Can Gender-Blind Algorithmic Pricing Eliminate the Gender Gap?*

Ozge Demirci†

April 23, 2023

Click here for the most recent version

Abstract

Insurance companies frequently use consumer attributes, such as gender or age, when setting a price for their services. Young male drivers, for example, are often charged more than young females for car insurance, as they are expected to be riskier. In 2019, California banned auto insurance companies from using information on gender in their pricing algorithms. I study how this ban affects the gender gap in prices, using a difference-in-differences strategy with older individuals and other states as control groups. I find that the ban reduced the gender gap in the insurance premiums paid by young drivers by around 65 percent, but it failed to eliminate it completely. My analysis of the pricing algorithm of a large insurance company in California indicates that algorithms are adjusted in a way that gender proxies receive larger weights after the policy. For instance, drivers using specific car models associated with young males were charged up to 22 percent more after the ban. My findings illustrate the limitations of anti-discrimination policies that impose group-blind pricing, with implications for the design of fairer algorithmic regulations.

Keywords: Algorithmic pricing; gender discrimination; machine learning

JEL Codes: D81, G22, J16

*I am grateful to Manuel Bagues, Mirko Draca, and Ludovica Gazze for their invaluable mentorship and advice. This paper has benefited greatly from feedback and discussions with Sascha Becker, Dan Bernhardt, Paul Belleflamme, Luis Candelaria, Christopher Carpenter, Daniele Condorelli, Maura Coughlin, Murat Demirci, Will Dobbie, Dita Eckardt, Oliver Falck, James Fenske, Thiemo Fetzer, Axel Gautier, Clement Imbert, Anne Kerkhof, Amrita Kulka, Morgane Laouenan, Dina Mayzlin, Carlo Perroni, Petra Persson, Christian Peukert, Michelangelo Rossi, Stephen P. Ryan, Stephan Seiler, Almudena Sevilla, Avner Strulov-Shlain, Caio Waisman, Joel Waldfogel, Ao Wang, and Natalia Zinovyeva. I also thank participants at different seminars and meetings at the University of Warwick, Paris Conference on Digital Economics, European Economic Association Meeting, European Association of Labour Economists Meeting, Royal Economic Society Conference, ACM Conference on Economics and Computation, ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization, ifo Institute, Discrimination, and Diversity Workshop. I gratefully acknowledge financial support from the Royal Economic Society.

†Department of Economics, University of Warwick. Email address: o.demirci@warwick.ac.uk.

1 Introduction

There is increasing reliance on algorithms for making important economic and managerial decisions. Advocates for the technology claim that automated computations can increase efficiency and objectivity in decision-making. However, as a growing literature reveals, in several fields algorithms are associated with biased and unfair outcomes.[1] One possible source of algorithmic bias is the use of group attributes to make individual-level predictions (Ascarza and Israeli, 2022). However, it is often difficult to measure and tackle algorithmic bias due to the black-box nature of algorithms. For instance, even if laws require specific group attributes (such as race or gender) to not be used, the high-dimensional data enables companies to use characteristics that closely correlate with the attributes that are banned. This issue raises the question of whether anti-discriminatory policies that ban the use of certain predictable attributes in algorithmic processes are effective, or whether algorithms adjust to circumvent the regulations.

In this paper, I use a natural experiment to investigate the impact of gender-blind algorithmic pricing on consumers and firms. I focus on a recent policy change in California that prohibits using gender information in pricing algorithms for automobile insurance. I investigate how this policy affects the difference in the insurance premiums paid by male and female drivers and the pricing algorithms companies use. To answer the first part of my research question, I rely on the triple difference-in-differences strategy, using data from the Consumer Expenditures Survey. Young male drivers under 25 often pay more for car insurance than young females; since young males are more likely to be riskier drivers (Huh and Reif, 2021, Moore and Morris, 2021), insurance companies use gender as a group characteristic to assign prices for young drivers. On the other hand, the gender gap does not exist among more experienced drivers, presumably because companies collect more information about individuals' driving history and use this information instead of gender when determining the price. The ban on using gender in insurance pricing in California is supposed to affect the prices mainly among young drivers. This setting allows me to use variation across age groups to identify the causal effects of gender-neutral pricing policy on prices. I compare the gender gap among young people with the gap among older people in California.

I find that, after controlling for a set of characteristics on demographics, location, and vehicles, the initial gender gap between young people has decreased by 13 percentage points following the gender-neutral pricing regulation.[2] Specifically, insurance prices increased for young females and decreased for young males. Although the ban reduced some parts of the

[1] See examples of algorithmic bias in hiring, Angwin et al. (2016), Arnold et al. (2021); and in advertisement targeting, Datta et al. (2014), Lambrecht and Tucker (2019).

[2] These characteristics include age, gender, marital status, education level, geographic region, population density, metropolitan statistical area status, vehicle type, vehicle year, vehicle brand, and new/used vehicle status. These are the features commonly used in the US auto insurance industry. However, in practice, we expect that insurance companies may use additional variables. Please refer to Section 4 for a more detailed discussion.

gender gap, 35 percent of the initial gap among young drivers remains after the policy. As an alternative identification strategy, I use other states with gender-based insurance pricing as a control group. As aligned with the findings obtained from the analysis within California, the policy reduced the gender gap. However, it fell short of eliminating price differences between young males and females in California compared to young drivers in control states.

Next, I investigate whether the remaining gender gap results from the adjustments in the pricing algorithms.[3] Specifically, I analyze whether pricing algorithms started to re-price features correlated with gender after the gender-neutral pricing policy. To do so, I construct a novel dataset of insurance pricing algorithms based on price filings from a large insurance provider in California. This dataset includes three essential features of pricing: (i) a list of factors used in pricing algorithms such as age, marital status, car models, etc., (ii) weights for each factor, i.e., how much it contributes to the pricing, and (iii) a pricing rule that combines price factors with their corresponding weightings to set a price. This setting allows me to estimate not only the changes in the insurance prices but also the changes in the price-generating processes. As the opaqueness of algorithmic processes is one of the main challenges in algorithmic bias literature, this dataset provides me with a rare opportunity to study bias in algorithmic pricing transparently.

To be able to analyze whether the factors associated with gender get more weight in the pricing algorithms after the ban, I first need to know which factors are highly correlated with gender. To detect these factors, I use National Household Travel Surveys. This dataset provides very detailed consumer characteristics that are aligned with the information collected by insurance companies. It is large enough that I can conduct a prediction analysis for gender with more than half a million drivers using more than 500 hundred features, including different individual, location, and vehicle characteristics. I adopt different estimation techniques, including logistic regressions and supervised machine learning (ML) models such as regularized regressions. I apply this broad set of methods so that I can determine the best subset of variables that predict being a young male, the riskiest age and gender group.[4] Compared to traditional stepwise selection methods based on logistic regressions, the supervised machine learning techniques have advantages in their performance for out-of-sample predictions and predictive accuracy. The prediction analysis selects 58 distinct features as a predictor of being

[3]The term pricing algorithm refers to the practice of automatically setting prices based on predetermined pricing rules and inputs. In auto insurance, this includes (i) variables used in pricing, (ii) their weights, and (iii) the price function that combines the first two components to determine prices. Throughout this paper, I will refer to these mechanisms as pricing algorithms.

[4]The prediction analysis focuses specifically on being a young male instead of a male or female. The main reason is that young males are the riskiest group for car insurance, as discussed previously and shown in the empirical analysis Section 5.1. Hence, banning gender information in pricing can create an incentive for firms to differentiate young males from others. As companies can still use age (or experience) in their pricing analysis after the regulation, features that differentiate young males from young females will be important. In that sense, in the prediction, by using age as a covariate (both linearly and non-linearly in different specifications), I predict young male characteristics conditional on being young and a wide range of other characteristics.

a young male.[5] This exercise allows me to investigate changes in these predictors in the pricing algorithm after the policy.

I analyze how weights for car models predicted as being correlated with 'young male' are changed in the pricing algorithm. I find that the algorithms started to increase the weightings of car models associated with young males by 10 percent compared to the weighting before the policy change. Additionally, weights for these young male cars increase by 11 percent relative to other car models that are not predicted as young male features. This result reflects the change in the pricing algorithm through changes in the weightings of some car models. A potential implication is spillover effects onto other demographic groups. After the policy change, the insurance price of these young male cars increases; hence any driver of these models starts to pay more for insurance.[6]

To guide the empirical analysis, I use a theoretical framework in which insurance firms predict individual risk for each consumer based on the information they collect. The model builds upon statistical discrimination studies (Aigner and Cain, 1977; Altonji and Pierret, 2001), but also it incorporates firms' responses to restrictions on monitoring some consumer characteristics. In the model, firms have limited information about consumers' risk attitudes while driving, which is called a signal for driver risk. Firms resort to group characteristics, such as gender, on the assumption that group characteristics are correlated with risk. Firms calculate an expected value of risk for each consumer based on the risk signal and the distributional prior for group characteristics. However, if there is a regulation that stops firms from basing calculations on group information, they predict group characteristics based on other observables. The model suggests two main predictions regarding firms' ability to predict groups. First, the expected risk differences between groups become smaller in the blind case where firms cannot observe group characteristics, compared to the non-blind scenario. This is driven by firms' limited ability to predict banned variables. Second, in the blind case, the signals correlated with banned group characteristics get larger weights in risk prediction than in the non-blind case. My empirical analysis provides evidence for both predictions.

This study contributes to algorithmic bias literature in two main ways. First, the growing algorithmic bias literature focuses on measuring and tackling algorithmic bias and discusses potential reasons for it (Lambrecht and Tucker, 2019, Cowgill, 2018, Cowgill and Tucker, 2020, Kleinberg et al., 2020, Arnold et al., 2021, Ascarza and Israeli, 2022). A common finding in this literature is that even when the algorithms do not intend to discriminate, they may create unfair and biased outcomes due to their design or the unrepresentative (or biased) training data. This

[5] Among these 58 features, 4 are demographic features such as education level and home ownership status that are elicited by the company but not used in the pricing algorithm. Three of them are vehicle categories such as car, pickup, or truck, 4 are car brands, and 47 are car models. As car models cover car types and brands, I continue my analysis with these 47 distinctive car models. For more details, please refer to Section 6.1.3.

[6] Using the consumer expenditure survey dataset, I also find confirmation that the owners of car models associated with young males started to pay more.

paper complements previous studies by providing empirical evidence on the effects of regulations on algorithmic bias. My findings highlight the (in)effectiveness of imposing gender-blind pricing, because there are underlying correlations between regulated variables and other covariates in the data. Second, the empirical evidence on algorithmic bias often relies on data from the realized outcomes, such as hiring decisions or prices, but not on the data from the decision-making process itself. In contrast, analyzing the auto insurance industry provides a salient advantage in studying algorithmic bias, as in most US states it is a regulatory requirement for companies to disclose their pricing algorithms. I took advantage of this to assemble a pricing algorithm dataset based on information from a large insurance provider in California. With this dataset I can observe the price-generating process itself and analyze the changes in the algorithm after the gender-blind pricing regulation.[7]

Another strand of literature this study contributes to is regulations in insurance pricing.[8] In the EU, together with the EU Gender Initiative in 2011, member states are required to use gender-neutral pricing for any insurance to achieve gender-equal outcomes. In the US, California banned gender-based pricing for car insurance in 2018. Many scholars focus on these regulatory reforms banning the use of group information in different insurance markets such as annuities (Finkelstein et al., 2009) and health insurance (Huang and Salm, 2020).[9] To the best of my knowledge, my work is the first to provide an empirical investigation of the gender-blind auto insurance policy in the US. Also, unlike the earlier studies that rely on consumer data, my paper contributes to the literature by examining the effect of a gender-ban in pricing on a firm's pricing behavior by compiling data on the pricing algorithm of an insurance company.

Similar to insurance markets, there are various regulations in labor markets that limit discrimination based on features of applicants in the hiring process. For instance, 'Ban the Box' policies stop firms from asking about criminal records in job applications until a late stage. These policies aim to prevent potential discrimination against ex-offenders in the labor market. However, a common finding is that under these policies, employers start to use other observables such as race to proxy for a criminal conviction (Agan and Starr, 2018, Doleac and

[7]A parallel strand of recent literature focuses on the relationship between algorithmic pricing and market structures. While Calvano et al. (2020); Klein (2021); Brown and MacKay (2021); Asker et al. (2021) provide theoretical evidence of independent algorithms leading to supra-competitive prices, Assad et al. (2020) discusses empirical evidence on the relationship between algorithmic pricing and market competition. This paper complements this work by providing evidence of the impact of regulations on algorithmic pricing.

