverage claim frequency, liability coverage claim frequency, and comprehensive coverage claim severity. We also estimated the average claim cost equations in table 6 controlling for fixed county effects. The coefficient estimate on PCTBLACK continues to be positive and statistically significant for collision and comprehensive coverage, but it is not significant for liability coverage. We also estimated alternative specifications that excluded some demographic variables and included other variables, but PCTBLACK generally remained positively related to average claim costs after controlling for the other variables. Among other specifications, we included an interaction variable between PCT55UP and PCTBLACK to allow for the possibility that comparatively few elderly blacks drive than elderly whites, thus increasing average claim costs in zip codes with greater values of PCTBLACK for any given value of PCT55UP. High correlations between the interaction variable and its components led to large standard errors for the coefficients for each variable.

TABLE 6 Weighted Least Squares Regression Results for the Logarithm of Average Claim Costs in $100s (AVGCC)

	Collision		Comprehensive		Liability	
	(1)	(2)	(3)	(4)	(5)	(6)
INTERCEPT	.060	−2.30	−.81	.29	.39	.10
	(6.81)	(8.76)	(37.21)	(.39)	(31.36)	(.63)
PCTBLACK	.50	.22	.75	.30	.34	.26
	(11.59)	(3.68)	(6.98)	(1.75)	(5.86)	(3.06)
log TOTPOP	. . .	−.01	. . .	−.13	. . .	.05
		(1.46)		(5.50)		(3.74)
PCT1824	. . .	−.64	. . .	−1.19	. . .	−.76
		(3.91)		(2.51)		(2.97)
PCT55UP	. . .	−.40	. . .	−.73	. . .	−.52
		(4.21)		(2.63)		(3.46)
MARRIED	. . .	−.18	. . .	−.83	. . .	−.23
		(2.34)		(3.77)		(1.89)
PCTUNEMP	. . .	2.34	. . .	1.97	. . .	1.37
		(5.79)		(1.69)		(2.24)
log AVGCARV	. . .	1.08	. . .	.05	. . .	. . .
		(11.11)		(.19)		
MEDLIMITS	. . .	. . .	. . .	. . .	. . .	.85
						(4.05)
HIGHLIMITS	. . .	. . .	. . .	. . .	. . .	.53
						(5.16)
OLS adjusted R^2	.33	.56	.17	.28	.14	.37
Predicted AVGCC: PCTBLACK = 0 ($)	106	109	51	53	135	136
Predicted AVGCC: PCTBLACK = 100 ($)	175	136	107	71	190	177

NOTE.—Observations are for 270 Missouri urban zip codes using aggregate insurance market data from 1988 through 1992; weight is the number of exposures; coefficients with absolute *t*-statistics are in parentheses below; coefficients on percent variables are multiplied by 100. PCTBLACK is the percentage of the population that is black, log TOTPOP is the natural logarithm of total population, PCT1824 is the percentage of the adult population between the ages of 18 and 24, PCT55UP is the percentage of the adult population that is 55 years or older, MARRIED is the percentage of households with a married couple, PCTUNEMP is the percentage of the population that is unemployed, log AVGCARV is the log of average insured car value based on classifications reported in the appendix, MEDLIMITS is the percentage of liability policies with medium limits (exposure class 2 in the appendix), and HIGHLIMITS is the percentage of liability policies with high limits (exposure classes 3, 4, and 5 in the appendix). Predicted AVGCC is the predicted value of average claim costs for the specified value of PCTBLACK with other variables equal to the sample means. OLS = ordinary least squares.

market share (in log odds form) on PCTBLACK and reduced-form regressions that include the control variables using least squares with White's heteroscedasticity-consistent standard errors.[31] Because drivers will usually purchase each coverage from one insurer (and are required to purchase liability coverage), we report results for a model that includes MEDLIMITS and HIGHLIMITS (defined for the entire volun-

31. The disturbance variance in the nonstandard share model should decline with both the number of nonstandard exposures and the total number of exposures.