[8]Insurance firms predict individual risk for each driver based on the collected information, including individual and vehicle characteristics. In that sense using all information, including gender, that may reflect risk, especially for young drivers, is economically efficient for companies. Adverse selection and moral hazard issues may arise if they cannot differentiate between groups with different perceived risks (Puelz and Snow, 1994, Finkelstein and Poterba, 2004, Cohen, 2005). On the other hand, using statistical information on group characteristics to infer individual risk may be discriminatory (Arrow, 2015, Phelps, 1972), and it is deemed illegal (Avraham et al., 2013). Therefore, the trade-off between efficiency and equality has created public policy debates in insurance pricing.

[9]There are also theoretical studies that focus on group-based pricing in insurance and find that restrictions on these characteristics in pricing cause efficiency loss (Hay and Leahy, 1982, Crocker and Snow, 1986, Rea, 1987, Polborn et al., 2006).

Hansen, 2020). Hence, these policies have unintended outcomes such as a lower likelihood of getting an interview or employment for African-American males. Other studies on credit history bans (Bartik and Nelson, 2016) and salary history bans (Agan et al., 2020, Hansen and McNichols, 2020) in the job market screening process provide similar results. In that sense, my study contributes to a growing literature emphasizing that anti-discrimination policies that remove information about different applicant characteristics might create unintended outcomes and sometimes cause even more harm. One crucial difference between these studies and my paper is the decision-making process leading to discrimination. In most labor discrimination studies, statistical discrimination is based on accurate or inaccurate human beliefs, and the decision-making process may not be fully monitored. In this study, insurance discrimination relies on algorithmic decision-making, and the decision process is relatively more discernible. Therefore, this study provides an alternative perspective to study the problem of removing information in screening processes by incorporating algorithms as decision-makers.

The structure of the paper is as follows. In Section 2, I describe the US automobile insurance market. In Section 3, I discuss the theoretical background and in Section 4, I describe data used in empirical analysis and main descriptive statistics. In Section 5, I present the empirical analysis of changes in insurance prices. In Section 6, I provide the empirical results on changes in pricing algorithms. In Section 7, I present a series of placebo and robustness checks. In Section 8, I conclude the paper with some policy discussions and a summary of the results.

2 Background on the Automobile Insurance

Auto insurance is one of the most common types of personal insurance, and it is mandatory in most US states. Nearly 215 million people (88 percent of all drivers) hold a car insurance policy (Council, 2021). In 2021, the size of the auto insurance industry was estimated at approximately 249.4 billion dollars for private automobile insurance, representing 35 percent of the total US insurance industry.[10] There are two broad categories for automobile insurance: liability and property coverage. Liability insurance covers property damage and bodily injuries to another person caused by an accident where the insured person is at fault. Property insurance provides coverage for one's car regardless of fault. All states require liability insurance, but different states mandate different minimum levels of liability insurance.[11]

Insurance Pricing Insurance companies in the US are required to submit detailed filings to regulatory agencies in each state and explain the details of their pricing.[12] Each state has its own regulations to ensure insurance pricing is fair and affordable for all drivers, and prevent

[10] Calculated based on financial statements published by NAIC, 2021.

[11] Although minimum requirements vary across states, all US states except New Hampshire mandate liability insurance.

[12] Wyoming is the only state that does not require publishing filings.

excessive pricing for some demographic groups. State-level insurance laws specify which features can and cannot be used as risk factors in pricing.

Companies have to show which risk factors they use in their pricing algorithm and the risk scores attached to each risk factor. These risk factors are grouped into three categories: driver, vehicle, and location characteristics. They usually include age, gender, marital status, ZIP code, vehicle make and model, and driving history. In addition to these factors, an insurance policy price has a base rate that varies over the state and insurance coverage choice such that choosing more extensive policy coverage increases its price. A risk score shows how risky a specific feature is seen relative to other features in that risk category.[13] For instance, a risk score for single male drivers with no driving experience indicates an expected risk relative to single male drivers with driving experience. Figures B1 to B4 in Appendix B provide some examples of risk score tables based on price filings from a major insurance provider.[14]

Industry regulations require auto insurance companies to disclose their pricing algorithm, including risk factors, risk scores, and pricing rules. Generally, the pricing rule can be summarized as:

$$\text{Price} = (\text{Base rate x driver x vehicle x location characteristics}) + \text{c} \tag{1}$$

where c is a term for a firm's administrative and maintenance costs. Driver and vehicle characteristics include age, gender, marital status, driving histories, annual miles, vehicle brand, and model.[15] The auto insurance price also depends on location-specific features such as ZIP code level population density or past claims. These customer characteristics and related risk scores for each characteristic (such as risk score for car model or individual features) enter the pricing algorithm as shown in Equation (1) and create a price. Figures B5 and B6 in Appendix B provide some examples of rate-making logic.

Regulation The auto insurance industry in the US is heavily regulated. In addition to the filing disclosure requirements discussed above, many states have legislation to prevent discrimination. Regulations are designed to make insurance prices fair and affordable to different demographic groups (Hunter et al., 2013) and there are different restrictions on the use of group-based characteristics in insurance pricing. Supporters of regulation often argue that using group-based characteristics beyond individuals' control rather than individual-level risk is discriminatory. On the other hand, pooling customers according to their risk characteristics helps firms alleviate adverse selection and moral hazard issues. Alongside these public debates,

[13] In insurance terminology, risk scores are often referred to as risk relativities.

[14] Examples provided in Appendix B were extracted from the price filings of a large insurance provider in California.

[15] The set of characteristics that are allowed to be used may vary across states depending on the regulations in each state.

some states such as Massachusetts, North Carolina, Pennsylvania, and Hawaii restricted using gender as a risk factor in the 1980s, and California joined them in 2019. Other than gender, different states have insurance pricing legislation on the use of credit scores, occupation, education, and income characteristics.

3 Theoretical Framework

In this section, I lay out a theoretical framework to explain the impacts of the gender-blind policy by adapting a signaling framework commonly used for studying statistical discrimination (Aigner and Cain, 1977, Altonji and Pierret, 2001, Lundberg and Startz, 1983). I introduce a model where firms try to predict individual-level risk by observables they have collected. The model makes theoretical predictions about what can happen when firms' ability to monitor risk predictors is constrained.

3.1 Environment

Assume that insurance firms cannot observe an individual i's actual risk while driving, but instead, they can observe a noisy signal of the individual risk level. The risk signal is defined as:

$$\mathrm{y_i} = \mathrm{x}_i + \eta_i$$

where $\mathrm{y_i}$ represents the risk signal of individual i, $\mathrm{x_i}$ is the true risk of individual i, and η_i denotes the noise in the signal. The noise η_i is assumed to be independently and normally distributed with zero mean and finite variance implying that $\eta_\mathrm{i} \sim \mathrm{N}(0, \sigma_{\eta_\mathrm{i}})$. Moreover, firms learn from experience that the individual risk $\mathrm{x_i}$ is normally distributed with mean $\bar{\mathrm{x}}$ and variance $\sigma_{\mathrm{x_i}}$. The accuracy of the signal is independent of individual risk, $\mathrm{E(x_i \eta_i)} = 0$.

The expected risk for individual i is the weighted average of the signal $\mathrm{y_i}$ and the distributional prior $\mathrm{p_i}$ for group characteristics. Here, signals can be considered as individual accident records, and distributional priors are group-based characteristics such as gender or postcode. The expected risk for individual i conditional on signal $\mathrm{y_i}$ and the individual characteristics $\mathrm{z_i}$ is given by

$$\mathrm{E(x_i|y_i, z_i)} = \gamma y_i + (1-\gamma) p_i$$

where $\mathrm{z_i}$ is the set of characteristics for individual i and $\gamma = \mathrm{E(x_i y_i)} \ / \ \mathrm{E(y_i y_i)} = \frac{\sigma_{\eta_i}}{\sigma_{\eta_i} + \sigma_{x_i}}$ which is the coefficient from a bivariate regression of $\mathrm{x_i}$ on $\mathrm{y_i}$ and a constant.

3.2 Restrictions on monitoring consumer characteristics

Let us define two cases depending on the firm's ability to monitor consumer i's characteristics.

1. Non-blind case: Firms can observe all signals and group characteristics. The expected risk for individual i will depend on the signal $\mathrm{y_i}$ and the distributional prior $\mathrm{p_i}$ for group characteristics. It can be written as:

$$\mathrm{E(x_i|y_i, z_i)} = \gamma y_i + (1-\gamma)p_i$$

2. Blind case: Firms cannot observe group characteristics, so first they estimate the probability of being from a particular group based on signal $\mathrm{y_i}$. For simplicity, let us assume there are two groups, $\mathrm{g_1}$ and $\mathrm{g_2}$. The probability of being from $\mathrm{g_n}$ where $\mathrm{n} \in 1{,}2$ is given by:

$$\mathrm{P(i \in g_n|y_i)} = \alpha_0 + \alpha_1 y_i + u_i$$

Next, the firm formulates an expected risk based on $\mathrm{E(p_i)}$, an expected value of distributional prior:

$$\mathrm{E(x_i|y_i)} = \gamma y_i + (1-\gamma)\mathrm{E(p_i)}$$

where

$$\mathrm{E(p_i) = P(i \in g_1|y_i) \times p_1 + P(i \in g_2|y_i) \times p_2}$$

denotes the weighted average distributional priors for different groups. There can be different outcomes, depending on firms' ability to predict the probability of being from a particular group, and hence the ability to calculate the expected value of distributional prior. For simplicity, let us consider two extreme cases:

- If firms cannot predict banned variables, then the expected value of risk will only depend on the signal:

$$\mathrm{E(x_i|y_i)} = \gamma y_i$$

- If firms can fully predict group characteristics, we can infer that the expected value of the distributional prior is equal to its true value, i.e., $\mathrm{E(p_i) = p_i}$. Hence, the expected value of risk conditional on the signal will be:

$$\mathrm{E(x_i|y_i, z_i)} = \gamma y_i + (1-\gamma)p_i$$

3.3 Predictions

Based on this framework and following Pope and Sydnor (2011), I generate different testable predictions. For the proofs of these testable predictions, please refer to Appendix A.

Prediction 1: The expected risk difference between two groups will be less in blind cases if firms cannot perfectly predict group characteristics. Let us assume group 1 is riskier than group 2 on average without loss of generality.

$$\begin{aligned}\mathrm{E}^{\mathrm{Non-blind}}(\mathrm{x_i}|\mathrm{y_i}, \mathrm{i} \in \mathrm{g_1}) - \mathrm{E}^{\mathrm{Non-blind}}(\mathrm{x_i}|\mathrm{y_i}, \mathrm{i} \in \mathrm{g_2}) >\\ \mathrm{E}^{\mathrm{Blind}}(\mathrm{x_i}|\mathrm{y_i}, \mathrm{i} \in \mathrm{g_1}) - \mathrm{E}^{\mathrm{Blind}}(\mathrm{x_i}|\mathrm{y_i}, \mathrm{i} \in \mathrm{g_2})\end{aligned}$$

Prediction 2: Signals $\mathrm{y_i}$ positively correlated with group characteristics will gain more weight in risk prediction.

$$\mathrm{E}^{\mathrm{Non-blind}}(\mathrm{x_i}|\mathrm{y_i}) = \gamma_0 y_i + (1 - \gamma_0) p_i$$

and

$$\mathrm{E}^{\mathrm{Blind}}(\mathrm{x_i}|\mathrm{y_i}) = \gamma_1 y_i + (1 - \gamma_1)\mathrm{E}(\mathrm{p_i})$$

Suggesting that, $\gamma_1 \geq \gamma_0$.

4 Data

To study the impact of gender-blind pricing on insurance prices and on pricing algorithms, I combine three main datasets on (i) car insurance prices, (ii) consumer characteristics, and (iii) pricing algorithms.

4.1 Car insurance prices

The data on car insurance prices are obtained through the US Consumer Expenditure Survey Microdata, a large-scale longitudinal dataset weighted to be representative of the US population. The repeated cross-sectional survey provides information about US consumers' quarterly expenditures, incomes, and consumption characteristics. It also includes information on how much people spend on car insurance and demographics such as age, marital status, occupation, education level, and location-based characteristics. The survey covers car characteristics such as car types, car brands, and car years. These characteristics are essential in the auto insurance context, as they are the main features collected by insurance firms to offer price quotes.[16] This paper focuses on the sample of consumers who own a single car and bought car insurance between 2010 and 2020 in all US states.[17] Having an extensive dataset on demographic and geographic characteristics allowed me to exploit spatial and age group variations in different identification strategies.

Table 1 provides information on the main summary statistics for the Consumer Expenditure Survey sample used in this paper. It shows insurance expenditures, vehicles, and individual characteristics for the population under and above 25 years old. I will refer to the people under (above) 25 years old as young (older) or inexperienced (experienced) drivers throughout the paper. Population under and above 25 are similar in different demographics except for the higher ratio of single people and university graduates among the younger population. More young drivers own second-hand vehicles, and we observe a significant difference in premiums paid by the two groups. Annual car insurance spending is $2,101 on average for people above 25 years old, whereas people under 25 pay approximately 38 percent more. This difference is greater for drivers under 20 (for more details on risky driving behavior and age, please refer to Section 5.1).

Although these young and old drivers do not differ much in their non-age demographics, they pay different prices for insurance for different reasons. First and foremost, young and old drivers can differ in expected risks, and companies try to infer underlying risks through the information they collect. However, the risk prediction may not always be very accurate. For instance, driving history can be a useful proxy for individual risk. However, young drivers do not have much driving experience, so there is no data for companies to collect. They can try to proxy it with group-based characteristics such as gender or marital status. Young people are statistically more likely to be involved in an accident, and this translates into higher premiums for inexperienced drivers.

4.2 Consumer characteristics

One of the main questions this paper aims to answer is whether firms change their pricing algorithms after the gender-neutral pricing policy and if they use gender-correlated features to compensate for the information loss. To analyze this, we need a dataset to observe the type

[16]The granularity of some variables in this dataset is not precisely the same as that collected by insurance firms. For instance, the consumer expenditure survey is anonymized so we cannot work at a ZIP code level. Still, I use other location-based features about population density. Most insurance companies use ZIP codes to infer population density and include this in their pricing calculations.

[17]The pricing rules for insurance policies with multiple cars are generally different than policies for single vehicles. As the analysis in this paper will focus on single-car insurance policies, I also restricted the sample to single-car owners for the expenditure survey dataset. Based on the balance check between the two groups, single car owners are more likely to be unmarried, own newer cars, have a bachelor's degree, live in a metropolitan statistical area, and own smaller cars compared to multiple car owners. Full descriptive statistics for the whole sample is available upon request.

Table 1: Summary of insurance prices dataset

	≤ 25 years old		> 25 years old		Full Sample	
	(1) Mean	St Dev	(2) Mean	St Dev	(3) Mean	St Dev
Female	0.49	0.49	0.51	0.49	0.51	0.49
Age	22.89	1.95	52.61	15.75	51.29	16.57
Single	0.72	0.44	0.42	0.49	0.44	0.49
White	0.81	0.39	0.82	0.38	0.82	0.38
Bachelor's degree	0.41	0.49	0.33	0.49	0.31	0.46
Living in an MSA	0.97	0.15	0.98	0.15	0.98	0.15
Insurance expenditure ($)	2,791	719	2,101	652.08	2,132	670
Automobile	0.76	0.42	0.69	0.46	0.69	0.46
Vehicle age	14.93	6.63	14.43	6.95	14.45	6.94
Used car	0.79	0.41	0.57	0.49	0.58	0.49
Observations	5,593		120,129		125,722	

Notes: This table is a summary of the main statistics for the US Consumer Expenditure Surveys between 2010 and 2020. Columns (1) and (2) focus on the population under 26 years old, and Columns (3) and (4) on the population above 25 years old. The last two columns give statistics for the entire sample used in this study. Female, single, white, bachelor's degree, living in an MSA, automobile, and used car are defined as dummy variables taking a value of 0 or 1. MSA refers to Metropolitan Statistical Area. Insurance expenditure reflects yearly expenditures in US dollars.

of information that car insurance companies collect, including gender and other characteristics that may correlate with gender. For this analysis, I use the US National Household Travel Survey (NHTS), a large-scale survey conducted every 8 years. This survey provides demographic characteristics (age, gender, education, occupation, etc.), detailed vehicle characteristics (vehicle brand, model, features, vehicle age, annual mileage), and location-specific characteristics. The sample used in the paper focuses on the last three waves: 2001, 2009, and 2017. It includes a stratified random sample of US households who own a vehicle at the survey time.

The information provided in the dataset largely overlaps with the data insurance companies collect from their customers. We can think of this dataset as very similar to a customer database for insurance companies.[18] In Section 6, using these extensive customer features and exploiting different statistical methods, I estimate features that predict gender. Specifically, I focus on characteristics associated with being a young male under 25 years old, as insurance companies define this as the riskiest group.

Table 2 shows descriptive statistics for 676,574 drivers and their characteristics for the entire sample, and for male drivers under 25. The average consumer is 53 years old, drives a 9-year-old vehicle, drives 9,660 miles per year, and lives with a population density of 3,577 people per square mile. Young males are more likely to use older cars and have higher annual miles and odometer readings than the full sample. We see other trends that appear to distinguish young males from the full sample in terms of car types. For instance, a larger share of young males

[18] The main difference is that this dataset is anonymized for postal code. However, it provides information about other geographical features, such as the core-based statistical area (CBSA) code which is a geographic area including one or more counties geographically grouped with at least 10,000 people.

has pickup trucks, whereas sport utility vehicles (SUVs) are less common. Young males also prefer some specific car models and brands, excluding SUVs and pickups, but this representation incorporates many car models. A more detailed analysis of vehicle types, including different car models and brands, is provided in Section 6. For a more detailed version of summary statistics, please refer to Table B1.

Table 2: Summary of consumer characteristics dataset

	Full Sample		Young males	
	(1) Mean	St Dev	(2) Mean	St Dev
Vehicle year	2002.53	8.39	2001.2	7.93
Vehicle age	9.56	7.46	10.87	6.88
Annual miles	9,660.48	11,035.35	10,612	13,676.83
Odometer reading	83,504.98	67,419.88	106,162.50	68,192.57
Car	0.49	0.49	0.59	0.45
Sport utility vehicle	0.21	0.38	0.13	0.25
Pickup	0.16	0.36	0.21	0.31
Age	53.83	18.38	18.65	3.47
White	0.84	0.36	0.78	0.41
Bachelor's degree	0.19	0.39	0.07	0.26
MSA over 1M population	0.16	0.37	0.18	0.39
Housing units per sq mile	1,580.62	2,983.88	1,708.23	3,073.41
Population density per sq mile	3,577.73	5,169.39	4,118.48	5,694.11
Observations	676,574		29,731	

Notes: This table summarizes the main statistics for the NHTS datasets for the entire sample used in this study and only for young males under 25 separately. Car, sport utility vehicle, and pickup are dummy variables taking values of 0 or 1 and representing different vehicle categories. MSA over 1M population refers to Metropolitan Statistical Areas with at least 1 million.

4.3 Data extraction: Pricing algorithms

Auto insurers in the US have to disclose their pricing details to insurance regulatory agencies in each state. These filings include information on (i) risk factors used in pricing calculations, (ii) risk scores attached to each risk factor and category, and (iii) a pricing algorithm that combines these risk factors with risk scores and translates into insurance prices for each customer. Each firm submits multiple files each year for its rate-making and shows changes in its pricing algorithms. These filings are often quite extensive and lengthy, to verify that firms charge fair and affordable prices.

I use the raw data from filings made by an auto insurance company in California between 2014 and 2021. Scanning hundreds of filings and thousands of pages, I select the text covering pricing rules, risk factors used in the algorithm, and risk scores attached to each factor (See Figures B1 to B4 for examples). Next, I extract text and tables of pricing information and compile them into one comprehensive dataset. The rich data allows me to construct a panel of insurance risk factors and scores used in pricing algorithms. Therefore, this feature of the data enables me to replicate the firm's pricing algorithm by incorporating all possible combinations of risk characteristics and risk scores. Although in many applications of algorithmic pricing, the algorithm itself is not observable to researchers, the setting in this study allows me to analyze the 'black box' of the pricing algorithm and investigate how the algorithm responds to the anti-discriminatory regulation. Therefore, the dataset I compiled provides a rare opportunity to study the consequences of algorithmic regulations in a transparent manner.

Insurance companies offer a range of levels of cover to consumers. For simplicity, this paper focuses on mandatory liability coverage, which consists of bodily injury and physical damage plans (abbreviated as BI and PD, in Appendix B). The compulsory nature of liability insurance mitigates concerns about the selection into different coverages by different customers. If a customer wants to buy additional coverage, this is included in the price as an additional cost.

Table 3 presents the descriptive statistics of risk scores extracted from pricing filings and provides information on risk classes based on the interaction of gender, marital status, and driving experience before and after the policy. Risk scores are raw measures of risk for the firm and can be interpreted as how risky a specific feature is, relative to other features. For instance, before the policy, the average risk score for single males with up to 10 years of driving experience is 2.02, and it is 1.13 for drivers with more than 10 years of experience in the same category. In that sense, inexperienced single young males are seen as almost twice as risky as similar drivers with experience. Although the gap in risk scores is not as striking as that for young men, inexperienced women drivers are also seen as significantly riskier than experienced women drivers.

Another notable dimension in this table is the gender gap in risk scores and how it changes with the experience level. The risk score gender gap among drivers with less than 10 years of experience is approximately 12 percent, in favor of young women. In comparison, the same gender gap for experienced drivers is less than 1 percent. In addition to experience, marital status also plays a role in risk scores, especially for drivers with fewer years of experience. Single drivers are seen as 28 percent riskier than married drivers. On the other hand, the gender gap among experienced single and married drivers is very similar.

Table 4 provides descriptive statistics about risk scores for car models grouped by their segments. Using a classification scheme of the US National Highway Traffic Safety Administration and the US Environmental Protection Agency, I coded 65,808 different car models based on their segments as compact, luxury, midsize, pickup, sport utility vehicle (SUV), and van. Pickup trucks have 10 percent higher risk scores than average among all car segments, whereas compact cars and luxury cars have the lowest risk scores. However, it is essential to note that each car segment contains hundreds of car models; hence this aggregated representation can

Table 3: Summary of risk scores for demographics

Risk scores before the policy	N	Mean	St. Dev.	Min	Max
Single males up to 10 years of experience	55	2.02	0.65	1.17	3.15
Single females up to 10 years of experience	55	1.83	0.5	1.22	2.65
Married males up to 10 years of experience	55	1.48	0.44	0.85	2.5
Married females up to 10 years of experience	55	1.25	0.32	0.89	2.1
Single males with 10+ years of experience	70	1.13	0.22	0.93	1.91
Single females with 10+ years of experience	70	1.22	0.1	1.05	1.45
Married males with 10+ years of experience	70	0.98	0.17	0.79	1.49
Married females with 10+ years of experience	70	0.95	0.09	0.81	1.14
Risk scores after the policy	N	Mean	St. Dev.	Min	Max
Single drivers with up to 10 years of experience	66	1.38	0.26	0.97	1.72
Single drivers with 10+ years of experience	84	1.08	0.16	0.79	1.32
Married drivers with up to 10 years of experience	66	0.98	0.22	0.71	1.29
Married drivers with 10+ years of experience	84	0.85	0.13	0.67	1.05
Full sample	800	1.22	0.44	0.67	3.15

Notes: This table summarizes the main statistics for the pricing algorithms dataset compiled through data extraction from insurance pricing filings of an insurance provider in California. It covers the years between 2014 and 2021. It focuses on risk scores related to gender, marital status, and driving experience (grouped as more and less than ten years of experience) only. It presents risk scores by grouping them before and after the policy (before and after 2019). The number of observations (N) represents the number of categories rather than the number of people in these categories.

miss heterogeneity in risk scores of car models within segments.

Table 4: Summary of risk scores for vehicle segments

Risk Scores by Vehicle Segment	N	Mean	St. Dev.	Min	Max
Compact car	6,728	0.96	0.16	0.54	1.66
Luxury car	7,310	0.97	0.17	0.53	1.66
Midsize car	15,454	1.02	0.15	0.53	1.55
Pickup car	11,998	1.17	0.23	0.55	2.42
SUV (Sport Utility Vehicle)	11,170	1.05	0.14	0.59	1.51
Van	3,022	1.06	0.21	0.63	1.58
Other	19,892	1.04	0.17	0.54	2.4
All vehicle models	65,808	1.05	0.19	0.53	2.42

Notes: This table summarizes the main statistics for the pricing algorithms dataset compiled through data extraction from insurance pricing filings of an insurance provider in California. It covers the years between 2014 and 2021. It focuses on risk scores for the vehicle model by their segment. The number of observations (N) represents the number of car models in each category rather than the number of people having these cars.

The pricing algorithm dataset includes all risk score tables for all risk factors used in pricing. To make it easier to follow for a reader, I only present risk scores related to demographic features and vehicle types. Still, it is possible to extend these to risk scores for ZIP codes, annual mileage, accident history, etc.