TABLE 7 **Ordinary Least Squares Regression Results for the Nonstandard Market Share (in Log Odds Form)**

	Collision		Comprehensive		Liability	
	(1)	(2)	(3)	(4)	(5)	(6)
INTERCEPT	−4.28	−2.89	−4.34	−2.99	−3.43	−1.74
	(145.40)	(8.32)	(148.72)	(8.40)	(114.97)	(4.86)
PCTBLACK	2.15	.94	2.15	.98	1.72	.30
	(17.00)	(5.78)	(17.15)	(6.13)	(13.42)	(1.82)
log TOTPOP	...	.02	...	.03	...	.02
		(1.21)		(1.40)		(1.18)
PCT1824	...	.34	...	.48	...	.24
		(.38)		(.53)		(.32)
PCT55UP	...	−1.65	...	−1.63	...	−1.62
		(4.14)		(4.07)		(4.39)
MARRIED	...	−1.20	...	−1.14	...	−1.18
		(5.81)		(5.56)		(5.17)
PCTUNEMP	...	2.30	...	2.19	...	3.37
		(2.20)		(2.13)		(3.72)
MEDLIMITS	...	.87	...	.78	...	.11
		(2.05)		(1.82)		(.27)
HIGHLIMITS	...	−1.27	...	−1.24	...	−1.65
		(5.89)		(5.73)		(8.32)
OLS adjusted R^2	.52	.72	.52	.72	.40	.72
Predicted NSSHARE: PCTBLACK = 0 (%)	1.4	1.5	1.3	1.4	3.1	3.6
Predicted NSSHARE: PCTBLACK = 100 (%)	10.6	3.7	10.1	3.8	15.3	4.7

NOTE.—Observations are for 270 Missouri urban zip codes using aggregate insurance market data from 1988 through 1992; coefficients with absolute t-statistics using White standard errors are in parentheses below; coefficients on percent variables are multiplied by 100. PCTBLACK is the percentage of the population that is black, log TOTPOP is the natural logarithm of total population, PCT1824 is the percentage of the adult population between the ages of 18 and 24, PCT55UP is the percentage of the adult population that is 55 years or older, MARRIED is the percentage of households with a married couple, PCTUNEMP is the percentage of the population that is unemployed, MEDLIMITS is the percentage of liability policies with medium limits (exposure class 2 in the appendix), and HIGHLIMITS is the percentage of liability policies with high limits (exposure classes 3, 4, and 5 in the appendix). Predicted NSSHARE is the predicted value of nonstandard share for the specified value of PCTBLACK with other variables equal to the sample means. OLS = ordinary least squares.

tary market) in each multivariate regression, as opposed to including the log of average car value in the collision and comprehensive equations (or including both types of variables in each equation). Table 7 also shows the predicted nonstandard market shares for PCTBLACK equal to 0% and 100% when the other explanatory variables are set to their sample means.[32]

As expected given the univariate comparisons in table 2, the simple

32. The predicted percentages equal $100[z/(1 + z)]$, where z is the predicted (unlogged) odds ratio. As an alternative, we approximated unbiased predictors of the odds ratio assuming normally distributed disturbances and using $\exp(u + .5\sigma^2)$ where u is the predicted log odds ratio and σ^2 is residual variance. The variance adjustment had little effect on the predicted percentages.

regressions indicate a strong positive relation between nonstandard market share and PCTBLACK. As predicted given the evidence in table 6 that PCTBLACK is positively correlated with omitted factors that increase claim costs, the reduced forms also indicate a positive and significant relation between PCTBLACK and nonstandard market share for each line of business, although the magnitude of the predicted relation is much smaller when the control variables are included. While the nonstandard market share analysis by itself cannot distinguish the omitted-variable explanation of this positive relation from a discrimination explanation, our loss ratio and average claim cost results clearly favor the omitted-variable explanation. Consistent with older, married, and wealthier individuals having lower expected claim costs (and perhaps greater persistency), nonstandard market share is negatively related to PCT55UP, MARRIED, and HIGHLIMITS. Results for a variety of other specifications had similiar implications for PCTBLACK (including models that omitted PCTUNEMP and included PCTPOOR, MEDHINC, PCTHIGHS, and fixed county effects).