5 Effects on consumers: Changes in insurance prices

The first step in an empirical investigation of the impact of a gender-blind policy is to estimate how much male and female consumers pay after the policy, and analyze whether the gender gap closes. The answer depends on the underlying distribution of other risk factors across demographic groups and companies' response to the policy. We can theorize that if firms cannot charge consumers based on gender (or its proxies) and the policy becomes fully effective, the gap should be minimal. On the other hand, if firms compensate for their information loss through different channels such as using proxies of the banned variable or collecting new information, then we can expect at least some part of the gender gap to remain, depending on firms' responsiveness or ability to recover their information set.

In this section I analyze whether the policy fully eliminates the gender gap or not, and how much the premiums of different demographic groups change. I will start by estimating the differential impact of the policy on the gender gap among different age and gender groups within California. Then, I will focus on the exogenous variation in the policy implementation across states and compare the young people in California to drivers in other states with gender-based pricing.

5.1 Age, gender and insurance prices

Gender is used as a rating factor in car insurance in many states where it is not prohibited by law. Insurers base their ratings on how risky male or female drivers are on average. These risk perceptions are based on previous customers' claims and accident statistics in the same risk group. However, one crucial aspect of gender-based pricing is its interaction with age or driving experience. For younger drivers with only a few years of driving experience, the gender gap is more apparent, and young males are associated with higher risk. Hence, they usually pay higher premiums than young female drivers. However, as young people gain more driving experience, firms base their insurance prices on individual driving and accident histories rather than gender, which is a noisier risk signal.

Figure 1 shows the risk scores based on gender and experience, as used in the firm's pricing algorithm. Risk scores can be seen as raw measures of a firm's risk perception and the risk of a specific risk group (for instance, people with x years of driving experience is a risk group) compared to the risk of other risk groups. They directly enter the firm's pricing algorithm and linearly contribute to the insurance price. In that sense, risk scores can be considered inputs into the pricing algorithm. Drivers with less than 5 years of experience are considered almost twice as risky as more experienced drivers. Likewise, the gender gap is more salient for less experienced drivers, especially in the first five years, but it significantly decreases as drivers gain more experience.[19] We should also note that risk scores for driving experience and gender are only one part of the price calculation. For the final insurance price, many other dimensions such as other individual factors (driving histories, etc.) and location and vehicle characteristics are included in the pricing algorithm.

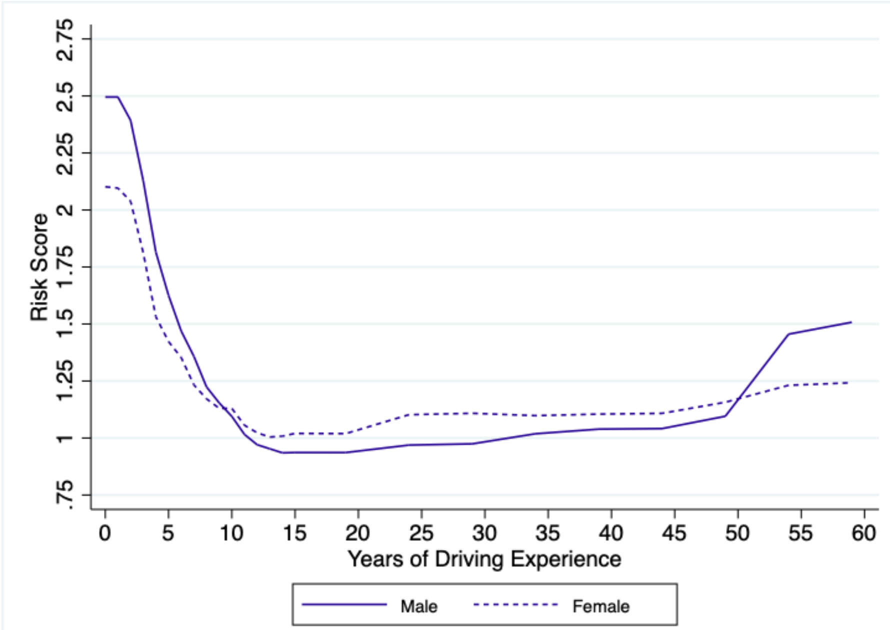
Another way of looking at the interaction between gender and experience is focusing on the premiums that different demographic groups pay. Figure 2 provides information on annual insurance premiums paid by male and female drivers of different ages.[20] Young drivers under 25 are charged more than twice what older drivers pay on average. When drivers reach age 25, insurance premiums significantly decrease, and the middle-aged group, between 50 and 70 years old, pays the lowest premiums. In the elder group, premiums show a slight increase. Regarding the differential in insurance premiums between men and women, the raw gap among drivers under 25 is about 20 percent, but over the age of 25, men and women pay very similar amounts for car insurance.[21] The similar patterns in Figure 1 and Figure 2 show evidence in favor of

[19] The gender gap in risk scores appears again for drivers with more than 50 years of experience. However, as there is a relatively small number of drivers in this category compared to other groups, this study will focus on two main groups (people with less and more than 10 years of driving experience) in terms of the gender gap.

[20] Throughout this work, I use the terms young (old) and inexperienced (experienced) interchangeably by thinking of age as a proxy for years of driving experience. There is various empirical evidence supporting this claim. According to Federal Highway Administration's report (2018), around 61 percent of US drivers obtain a driving license by age 18, 84 percent by age 25, and 94 percent by age 35. Besides, significant similarities in Figure 1 and Figure 2 suggest that young and inexperienced drivers have almost identical patterns.

[21] As opposed to the gender gap in risk scores of people with more than 50 years of driving experience, there

Figure 1: Risk scores by age and gender



Notes: This figure shows the relationship between risk scores and driving experience by gender. It uses the pricing algorithm dataset from an insurance provider in California. It covers the years between 2014 and 2018 (the period before the policy).

the insurance premiums dataset being an accurate reflection of the correlation between being a member of a group with higher risk (i.e., young) and paying higher premiums.

Next, I estimate the following equation to analyze the different costs faced by young males compared to other demographic groups:

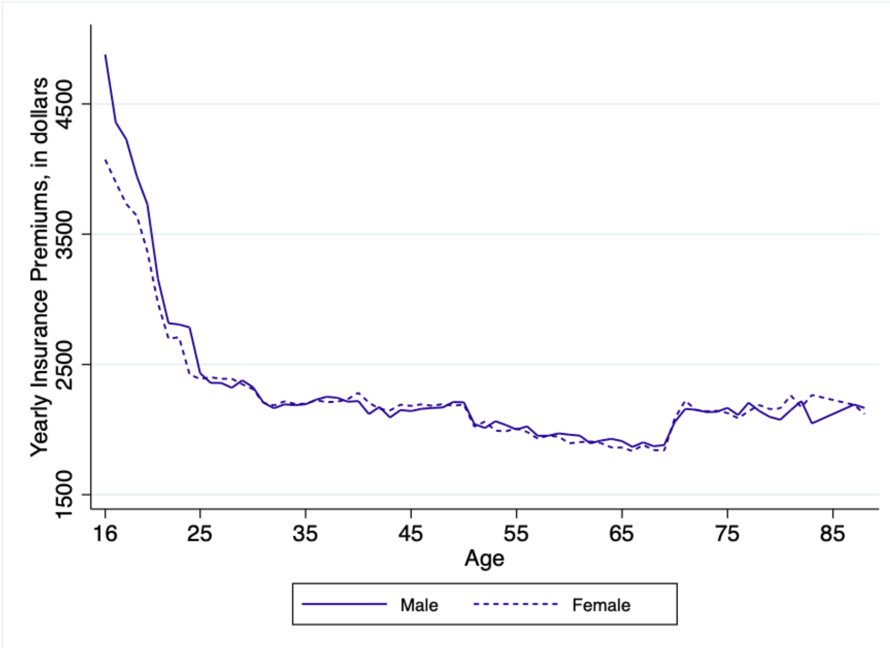
$$\text{Log}(\text{Price}_{ist}) = \beta_0(\text{Young Male in CA}_{is}) + \text{X}_{ist} + \gamma_s + \alpha_t + \epsilon_{it} \qquad (2)$$

where i is the individual i and t is the year. $\text{Price}_{\text{ist}}$ represents the insurance premium paid by individual i at state s and at time t. X_{ist} is a vector of controls, including individual, location, and vehicle characteristics. ($\text{Young Male in CA}_{\text{is}}$) takes value 1 when individual i is a young male in California, 0 otherwise. γ_{s} and α_{t} are state and year fixed effects, respectively.

Table 5 reports estimation results for Equation (2). It supports the two main observations we infer from Figure 1 and Figure 2. First, a significant gender gap exists among young people. Column (1) indicates that young males in the treated state (California) paid 19 percent more than young females before the policy was effective. Secondly, young males pay significantly more than everyone else. Columns (2) to (5) focus on different subsamples and suggest a significant gap between young male drivers and other demographic groups in California and non-treated

is no gender gap in terms of premiums for this group. It provides further evidence in support of focusing on only two groups when it comes to the gender gap (below and above 25 years old).

Figure 2: Insurance premiums paid by age and gender



Notes: This figure shows the relationship between age and annual insurance premiums for male and female drivers. It uses the US Consumer Expenditure Survey data for states with gender-based pricing (excluding Massachusetts, Pennsylvania, North Carolina, Hawaii, and Montana) between 2010 and 2020 and California between 2010 and 2018 (before the policy).

states. In Columns (1) and (2), differences between young males and other groups in California can be attributed to young males being riskier than others on average. However, in Columns (3) and (4), young males in California are paying more than their counterparts in other states can be attributed to pre-existing differences across states. Section 5.3 provides a more detailed discussion of these differences.

5.2 Variation across age groups

5.2.1 Empirical strategy and identification

The empirical evidence in the previous section shows that the insurance gender gap only exists for young drivers. Therefore, we expect the gender-neutral pricing policy to only directly affect young people. It should not impact the gender gap among older drivers, as gender is not considered a significant risk factor after certain years of driving experience.[22] The main idea behind the identification is to compare the evolution of the gender gap among young individuals in California with the gender gap among older people before and after the policy. The identifying assumption is that the gender gap among older people is a good counterfactual for how the gender gap among the young would have evolved in the absence of the new rule.

[22] There can be an indirect effect on older people if firms start to use characteristics that are highly correlated with gender as a proxy, and I will discuss this in Section 6.

Table 5: Premiums paid by young males in California before the policy

Dependent variable: log(price)	(1) Young (California)	(2) All people (California)	(3) Young (All states)	(4) Young males (All states)	(5) Full sample
Young Male	0.19***	0.33***			
	(0.02)	(0.02)			
Young Male * California			0.08***	0.13*	0.32***
			(0.02)	(0.07)	(0.00)
Demographic characteristics	✓	✓	✓	✓	✓
Location characteristics	✓	✓	✓	✓	✓
Vehicle characteristics	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
State FE			✓	✓	✓
State x Year FE			✓	✓	✓
Observations	646	13,061	4,375	2,466	92,431
Adjusted R-squared	0.42	0.09	0.35	0.42	0.07

Notes: The logarithm of insurance price is regressed on the young male dummy and set of controls. Demographic controls are age, marital status, and education level. Location characteristics are geographic region, population density, and metropolitan statistical area dummy. Vehicle characteristics are vehicle type, vehicle year, vehicle brand, and used status. The sample used in these regressions covers the years between 2010 and 2018 (pre-policy period) and states that have gender-based pricing policies. Heteroskedasticity-robust standard errors are clustered at the state level in columns (3)-(5). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

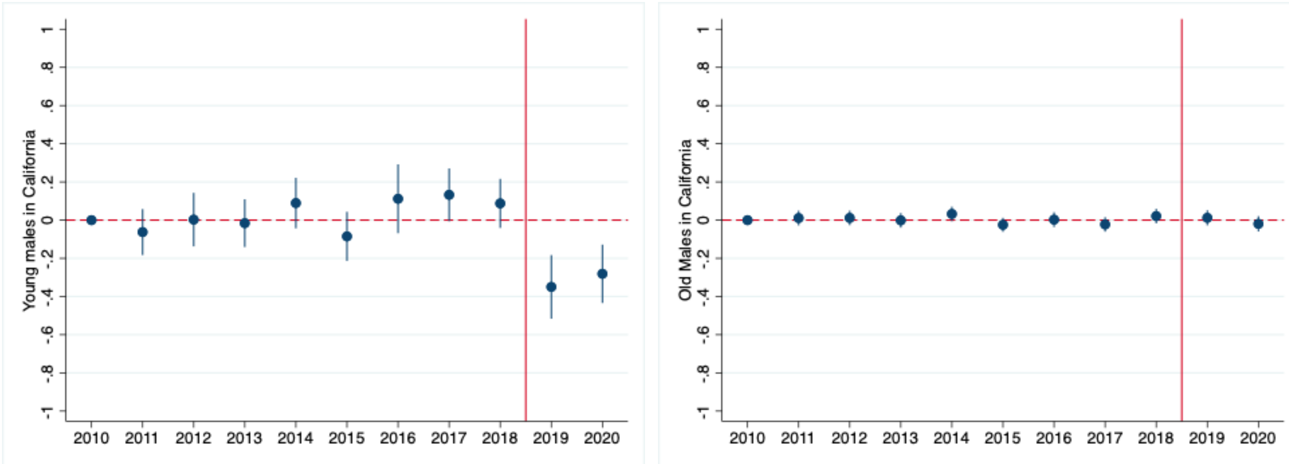
Figure 3 provides event study estimates for young and older people separately. The panel on the left compares only young people and shows that young males started to pay significantly less than young females after the policy. The right panel plots the change in premiums for older males compared to older females, and there is no significant gap between these two groups before or after the policy. Additionally, Figure C1 is an annual breakdown of the insurance premiums paid by each group. This figure also provides essential evidence for parallel trends across younger and older groups before the policy was effective. Figure C2 and Figure C3 show further evidence for time trends for each group.

Following a similar approach to Gruber (1994), I estimate the following triple difference-in-differences model to estimate the differential impact of policy on the young gender gap:

$$\begin{aligned}\text{Log}(\text{Price}_{\text{it}}) = \beta_0(\text{Post}_t \times \text{Young}_i \times \text{Male}_i) + \beta_1(\text{Post}_t \times \text{Young}_i) + \beta_2(\text{Young}_i \times \text{Male}_i) + \\ \beta_3(\text{Post}_t \times \text{Male}_i) + \beta_4\text{Post}_t + \beta_5\text{Young}_i + \beta_6\text{Male}_i + \text{X}_{it} + \alpha_t + \epsilon_{it} \quad (3)\end{aligned}$$

where for each individual i and year t, $\log(\text{Price}_{\text{it}})$ represents the logarithm of insurance premium price paid by individual i at year t. Young_{i} and Male_{i} represent whether individual i is under 25 years and male, respectively. Policy_{t} takes value 1 if the policy is implemented at time t, 0 otherwise. X_{it} is a vector of individual-level control variables such as demographic, location,

Figure 3: Premiums for young and older males in California



Notes: The left panel focuses on the sample of young people under 25, and the right panel is on people above 25. This figure presents the estimates for male coefficients in a regression where the logarithm of price is regressed on a male dummy and demographic, location, vehicle characteristics, and year fixed effects. For a complete list of variables, see Table 6. It uses data from consumer expenditure data between 2010 and 2020. 90% confidence intervals are plotted. Heteroskedasticity-robust standard errors are used for each regression.

and vehicle characteristics for individual i at time t.[23] Year fixed effects denoted by α_t are also included in each regression. By estimating Equation (3), the coefficient of interest β_0 reflects the change in gender gap among young people with respect to the gender gap among older people in the treated state before and after the policy. β_5 and β_6 are treatment group specifics to account for average time-invariant differences for being young and male, respectively. β_4 captures common time trends for treatment and control groups.

The theory discussed in Section 3 suggests a differential impact on young people where gender is considered informative about risk. Specifically, after the policy, companies lose one observable (gender) that provides information about risk for young people. The theory suggests that since companies cannot use gender in pricing anymore, we expect the gender gap in prices to decrease significantly. On the other hand, if companies make some adjustments in the pricing algorithm to recover for missing information (i.e., using proxies), we expect some gender gap to remain and some spillover effects on different demographic groups through proxies.

5.2.2 Results: Variation across age groups

As a part of the theoretical framework in Section 3, removing gender information is expected to decrease the gender gap among young people as companies cannot perfectly differentiate between two genders. To empirically test this prediction, I use the variation in the effectiveness of the policy across age groups. Specifically, I use a triple difference estimator by focusing on the gender gap among the young with respect to the old before and after the policy.

Table 6 presents estimation results for Equation (3) and the gender gap among young and

[23]For a detailed list of variables, please refer to Table 6.

old separately with and without controls. Columns (1) and (2) report estimation results for young people. The results imply that the unconditional gender gap among the young decreased by 13 percent after the policy. When we control different demographic, geographic, and vehicle characteristics, the conditional gap becomes 21 percent. Table C1 presents further evidence of how each control variable contributed to the reduction in the gap. By comparison, Columns (3) and (4) indicate that the policy did not significantly impact the gender gap among older people. This finding aligns with the evidence presented in Figure 3. Finally, Column (5) reports the triple DID estimates from Equation (3). The gender gap among young drivers has decreased by 13 percent compared to the gender gap among older adults in California after the policy. This impact corresponds to an approximately $236 decrease in the relative gender gap in insurance costs. The triple difference method reflects the relative changes in the gender price gap among young and old drivers by differencing two difference-in-differences estimators. The effect sizes are consistent with the sample sizes for young and old samples.

Table 6: The impact of the gender-neutral pricing policy on drivers in California

Dependent variable: log(price)	(1) Young (California)	(2) Young (California)	(3) Old (California)	(4) Old (California)	(5) California (Full sample)
Post * Male	-0.136**	-0.212***	0.001	0.001	
	(0.063)	(0.051)	(0.015)	(0.015)	
Post * Young * Male					-0.133***
					(0.049)
Demographic characteristics		✓		✓	✓
Location characteristics		✓		✓	✓
Vehicle characteristics		✓		✓	✓
Year FE	✓	✓	✓	✓	✓
Observations	730	730	14,779	14,779	15,509
Adjusted R-squared	0.111	0.473	0.084	0.084	0.129
Empirical Strategy	DID	DID	DID	DID	Triple DID

The logarithm of insurance price is regressed on a male dummy and set of controls in Columns (1) to (4). Column (5) reports estimation results from equation 1. Demographic controls are age, marital status, and education level. Location characteristics are geographic region, population density, and metropolitan statistical area dummy. Vehicle characteristics are vehicle type, year, brand, and used status. The sample used in these regressions covers the years between 2010 and 2020, only for California. The heteroskedasticity-robust standard errors are reported in each regression. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

These results provide significant empirical support for the claim that it is young males who are mainly affected by this policy. They also support the intuition that the firms need to differentiate young males from young females after the policy, as gender does give information about risk for these groups. Depending on the firm's response (whether they will react or use proxies for 'young and male'), we theoretically expect young males to pay less as companies may not predict this feature perfectly. On the other hand, the impact on young females is not easy to predict as it will mainly depend on the underlying distribution of features among this group. As firms can still charge people according to age or experience, young females would be worse off because they are young, but firms may not have a further incentive to adjust this group's pricing. In that sense, the changes in the young females' premiums can depend on the firm's use of proxies for young males and the distribution of young male features among young females. Further discussion will be provided in Section 6 where I analyze the impact of the policy on the firm's pricing. Also, additional empirical evidence on how young males and females are separately affected by the policy is provided by event study estimates in Figure C2.

The empirical analysis in part focuses on different age groups and how each group is affected relative to the other. Using the fact that after a certain level of experience, gender is no longer used as a risk factor suggests that older (experienced) groups will not be affected by the removal of gender as a variable in pricing. Findings in Figure 3 and Table 6 support this and provide further evidence that the gender gap among young people decreased, compared to those aged over 25. One caveat for interpreting these findings is considering changes in the pricing of other factors and spillover effects for different demographic groups through these factors. Specifically, after the policy, companies can use other features to predict risk, which can be correlated with gender. Although the analysis in this part mainly groups drivers in terms of age and driving experience, there can be different effects within these groups. For instance, based on this finding, we expect older men not to be affected by the policy on average. Still, older men with specific features (specific car models associated with young males, for instance) may be negatively affected. Section 6 discusses this type of effect within groups.

Furthermore, a gender-neutral pricing policy requires men and women to pay the same insurance price conditional on the same observables. This analysis points out that a conditional gender gap remains for the younger group. There can be different explanations for this remaining gap, such as the underlying distribution of observables across different groups, the granularity of the dataset used in this analysis, or the variation across collected information by various insurance companies. The granularity in the consumer expenditure survey has some natural limitations. In that sense, the data provided in the surveys may not necessarily reflect the same set of information as insurance firms. For instance, insurance companies can ask customers for more specific information such as driving histories, location, and detailed car features. However, as the treatment group (young drivers) does not have accident histories anyway, some of the missing information has a lesser impact on the analysis. In terms of detailed car features and location, the more detailed analysis using extensive data on driver characteristics is provided in Section 6.

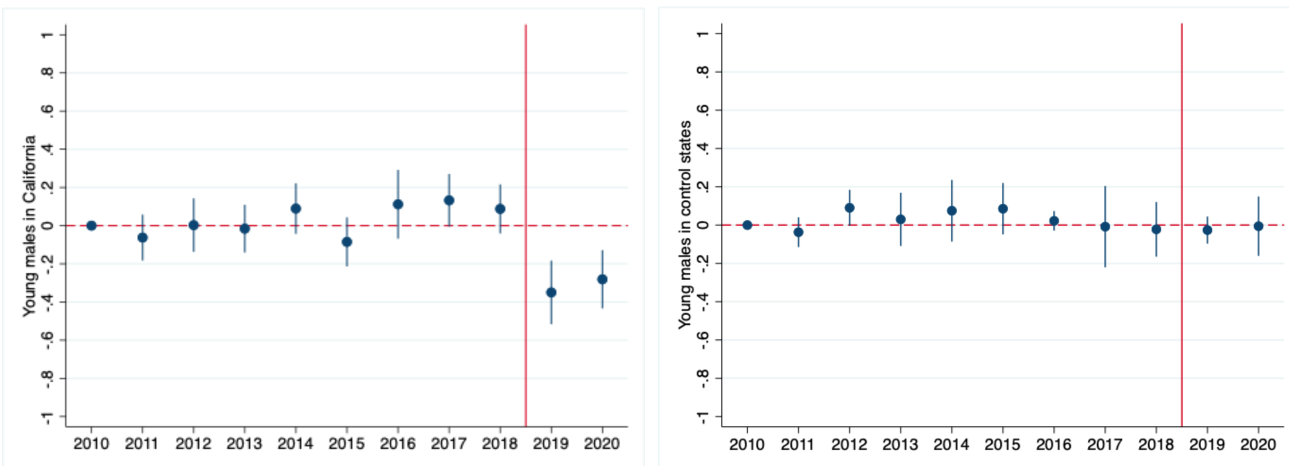
Using the variation across age groups has some advantages. As both treatment and control groups have the same contemporaneous shocks within the same state, using difference-in-differences estimator differences out these shocks. However, an alternative way of looking at the effect of this policy is to look at variation across states. In the following section, I focus on states that do not have a gender-neutral policy and compare them to California.

5.3 Variation across states

5.3.1 Empirical strategy and identification

An alternative way to identify the effect of gender-blind pricing on young drivers is to focus on the variation across states. Specifically, I compare young individuals in the treated state to young people in states with gender-based insurance pricing. The identification assumption requires that changes in the young driver gender gap in the treated state would have been similar to the gender gap across young drivers in control states, in the absence of a gender-blind policy.

Figure 4: Premiums for young males in California and control states



Notes: The left and right panels focus on the sample of young people under 25 in California and control states, respectively. The figure presents the event study estimates for young males compared to young females. The logarithm of insurance premiums is regressed on male dummies, demographic, location, vehicle, and characteristics. Fixed effects for the year, state, and state*year are included in each regression. The coefficients for the male and year interaction term are plotted for young males in California and control states by taking 2010 as the base year. Data comes from consumer expenditure data between 2010 and 2020. 90% confidence intervals are plotted. Heteroskedasticity-robust standard errors are clustered at the state level for the figure on the right.

Figure 4 is a plot of event study estimates using Equation (4) and it compares young males relative to young females in California and control states, respectively.

$$\text{Log}(\text{Price}_{\text{ist}}) = \beta_0 \text{Male}_{is} + \text{X}_{ist} + \gamma_s + \alpha_t + \epsilon_{ist} \tag{4}$$

where i is individual, s is state, and t is year. $\text{Price}_{\text{ist}}$ denotes insurance premiums for individual i at state s and at time t. X_{ist} is a set of control variables, including individual, location, and vehicle characteristics. Male_{is} takes a value of 1 if individual i at state s is male, 0 otherwise. γ_{s} and α_{t} are state and year fixed effects, respectively.