We also used instrumental variables estimation to estimate nonstandard share models that included the predicted log of average claim costs from a regression of the log of average claim costs on the log of average premiums, given the random variation (measurement error) in average claim costs compared with expected claim costs (also see n. 27). Because the average level of expected claim costs for drivers in a zip code should be related to omitted cost factors, including this variable as a regressor might reduce omitted-variable bias. However, a positive relation between nonstandard share and PCTBLACK is still predicted to the extent that, for a given mean of expected claim costs, PCTBLACK is positively correlated with the proportion of drivers with high enough expected claim costs to make coverage in the nonstandard market likely (or a negative correlation with factors related to persistency). The coefficient on the average claim-cost variable was positive and highly significant for each line, and the coefficient for PCTBLACK declined in size compared with table 7 (and was not significant for liability coverage). Given the results of our loss ratio and average claim-cost equations, this finding suggests that PCTBLACK is correlated with omitted cost factors related to the proportion of high-risk drivers and/or omitted economic factors that could affect renewal rates, nonpayment of premiums, and so on.[33]

33. To explore whether crime rates might be important omitted variables in the average claim cost and nonstandard share equations, we obtained 1993 crime rates (offenses per thousand people) for a variety of violent crimes and motor vehicle theft for nine St. Louis and four Kansas City police districts. By overlaying police-district maps and zip-codes maps, we estimated crime rates for 47 zip codes in St. Louis and Kansas City. We then reestimated the average claim cost, nonstandard share, and, for completeness, loss ratio equations using just these 47 zip codes. While PCTBLACK and the estimated crime rates

VI. Summary and Conclusion

Following Becker's (1993) suggestion to analyze profitability to test for discrimination, we examine whether automobile-insurance market loss ratios are related to percent minority population. Compared with tests for discrimination in lending markets that use rejection rates or default rates, our analysis of loss ratios is much less susceptible to bias from omitted variables that could be positively correlated with both the risk of loss and percent minority population. Controlling for available demographic and coverage related factors that could influence auto-insurance loss ratios, our main finding is that loss ratios are not negatively related to percent minority population. This result is inconsistent with the hypothesis that racial discrimination increases premiums relative to expected claim costs for minorities. Absent significantly lower transaction costs or reductions in quality for minorities that would materially increase insurer profits for any given level of the loss ratio, this finding implies that insurer profits are not higher for minorities.

Our loss ratio results suggest that auto-insurance-affordability problems in urban areas with large minority populations reflect higher costs of providing coverage rather than discrimination. Indeed, univariate comparisons with the Missouri data indicate a strong positive relation between average claim costs and percent minority population. In addition, our regression equations for average claim costs that include a variety of control variables generally indicate a significant positive relationship between average claim costs and percent minority population, which is consistent with omitted cost factors that are positively related to percent minority population. As predicted given this result, we also find that nonstandard market share is positively and significantly related to percent minority population after controlling for our demographic and coverage variables. When viewed in isolation (and as is also true for studies of mortgage lending that focus on rejection rates), our nonstandard market share results cannot distinguish whether this positive relation is due to discrimination or omitted cost factors. However, the combination of our loss ratio results, which are inconsistent with discrimination, and our average claim-cost results, which are consistent with omitted cost factors correlated with percent minority population, provide strong support for the omitted-variable explanation.