Figure 4 also provides evidence for parallel time trends between treatment and control states. Further evidence for raw trends of premiums paid by young drivers in California and control states is provided in Figure C5. I estimate the following triple difference-in-differences model:

$$\text{Log}(\text{Price}_{\text{ist}}) = \beta_0(\text{Post}_t \times \text{Treated}_s \times \text{Male}_i) + \beta_1(\text{Post}_t \times \text{Treated}_s) + \beta_2(\text{Treated}_s \times \text{Male}_i) + \beta_3(\text{Post}_t \times \text{Male}_i) + \beta_4\text{Post}_t + \beta_5\text{Treated}_s + \beta_6\text{Male}_i + \text{X}_{ist} + \gamma_s + \alpha_t + \epsilon_{ist} \quad (5)$$

where i is the individual, s is state, and t is year. Price_{ist} represents the insurance premium paid by individual i at state s and at time t. X_{ist} is a vector of controls, including individual, location, and vehicle characteristics. Post_{t} takes value 1 after 2018 when the policy was implemented, 0 otherwise. $\text{Treated}_{\text{s}}$ and Male_{i} are dummy variables for being in the treated state and being male, respectively. The coefficient of interest β_0 for the triple interaction term reflects the relative change in the gender gap among young drivers in California with respect to the gender gap among other groups in control states before and after the policy.

To identify the causal impact of the policy, I must account for any systemic shocks in different states that may be correlated to the law. To control for these shocks, I include year fixed effects (denoted by α_t) that capture time trends, state fixed effects γ_s to account for differences in California and control states that do not pass the law, and state-by-year fixed effects to control for state-specific shocks over time. Thus, I compare the treatment individuals (young drivers in California) to drivers in control states and estimate the relative change in premiums of treatment individuals compared to drivers in control states.

Similar to the predictions provided earlier in Section 5.1, the theoretical framework suggests a relative decrease in the gender gap between treated young individuals in California, compared to those in control states. Depending on the control group, one can make different predictions. For instance, after controlling for state-specific shocks, we expect the young gender gap in California to decrease compared to the young gender gap in control states. In line with the earlier discussion, a decrease in premiums for treated young males in California should mainly drive the decline in the gender gap.

5.3.2 Results: Variation across states

Table 7 presents different estimation results for triple DID and DID estimations.[24] Columns (1), (2), and (3) report estimation results for difference-in-differences estimations, and Column (4) presents the triple difference results. Specifically, Column (1) shows that the policy does not impact the unconditional gender gap among young drivers in control states. The null effect is persistent when we control for demographic, geographic, and vehicle characteristics. These results are consistent with the prediction that control groups are not affected by the policy.

By contrast, as shown in Column (3), the gender gap for young drivers in California signif-

[24] The sample of control states in Table 7 is restricted to the states with at least 200 observations for the young population. But the estimation results are robust to non-restricted versions as well. Table C2 lists empirical results without this sample restriction.

Table 7: The impact of the gender-neutral pricing policy on young drivers across states

Dependent variable: log(price)	(1) Young (Control states)	(2) Young (Control states)	(3) Young (California)	(4) Young (All states)
Post * Male	-0.042	-0.047	-0.212***	
	(0.061)	(0.061)	(0.051)	
Post * Male * California				-0.178**
				(0.057)
Demographic characteristics		✓	✓	✓
Location characteristics		✓	✓	✓
Vehicle characteristics		✓	✓	✓
Year FE	✓	✓	✓	✓
State FE	✓	✓		✓
State x Year FE	✓	✓		✓
Observations	1,438	1,438	730	2,168
Adjusted R-squared	0.035	0.336	0.473	0.373
Empirical strategy	DID	DID	DID	Triple DID

The logarithm of insurance price is regressed on a male dummy and set of controls in Columns (1) to (3). Column (4) reports the coefficient for the triple interaction term in equation 3. Demographic controls are age, marital status, and education level. Location characteristics are geographic region, population density, and metropolitan statistical area dummy. Vehicle characteristics are vehicle type, year, brand, and used status. The sample used in these regressions covers the years between 2010 and 2020, only young people. Control states are all states that have gender-based insurance pricing and have at least 200 observations in the sample. There are 36 control states included. Fixed effects for the year, state, and state*year are included in each regression. The heteroskedasticity-robust standard errors are clustered at the state level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

icantly decreased after the policy (as discussed in Section 5.2). In line with these results, the triple difference estimator in Column (4) indicates a 17 percentage points relative decrease in the young gender gap in California compared to the young gender gap in control states.

Within California estimates from Table 6 also suggest a significant relative decrease in the gender gap for young drivers, falling 13 percentage points compared to the gender gap for older drivers. These results support the robustness of the main effect across different control groups and empirical strategies. We obtain different effect sizes as the control groups differ in two empirical strategies (Section 5.2.2 and Section 5.3.2). The first empirical strategy compares the change in the young gender gap in California (21 percent) with the older gender gap (0.001 percent and insignificant). The second empirical strategy compares it with the change in the young gender gap in states with gender-based pricing (4 percent and not significant). Differences in sample sizes and standard errors across these two control groups drive different effect sizes in Table 6 (Column 5) and Table 7 (Column 4).

6 Opening the black box of the pricing algorithm: Effects on pricing

In this section, I focus on the impact of gender-neutral pricing on the firm's pricing algorithm. I analyze whether the features that predict being a member of the riskiest demographic group, i.e., young males, are repriced and given more weight in price calculations. As a first step, I estimate the features that predict being a young male by using a rich dataset that includes individual, vehicle, and location characteristics. I use different estimation techniques, including logistic regressions, and supervised machine learning methods such as regularized regressions, to select features that predict gender. Next, I analyze the impact of policy on pricing algorithms and determine whether variables that predict gender are repriced in the algorithm. Using the pricing algorithm dataset, I estimate changes in risk scores of gender-predicting features (such as different car models). Lastly, I focus on the consumer side and investigate what drivers with features that are highly correlated with gender start to pay after the policy change.

6.1 Predicting gender with different statistical methods

Insurance companies collect extensive information to calculate the prices they quote to customers. This information often includes individual characteristics such as age, gender, marital status, education, occupation, location-based characteristics such as ZIP code, and vehicle characteristics such as car brand, model, year, etc.[25] After the policy change, Californian insurance providers cannot use gender directly to determine prices, and so rely on proxies for gender instead.

In this section, using an extensive household travel survey that matches the individual and car characteristics of consumers, I will estimate features that predict gender. This dataset includes more than half a million individuals and 682 variables, including a large set of individual and vehicle characteristics (Listed in Table C3). I focus on the predictors for young males given other features, using information from all drivers since the gender gap is not significant for middle-aged and elderly groups.[26]

[25]Different states have different regulations about which variables can be used in determining risk and pricing. Further information can be found on each state's insurance department's websites.

[26]Young males are the riskiest group for car insurance, and firms may need to identify them after the policy change. After the gender-neutral pricing regulation, firms can still ask about age or driving experience; therefore, they can infer if a driver is young or inexperienced. In that sense, the firm's problem can be considered as differentiating young males from young females. This is the primary reason for focusing on young male characteristics in this part of the analysis. Section 5.1 and Section 5.2 provide evidence that supports the proposition that young males are the primary treatment group.

6.1.1 Logit Model

Logit models are commonly used in discrete choice, forecasting, and prediction applications (Demyanyk and Van Hemert, 2011, Elul et al., 2010). They are considered an alternative to ordinary linear regressions when the outcome variable is discrete. In these models, the logit link function connects the outcome variable (probability values between 0 and 1) to the linear combination of parameter estimates for each predictor in the model. To estimate the probability of being a young male given other features, I use the following logit model:

$$\log(\frac{\theta_{\mathrm{i}}}{1-\theta_{\mathrm{i}}}) = \beta_0 + \beta_1 \mathrm{x}_{i1} + \beta_2 \mathrm{x}_{i2} + ... + \beta_k \mathrm{x}_{ik} \tag{6}$$

where θ_{i} is the probability of being a young male for individual i, and each x_{i} represents covariates in the model. The covariates include individual characteristics (age, education, occupation), household characteristics (homeownership, number of vehicles/drivers in household), location-based characteristics (metropolitan statistical area size, population, and housing density in living area), car characteristics (vehicle brand, vehicle model, vehicle age, fuel type, vehicle type, annual miles, odometer reading).

I estimate two different versions of these models, which differ in terms of whether some of the covariates are used in a linear or nonlinear form to predict being a young male. In the first model, control variables that may also take categorical forms such as vehicle age, annual miles, odometer reading, and individual's age are used in the linear form together with various dummy variables. In the second model (nonlinear logit), I relax the linearity assumption and add them as categorical variables to capture possible nonlinear relationships. It allows us to include potential nonlinear relationships between the dependent variable and some covariates. The remaining variables are the same in both models. Table C3 presents a complete list of variables included in prediction analysis.

6.1.2 Supervised Machine Learning Methods: Regularized Regressions

Regularized regressions are machine learning algorithms where an objective function aims to balance out-of-sample predictions with in-sample using a penalty term. This penalty term aims to prevent overfitting and drops many covariates from the regressions by assigning a coefficient of zero.[27] There is a large literature in economics, business, and computer science that uses regularization methods for different problems, such as prediction in policy analysis (Chandler et al., 2011, Chalfin et al., 2016; Kleinberg et al., 2015; Peysakhovich and Eckles, 2018; Kleinberg et al., 2018), causal inference (Angrist and Frandsen, 2022; Andini et al., 2018), and model selection (Belloni et al., 2014a; Belloni et al., 2014b).[28]

[27]Ridge estimator whose interpretation is slightly different from lasso estimator leads to many nonzero coefficients. See more detailed discussion in Hastie et al. (2017).

[28]For a more general discussion about machine learning methods in economics, see Varian (2014); Kleinberg et al. (2015); Gentzkow et al. (2019); Mullainathan and Spiess (2017); Athey (2017); Athey and Imbens (2019);

I adopt simple machine-learning methods to capture the subsets of variables that best predict being a young male. Traditional stepwise selection methods may have issues with prediction accuracy in high-dimensional data settings. Compared to these methods, regularized regressions improve prediction accuracy by shrinking and penalizing the size of coefficients, especially when there are many predictor variables. Another dimension where regularized regressions can be advantageous is their performance for out-of-sample forecasts. Using a penalty term, they select predictors with cross-validation and are inclined to provide better out-of-sample predictions (James et al., 2013). In this part, I use different regularization methods such as Least Absolute Shrinkage and Selection Operator (Lasso) (Tibshirani, 2011), ridge regression (Hoerl and Kennard, 1970), and elastic net (Zou and Hastie, 2005; Zou, 2006; Zou and Zhang, 2009).

In a generalized form, we can consider a linear regression model to predict $\mathrm{y_k}$ with a set of covariates $(\mathrm{x_1, x_2, ...x_k})$ and a parameter vector $(\mathrm{b_1, b_2, ...b_k})$. Regularized regression methods additionally have a penalty term. Minimizing the sum of residuals and a penalty term can be denoted as follows:

$$\lambda \sum_{\mathrm{k}=1}^{\mathrm{K}} (1-\alpha)|\beta_{\mathrm{k}}| + \alpha|\beta_{\mathrm{k}}|^2 \tag{7}$$

where β_{k} represents coefficients for predictor value $\mathrm{x_k}$ where $\mathrm{k} = 1, 2, ..., \mathrm{K}$. The general form of Equation (7) gives elastic net regression. Depending on the values of λ and α, it creates different regularized prediction methods as special cases. In the simplest case, if there is no penalty term ($\lambda = 0$), this is the ordinary least squares method. If there is a penalty ($\lambda > 0$) but $\alpha = 0$, this will give the functional form for lasso regression. If there is an additional penalty with a quadratic term ($\alpha = 1$), it is ridge regression. These models tend to penalize the complexity of the models to overcome overfitting problems, and their estimators are considered computationally efficient, especially with large samples (Varian, 2014). Adaptive lasso is a modified version of lasso where weights are also applied to each parameter β_{k} in the penalty term (Zou, 2006).

6.1.3 Results: Predicting gender

I estimate covariates predicting being young male using different statistical methods discussed above. I included an extensive set of variables, including individual, demographic, household, and vehicle characteristics. Table C4 reports the results of these prediction models. Columns (1) to (4) show covariates selected by regularized regressions: lasso, elastic net, ridge, and adaptive lasso, respectively. Column (5) and (6) presents logistic regression results without any

Athey (2019).

penalty term.[29]

After the policy, firms can still include age (or experience) in their pricing; therefore, firms' main problem is differentiating young males from young females, but not from older males, for instance. One of the crucial aspects of interpreting results in Table C4 is that age is also included as a predictor in this analysis. In that sense, the selected features represent distinctively young male characteristics conditional on being young.

Table C4 gives various insights into the relationship between being a young male and the types of characteristics. The car characteristics at different levels such as vehicle category (car, pickup, SUVs, etc.), vehicle brand, and vehicle model are selected as predictors. For instance, young males are more likely to use cars, pickups, or trucks (compared to SUVs or vans). Additionally, they are associated with some car brands, such as Ford, Buick, Lexus, and Victory. However, considering that these are quite heterogeneous categories, including tens of different car models, focusing on car models can give better insights.

Four hundred and thirty-eight vehicle models are included in the analysis, and 45 of them are selected as associated with young males. They constitute 10.2% of all car models in the dataset and 16.3% of all drivers in the sample using these car models. They vary in their segments, types, and costs. For instance, young males prefer to use old car models such as Cadillac Seville, Chevrolet Chevelle, Chevrolet Nova, Pontiac Bonneville, and some luxury two-seater sports cars such as Porsche Boxster and BMW Z3. Additionally, they notably prefer pickup trucks such as GMC Canyon and Dodge Ram. These findings are consistent with the empirical evidence in the literature (Chandra et al., 2017).

As discussed in Section 2, car insurance companies also collect information on detailed individual and location characteristics. Hence, demographic factors are also included in the predictive analysis. These characteristics are expected to be selected if noticeably different among young males and others. After having various features regarding individual characteristics (age, education level, occupation, homeownership at household level) and location-specific features, only a few were selected by the models. For instance, being a middle school and high school graduate (compared to college or higher-level degrees) is more common among young males under 25 years old. Also, being a household member with homeownership and living in a metropolitan statistical area are predicted as other features of this demographic group. However, although these demographic features are selected, their coefficients are significantly smaller compared to those of selected car models (with the exception of being middle school

[29] Table C4 reports only positive coefficients for a few reasons. First, negative coefficients represent a potential negative correlation between being a young male and a given feature. However, this case usually is less informative than getting a positive coefficient for the firm. For instance, if a car model is predicted with a negative coefficient, it is more likely to be used by any other demographic group but not young males. Considering the firm's problem of predicting risky (young male) features, the negative coefficients will not be as informative. Second, this shorter representation is possibly easier to follow for the reader. Also, due to the same reason, Table C4 lists only the predictors selected by at least two different methods. Only positive and statistically significant ($p<0.05$) covariates are listed for logit models.

graduates). Therefore, the main focus will be on tracking car models as a young male feature in the pricing algorithm.

One caveat about this analysis is the firms' reasoning for this kind of predictive analysis. The main intuition behind this exercise is to predict the riskiest group's features, and there can be two important reasons supporting the exercise. The first and natural one is that as firms cannot observe gender after the regulation, they can use proxies for this feature, and they do so, particularly for the riskiest group, young males. The second reason is that firms now can try to recalculate the risk prediction problem under this new constraint (information set without gender) and predict risky features. Essentially, the results of both cases are not mutually exclusive, and in both, firms need to infer the risk through a prediction. One might argue that under the new regulation, firms can try to predict risky behavior through 'better' risk predictors such as accident histories or claim histories instead of predicting young male features. However, as young people do not have these features based on driving or claim records, it is not always possible to do this. In that sense, predicting young male characteristics is another way of predicting risk under a new information set without gender.

6.2 Changes in pricing algorithm: Repricing gender predicting features

In this section, I analyze how firms change their behavior in response to gender-neutral pricing policies. This behavior can reveal itself in two ways. The first is adjustments in the pricing algorithm in terms of risk factors estimated as correlated with gender. The latter is how much people with these characteristics start to pay after the policy change. To analyze the changes in the risk scores, I use the pricing algorithm dataset that includes risk factors used in pricing, risk scores attached to each risk factor, and the pricing rule. In that sense, it enables me to analyze changes in the raw risk scores for different attributes such as vehicle characteristics, which directly affect the insurance price by entering into the pricing algorithm.

6.2.1 Changes in insurance pricing of car models

In this section, I focus on the changes in the pricing algorithm in terms of car-related characteristics. Specifically, using the predictive analysis results obtained in Section 6.1, I analyze the impact of policy on risk scores for car models associated with young male drivers. First, I create a variable for young male cars that incorporates car models presented in Table C4.[30]

[30]The prediction analysis uses 438 different car models and selects 45 models as a predictor for being young males by using the consumer characteristics NHTS dataset. Analyzing changes in risk scores for these selected car models requires matching the car model names in the pricing algorithm dataset. In this step, 7 car model names (Chevrolet Geo Metro, Chevrolet Nova, Saturn SC, Volvo 240 series/DL/GL/GLT, Acura Integra, Acura Legend, Lexus NX) did not match across two datasets. Due to similarities, three car models (Nissan Z-Car, Nissan Maxima, and Nissan Datsun 350Z) are named together under the Nissan Z-car name. Therefore, the young male car variable used in this part onward includes the remaining 35 distinct car models.

These car models can be interpreted as models used distinctively by young males. Besides, to create a benchmark group, I use another variable for gender-neutral cars based on the share of young males and young females using these cars. The latter group includes car models with very similar shares among young male and female drivers.[31] In this analysis, the gender-neutral cars can be regarded as a control group where we do not expect any significant change after the policy.

I estimate the following equation to understand changes in risk scores for these car categories compared to other cars.

$$\log(\text{Risk Score})_{mbt} \;=\; \beta_0(\text{Year}_t \times \text{Car type}_m) + \beta_1(\text{Car type}_m) + \gamma_m + \eta_b + \theta_t + \epsilon_{mbt} \quad (8)$$

where $\log(\text{Risk Score})_{mbt}$ represents the logarithm of a risk score for each car model m with brand b at year t. Car type_m takes value 1 for young male car models and for gender-neutral car models for each estimation, respectively.. γ_{m}, η_{b} and θ_{t} are fixed effects for car models, car brands, and year, respectively.

Figure 5 shows event study estimates on car risk scores for young male and gender-neutral cars. Risk scores for young male cars show a significant increase after the policy around 10% compared to the pre-policy period. On the contrary, the risk scores for gender-neutral cars are stable over time and have no change after the policy. Figures C6 and C7 present the event study estimates for young males, gender-neutral cars, and all other cars (used as a control group in Figure 5) separately. Additionally, Figure C8 shows car risk scores for each young male car model over time separately. Furthermore, I estimate the following difference-in-differences equation to estimate changes in the young male car risk scores compared to all other car models.

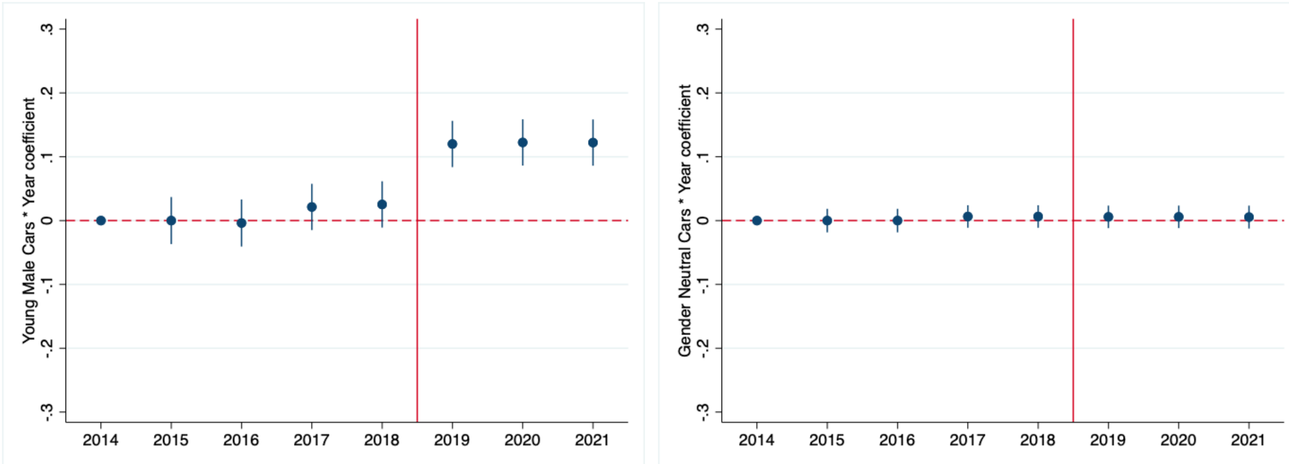
$$\log(\text{Risk Score})_{mbt} = \beta_0(\text{Post}_t \times \text{YM car}_m) + \beta_1(\text{YM car}_m) + \beta_2\text{Post}_t + \gamma_m + \eta_b + \theta_t + \epsilon_{mbt} \quad (9)$$

where $\log(\text{Risk Score})_{mbt}$ represents the logarithm of a risk score for each car model m with brand b at year t. YM car_m takes value 1 if a given car model is predicted as being used distinctively by young males among young drivers. γ_{m} and η_{b} are fixed effects for car models and brands. Post_{t} is 1 after the policy is implemented at time t, 0 otherwise. Year fixed effects denoted by θ_{t} are also included in each regression.

Table 8 shows that risk scores for young male cars have increased by 10 percent compared

[31] The criteria for gender-neutral cars is defined as for a given car model, the share of young males using this car must be between 0.45 and 0.55 among all young people. Twenty car models fall into this category, which corresponds to 66% of drivers using these models.

Figure 5: Risk scores for young male and gender-neutral cars



Notes: The analysis uses data from the pricing algorithm dataset compiled from insurance pricing filings of an insurance provider in California. The figure shows the event study estimates for young male cars and gender-neutral cars, respectively. The risk score for a given car group is regressed on year, car brand, and car model fixed effects for each regression. The omitted group is all remaining car models as not predicted as young male cars. The interaction of year and car category (young male car or gender-neutral car) is plotted. It covers the years between 2014 and 2021. The base year 2014 is omitted in each figure. 95% confidence intervals are plotted. Heteroskedasticity-robust standard errors are used for each regression.

to all other vehicles. When we also control for car brands and models, this effect becomes 11 percent. It gives similar results when we constrain our sample to young male and gender-neutral cars. Consistent with event study graphs, risk scores for young male cars rise by 10 percent after the policy.

The interpretation of risk scores and how they changed relative to each other for different groups provide strong evidence for insurance pricing algorithms' modifications. However, how risk scores translated into insurance prices is not easy to interpret with this analysis alone. We need to analyze insurance prices to examine how each customer is affected by these algorithmic adjustments. Also, the insurance price for each consumer is calculated as a combination of risk scores in different dimensions, such as risk scores according to their car model, location, or demographics. In that sense, how people will be affected by adjustments in the algorithm depends on their unique features. To illustrate, the increase in young male car risk scores could affect anyone who uses these cars regardless of age and gender. In that sense, people who have young male features in terms of observables can pay more after these adjustments. These unintended outcomes are one of the most controversial aspects of these types of anti-discriminatory laws, where decision-makers ban the use of an applicant feature.

6.2.2 Consumers with young male characteristics

In this section, I analyze changes in insurance expenditures for consumers with young male features after the policy. Previously, I provide evidence of adjustments to the pricing algorithm. In this part, I discuss how these changes translate into insurance prices based on consumer

Table 8: The impact on young male cars risk scores

Dependent variable: log(Risk Score)	(1) Full sample	(2) Full sample	(3) Full sample	(4) Young male & Gender neutral
Post * Young male car	0.104***	0.106***	0.112***	0.109***
	(0.014)	(0.016)	(0.009)	(0.007)
Year FE	✓	✓	✓	✓
Car brand FE		✓	✓	✓
Car model FE			✓	✓
Observations	59,660	59,660	59,660	1,633
Adjusted R-squared	0.026	0.317	0.779	0.859

Notes: The logarithm of the car model risk score is regressed on a young male car and post dummies and year fixed effects in all regressions. Column (2) also has car brand fixed effects. Columns (3) and (4) also have car model fixed effects. The analysis uses data from the pricing algorithm dataset compiled from the insurance pricing filings of an insurance provider in California. It covers the years between 2014 and 2021. Columns (1) to (3) use all car models, whereas Column (4) focuses on only young males and gender-neutral cars. The heteroskedasticity-robust standard errors are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

expenditure data. To measure the differential impact of the policy on consumers with young male features, I estimate the following equation:

$$\log(\text{Price})_{it} = \beta_0(\text{Post}_t \times \text{YM Car}_i) + \beta_1(\text{YM Car})_i + \gamma X_{it} + \theta_t + \epsilon_{it} \tag{10}$$

where $\log(\text{Price})_{it}$ represents the logarithm insurance premium price paid by individual i at year t. YM Car_i takes value 1 if individual i has a young male car type. X_{it} represents individual characteristics of individual i such as different demographics and car brands. Post_{t} is 1 after the policy is implemented at time t, 0 otherwise. Year fixed effects denoted by θ_{t} are also included in each regression.