It would be desirable to have further research on omitted cost factors that could be correlated with claim costs and percent minority population, as well as research on whether quality differs materially between blacks and whites, after controlling for other factors that affect the demand for quality, and, if so, whether reduced quality for minorities increases insurer profits. These types of research will be difficult with

---

generally were positively correlated, the inclusion of the crime-rate variables did not alter the main conclusions.

existing data. Because of unavailability of data, we could not control for possible differences in crime or theft rates in large numbers of zip codes, for possible differences by race in the age and sex distribution of drivers, or for differences in transactions costs that could be related to product quality. In principle, it might be desirable to measure some of the demographic variables for the driving population, as opposed to the total population, or possibly to use data at the individual policyholder level rather than aggregate data at the zip-code level. However, requiring insurers or government agencies to develop these types of data would be costly, and it would still be necessary to use population measures for a number of environmental factors.

Appendix

Additional Data Description

Insurance Market Data

The five policy (insurer) types for which data are reported to the Missouri Department of Insurance are defined by the Missouri Department as follows:

A. *Preferred family.*—A policy form at least equal to family automobile insurance ordinarily offered to risks meeting high acceptance standards at rates less than the industry average.
B. *Standard family.*—A policy form at least equal to family automobile insurance ordinarily offered to risks categorized as better than average at rates at or near the industry average.
C. *Standard risk.*—A policy form of the basic automobile type ordinarily offered to risks evaluated as average or slightly below average at rates at or slightly above the industry average.
D. *Nonstandard basic.*—A policy form of the basic automobile type ordinarily offered to risks evaluated as poor or below average at rates considerably greater than the industry average.
E. *The JUA Basic.*—A policy written under the Joint Underwriting Association.

The category definitions imply that rates increase and underwriting standards decline as one moves from category A to category D. Conversations with the Missouri Department of Insurance indicated that coverage terms were similar across these categories and that differences among the categories primarily reflect differences in rates and underwriting standards across insurers. The bulk of reported exposures (90%–95%) for the voluntary market are in categories A and B with relatively few exposures in category C.

The exposure class definitions are shown in table A1. We define average car value as the weighted average (using the percentage of policies in each exposure class as weights) of the midpoints of the auto-value intervals for each class.

TABLE A1 Exposure Class Definitions

Class	Collision and Comprehensive Car Value ($)	Liability Limits ($)	
		Split	Single
1	0–3,700	25,000/50,000	60,000–100,000
2	3,701–8,000	50,000/100,000	100,000–300,000
3	8,001–17,500	100,000/300,000	300,000–500,000
4	17,501–24,000	250,000/500,000	500,000–1,000,000
5	24,001 and above	500,000/1,000,000	1,000,000 and above

TABLE A2 Correlation Coefficients for Primary Demographic and Insurance Coverage Variables

	PCT-BLACK	log TOTPOP	PCT1824	PCT55UP	MARRIED	PCT-UNEMP	log AVGCAR	MED-LIMITS
log TOTPOP	.29							
PCT1824	.21	.28						
PCT55UP	.07	−.09	−.33					
MARRIED	−.70	−.39	−.38	−.10				
PCTUNEMP	.72	.12	.24	.15	−.55			
log AVGCARV	−.26	.21	−.10	−.09	.15	−.46		
MEDLIMITS	−.10	.02	.03	−.13	.16	.06	−.10	
HIGHLIMITS	−.29	.17	−.14	−.08	.21	−.49	.98	−.25

Control Variables and Loss Ratios

Definitions and descriptive statistics for our demographic and coverage variables are shown in tables 1 and 3, respectively. Table A2 shows the correlation matrix for the primary control variables used in the analysis.

As noted in the text, λ_i (the present value of insurers' costs per dollar of expected claim costs) will depend on (1) the timing of claim payments, (2) the risk-adjusted discount rate, (3) the amount of capital held by insurers, and (4) the expense loading for insurer underwriting, distribution, claim settlement, and policy service costs. The speed with which claims are paid and the amount of insurer capital are dimensions of product quality, the demand for which is likely to vary with income or wealth.[34] Other quality characteristics that could increase insurer operating expenses also should depend on differences in demand. Insurer underwriting expense ratios (total underwriting and distribution expenses as a proportion of premiums) and thus loss ratios vary with insurer distribution systems, with direct writers generally having lower expense ratios and higher loss ratios than insurers that use independent agents.[35] Direct writers also are known on average

34. Unobserved differences in property damage versus bodily injury liability mix could be related to income or wealth, although the effect is theoretically uncertain. Claim-settlement expenses will be lower but claims will tend to be paid faster in zip codes with higher proportions of property damage liability claims. The effect of this mix on λ_i is thus ambiguous.

35. Whether this difference reflects a difference in service quality has been disputed (e.g., Cummins, Weiss, and Berger 1997).

for insuring drivers with lower expected claim costs than auto insurers that use independent agents, and direct writers tend to focus on business with high persistency. Insurers with customer bases with higher policyholder persistency (renewal rates) will tend to have lower expense loadings because front-end expenses can be amortized over a longer period. Persistency could depend on demographic characteristics, such as population mobility and income. Expense ratios and thus loss ratios also could depend on economies related to multiple-car discounts and possible economies of scope from cross-selling between the marketing of auto insurance and other types of insurance to a given household.

It is possible that expense loadings and loss ratios also might be related to expected claim costs. For example (and as noted in n. 12), if some underwriting expenses do not increase in direct proportion to expected claim costs (e.g., fixed costs of processing an application), loss ratios could be related positively to factors that increase expected claim costs. Thus, any underwriting expenses that are less than proportional to expected claim costs could cause loss ratios to be higher, for example, for young drivers and in large urban areas with higher expected claim costs. However, truly fixed costs are likely to be small, and it is arguable whether transaction costs within a state would be expected to increase less than proportionately in these cases. For example, it is possible that required compensation for agents and rental costs of facilities could increase proportionately (or even more than proportionately) with expected claim costs in urban areas or that risk selection costs could be proportionately higher for young drivers.

With respect to our specific control variables, increased product quality, which should reduce the loss ratio, will likely be positively related to income, wealth, and coverage levels. The market share of direct writers, which also should be positively related to the loss ratio due to lower expense ratios, will likely be positively related to income, wealth, coverage levels, and perhaps PCT55UP and MARRIED, and possibly negatively related to PCT1824, given that direct-writer marketing strategies often emphasize business with greater persistency and opportunities for cross-selling other coverages. The effects of education could be similar to those of increased income because education could be correlated with unmeasured differences in income and wealth. We have no prior expectations concerning the effects of age distribution and marital status on the demand for product quality, other factors held constant, or concerning the possible effects of population density. Increases in population density could tend to increase the loss ratio if expected claim costs increase with density and transaction costs increase less than proportionately.

References

Ashenfelter, O., and Rees, A. 1973. *Discrimination in Labor Markets.* Princeton, N.J.: Lexington.

Becker, G. 1993. Nobel lecture: The economic way of looking at behavior. *Journal of Political Economy* 101 (June): 385–409.

Belsley, D. 1991. *Conditioning Diagnostics: Collinearity and Weak Data in Regression.* New York: Wiley.

Berkovec, J.; Canner, G.; Gabriel, S.; and Hannan, T. 1994. Race, redlining, and residential mortgage loan performance. *Journal of Real Estate Finance and Economics* 9 (November): 263–94.

Butler, R. J. 1994. Economic determinants of workers' compensation trends. *Journal of Risk and Insurance* 61 (September): 383–401.

Chan, T. 1997. Service quality and unfair racial discrimination in homeowners insurance. Working paper. Jefferson City: Missouri Department of Insurance.

Cummins, J. D., and Tennyson, S. 1992. Controlling automobile insurance costs. *Journal of Economic Perspectives* 6 (Spring): 95–115.

Cummins, J. D.; Weiss, M.; and Berger, A. 1997. The co-existence of multiple distribution systems for financial services: The case of property-liability insurance. *Journal of Business* 70 (October): 515–46.

D'Arcy, S. P., and Doherty, N. A. 1990. Adverse selection, private information, and lowballing in insurance markets. *Journal of Business* 63 (April): 145–64.

Dionne, G., and Doherty, N. A. 1994. Adverse selection, commitment, and renegotiation: Extension to and evidence from insurance markets. *Journal of Political Economy* 102 (April): 209–35.

Ferguson, M., and Peters, S. 1995. What constitutes evidence of discrimination in lending? *Journal of Finance* 50 (June): 739–48.

Ferguson, M., and Peters, S. 1997. Is lending discrimination always costly? Working paper. Bloomington: Indiana University, School of Business.

Harrington, S. 1987. A note on the impact of auto insurance rate regulation. *Review of Economics and Statistics* 69 (February): 166–70.

Harrington, S. 1992. Rate suppression. *Journal of Risk and Insurance* 59 (June): 185–202.

Harrington, S., and Niehaus, G. 1992. Dealing with insurance availability and affordability problems in inner cities: An analysis of the California proposal. *Journal of Insurance Regulation* 10 (Summer): 564–84.

Harrison, G. 1998. Mortgage lending in Boston: A reconsideration of the evidence. *Economic Inquiry* 36 (January): 29–38.

Holmes, A., and Horvitz, P. 1994. Mortgage redlining: Race, risk, and demand. *Journal of Finance* 49 (March): 81–99.

Insurance Redlining Practices. 1993. Hearings before the Subcommittee on Commerce, Consumer Protection, and Competitiveness of the Committee on Energy and Commerce, U.S. House of Representatives.

Klein, R. 1989. Competition in private passenger automobile insurance: A report to the Personal Lines (C) Committee. Working paper. Kansas City, Mo.: National Association of Insurance Commissioners.

Klein, R. 1995. Urban homeowners insurance markets: Problems and possible solutions. Working paper. Kansas City, Mo.: National Association of Insurance Commissioners.

Munnel, A.; Browne, L.; McEneaney, J.; and Tootell, G. 1996. Mortgage lending in Boston: Interpreting HMDA data. *American Economic Review* 86 (March): 25–53.

Myers, S. C., and Cohn, R. 1986. A discounted cash flow approach to property-liability insurance rate regulation. In J. D. Cummins and S. E. Harrington (eds.), *Fair Rate of Return in Property-Liability Insurance.* Boston: Kluwer.

Peterson, R. 1981. An investigation of sex discrimination in commercial banks' direct consumer lending. *Bell Journal of Economics* 12 (Autumn): 547–61.

Rachlis, M., and Yezer, A. 1993. Serious flaws in statistical tests for discrimination in mortgage markets. *Journal of Housing Research* 4:315–36.

Rothschild, M., and Stiglitz, J. 1976. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics* 90 (November): 629–50.

Shaffer, S. 1996. Evidence of discrimination in lending: An extension. *Journal of Finance* 51 (September): 1551–4.

Smallwood, D. 1975. Regulation and product quality in the automobile insurance industry. In A. Phillips (ed.), *Promoting competition in regulated markets.* Washington, D.C.: Brookings.

Squires, G. D.; DeWolfe, R.; and DeWolfe, A. 1979. Urban decline or disinvestment: Uneven development, redlining, and the role of the insurance industry. *Social Problems* 27 (October): 79–95.

Squires, G. D.; Velez, W.; and Taeuber, K. E. 1991. Insurance redlining, agency location, and the process of urban disinvestment. *Urban Affairs Quarterly* 26 (June): 567–87.
Stegman, M., and Goering, J. (eds.). 1996. Race and default in credit markets: A colloquy. *Cityscape: A Journal of Policy Development and Research* 2:1–126.
Stiglitz, J., and Weiss, A. 1981. Credit rationing in markets with imperfect information. *American Economic Review* 71 (June): 393–410.
Yezer, A.; Lindsey, L.; and Issac, W. (eds.). 1995. *Fair Lending Analysis: A Compendium of Essays on the Use of Statistics.* Washington, D.C.: American Bankers Association.