To estimate changes in insurance prices for young male cars, I use the consumer expenditure dataset for the prices and prediction results for young male features from Section 6.1. However, one limitation in this exercise is that the consumer expenditure dataset has only vehicle brands, vehicle types (car or pickup trucks), and vehicle ages but not the model. To overcome this challenge, I focus on the subset of predicted car models that can be distinctive by using brand and year information from the expenditure dataset and in line with predicted young male cars. For instance, results in Section 6.1.3 suggest that young males have specific preferences for some car brands. All Chevrolet, Pontiac, and Saturn brand cars are old models of the given brands. These particular trends allow me to create car categories by interacting with their brand and age. Furthermore, based on predictive analysis, pickup trucks are also common among young males. Based on this preference, I also created another variable for pickup trucks and analyzed insurance prices for these vehicles over time.[32]

[32]There are multiple pickup truck models and brands predicted in Section 6.1. However, these vehicles are

Table 9 presents the estimation results from equation (10). Column (1) gives the estimation results for how insurance prices for young male characteristics changed. It uses an aggregate term called young male prediction. This term is created by regressing the young male dummy on all demographic, vehicle, and location characteristics used in the main analysis. The predicted value from this regression has interacted with a policy dummy. The results in Column (1) indicate that the insurance price of young male features increased by 10 percent after the policy. Column (2) shows that the insurance price for pickup trucks increased by 7 percent after the policy. Results in the remaining columns indicate that insurance prices for car types such as old models of Chevrolet, Pontiac, and Saturn brands also increased significantly. Overall, the results presented in Table 9 align with the theoretical prediction of firms proxying gender with other specifications and findings in previous sections. Also, these results further imply that a driver using young male cars needs to pay up to 22 percent more even if they are from another demographic group.

Table 9: The impact on consumers with young male features

Dependent variable: log(price)	(1) California	(2) California	(3) California	(4) California	(5) California
Post * Young Male Prediction	0.107***				
	(0.026)				
Post * Pickup Trucks		0.071***			
		(0.013)			
Post * Old Chevrolet Cars			0.102**		
			(0.047)		
Post * Old Pontiac Cars				0.221***	
				(0.087)	
Post * Old Saturn Cars					0.173**
					(0.088)
Year FE	✓	✓	✓	✓	✓
Car brand FE	✓	✓	✓	✓	✓
Observations	15,509	15,509	15,509	15,509	15,509
Adjusted R-squared	0.017	0.131	0.127	0.127	0.127

Notes: The logarithm of insurance price is regressed on dummy variables for having a given young male car brand and its interaction with post dummy and individual characteristics. Fixed effects for year and car brands are also included in each regression. The analysis uses US Consumer Expenditure Survey between 2010 and 2020, and the sample used for this regression only includes individuals in California. The heteroskedasticity-robust standard errors are reported in parentheses $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

a small portion of the population; therefore, each brand has only a few observations. Due to the small sample size of pickup trucks at the car brand level, I aggregate all pickup trucks as one variable and estimate changes in their insurance after the policy. The fact that pickup trucks are a popular and distinctive feature of being a young male is also supported in other studies. See Chandra et al. (2017).

7 Placebo analysis and permutation tests

The fundamental identifying assumption in Section 5.2 is that the outcome in California would not have evolved differently from other states in the absence of the gender-neutral pricing policy. Figure 4, Figure C4, and Figure C5 already provide empirical evidence for the time trends in control states. Furthermore, I estimate a set of placebo regressions to explore this assumption. I replicate the main analysis of how male and female drivers started to pay after the policy for the control states. Specifically, I estimate Equation (3) and Equation (5) for states that do not have gender-neutral pricing policy for auto insurance.[33]

Table 10: Placebo estimations for the gender gap in control states

Dependent variable: log(price)	(1) Placebo states	(2) Placebo states	(3) Young (All states)	(4) Young (All states)
Post * Young * Male	0.002	0.004		
	(0.028)	(0.029)		
Post * Placebo States * Male			-0.038	-0.073
			(0.048)	(0.049)
California estimates	-0.075	-0.133	-0.055	-0.178
Placebo state estimates:				
5th percentile	-0.058	-0.055	-0.135	-0.172
95th percentile	0.059	0.064	0.059	0.026
Two-tailed test p-value	0.993	0.868	0.434	0.144
Demographic characteristics		✓		✓
Location characteristics		✓		✓
Vehicle characteristics		✓		✓
Year FE	✓	✓	✓	✓
State FE	✓	✓	✓	✓
State x Year FE	✓	✓	✓	✓
Observations	110,213	110,213	5,593	5,593
Adjusted R-squared	0.060	0.070	0.019	0.419

The logarithm of insurance price is regressed on the interaction of dummies for young, male and post, and a set of controls in Columns (1) to (2). Column (3) and (4) reports the coefficient for the triple interaction term of being male, in a given placebo state and post. Demographic controls are age, marital status, and education level. Location characteristics are geographic region, population density, and metropolitan statistical area dummy. Vehicle characteristics are vehicle type, year, brand, and used status. The sample used in these regressions covers the years between 2010 and 2020, only young people. Control states are all states except California. Fixed effects for the year, state, and state*year are included in each regression. The heteroskedasticity-robust standard errors are clustered at the state level in Column (3) and Column (4). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 10 presents estimation results for this analysis. Columns (1) and (2) are estimates of the change in the young gender gap compared to the older gender gap in control states with and without controls. The remaining columns are comparisons of the gender gap in control states with gender-neutral pricing states.[34] The fact that none of the placebo regressions give

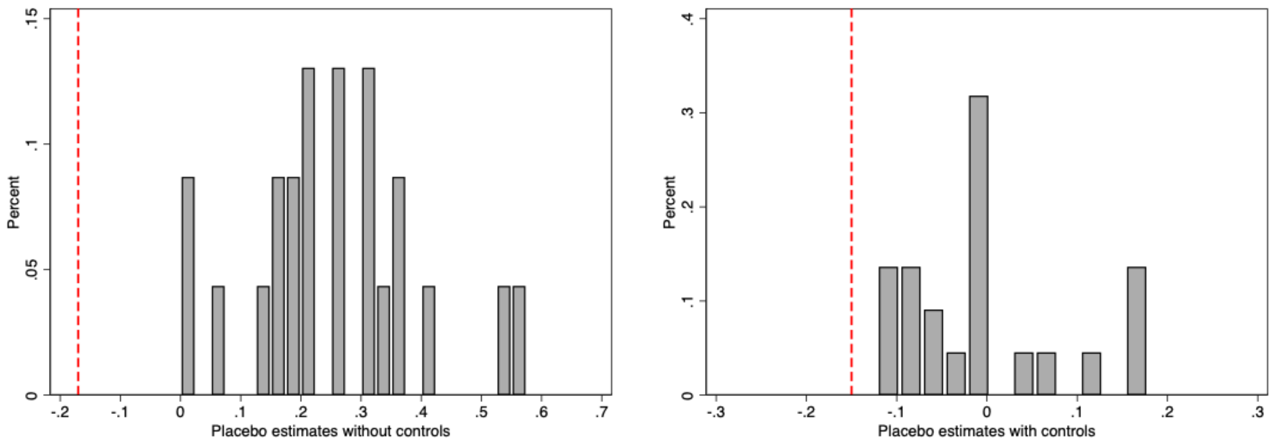
[33]Control states are all the states except California to provide a comprehensive understanding. I also conduct the analysis when I restrict control states to those with at least 200 observations for young populations to be consistent with Section 5.2. The results are robust to the findings in this part.

[34]The gender-neutral pricing states include California, Hawaii, Montana, Massachusetts, Montana, North

statistically significant coefficients for control states provides further evidence for the identifying assumption and the differential impact of the policy on California.

Next, by adopting a similar approach to Cunningham and Shah (2018), I conduct a set of randomization tests and plot them. Specifically, I estimate Equation (3) for each control state separately with and without controls. In that sense, I repeat the analysis reported in Columns (1) and (2) in Table 10 for each control state and compare these estimates with the estimate from the main analysis conducted for California.

Figure 6: Permutation tests for control states and California



Notes: These figures show the effects estimated from permutation tests in Columns (1) and (2) from Table 10. The left and right figures show the estimated effects with and without controls, respectively. The dashed red lines are the policy estimate for California. It uses data from Consumer Expenditure Survey between 2010 and 2020.

As shown in Figure 6, the estimates for California (red dashed lines) lie at the left point in the histograms, implying that they are at the bottom of the placebo distribution. These results help alleviate the potential concerns about the endogeneity of the policy implementation and support the setting for difference-in-differences as an empirical strategy.

8 Conclusion

Today, firms are deploying algorithms in different settings, such as determining prices, resume screening or even judicial decisions. This increasing reliance on algorithms raises issues about their fairness. The main arguments are that they may use biased or unrepresentative training data or group attributes to infer individual-level merits. However, identifying and measuring such algorithmic bias is not always easy, due to the opaqueness of these processes. For example, sometimes, even if some regulations ban the use of specific group attributes (such as race or gender), the high-dimensional data enables algorithms to use correlated characteristics. This prompts questions about whether anti-discriminatory policies that require group blind pricing

Carolina, and Pennsylvania.

are adequate, or whether algorithms using correlated variables can lessen the effectiveness of these policies.

I investigate the impact of gender-neutral insurance pricing policies on the gender gap in insurance premiums and firms' reactions. I analyze how the policy affects premiums paid by male and female drivers and whether it eliminates the gender gap. To answer these questions, I focus on a recent law change in California requiring gender-neutral pricing in auto insurance. I find that the policy decreased the gender gap among young people in California by 70 percent, but a significant part of the gap remains.

I analyze insurance firms' responses to this change, specifically the changes in the pricing algorithms after the gender-blind requirements. I construct a novel dataset by extracting data from the pricing filings of a large insurance provider in California. This dataset allows me to study not only the changes in insurance prices but also the changes in the price-generating mechanism. My analysis indicates that the pricing algorithms were adjusted after the regulation, and they started to give larger weights to gender proxies. Specifically, they consider characteristics associated with young males, the riskiest group, 10 percent riskier than before. These readjustments reflect around an 11 percent increase in insurance premiums for people whose observables match those of young males.

My finding is a significant example of how algorithmic pricing may perpetuate the bias embedded in previous customer data, even though the anti-discrimination policies aim to achieve more equity. A critical implication of this paper is understanding to what extent anti-discrimination policies achieve their aims, especially when dealing with inherent bias in algorithms. The anti-discrimination initiatives have good intentions and recognize that discrimination is a significant problem. However, these well-intended policies may not necessarily bring equity in all settings. In that sense, discrimination may still exist in other forms and may be built into the algorithms designed to automate decision processes.

References

Agan, A., Cowgill, B., & Gee, L. (2020). *Salary disclosure and hiring: Field experimental evidence from a two-sided audit study* (tech. rep.). Working paper.

Agan, A., & Starr, S. (2018). Ban the box, criminal records, and racial discrimination: A field experiment. *The Quarterly Journal of Economics*, *133*(1), 191–235.

Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Ilr Review*, *30*(2), 175–187.

Altonji, J. G., & Pierret, C. R. (2001). Employer learning and statistical discrimination. *The quarterly journal of economics*, *116*(1), 313–350.

Andini, M., Ciani, E., de Blasio, G., D'Ignazio, A., & Salvestrini, V. (2018). Targeting with machine learning: An application to a tax rebate program in italy. *Journal of Economic Behavior & Organization*, *156*, 86–102.

Angrist, J. D., & Frandsen, B. (2022). Machine labor. *Journal of Labor Economics*, *40*(S1), S97–S140.

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias. propublica, may 23, 2016.

Arnold, D., Dobbie, W., & Hull, P. (2021). Measuring racial discrimination in algorithmsy. *AEA Papers and Proceedings*, *111*, 49–54.

Arrow, K. J. (2015). The theory of discrimination. In *Discrimination in labor markets* (pp. 1–33). Princeton University Press.

Ascarza, E., & Israeli, A. (2022). Eliminating unintended bias in personalized policies using bias-eliminating adapted trees (beat). *Proceedings of the National Academy of Sciences*, *119*(11), e2115293119.

Asker, J., Fershtman, C., & Pakes, A. (2021). *Artificial intelligence and pricing: The impact of algorithm design* (tech. rep.). National Bureau of Economic Research.

Assad, S., Clark, R., Ershov, D., & Xu, L. (2020). Algorithmic pricing and competition: Empirical evidence from the german retail gasoline market.

Athey, S. (2017). Beyond prediction: Using big data for policy problems. *Science*, *355*(6324), 483–485.

Athey, S. (2019). The impact of machine learning on economics. the economics of artificial intelligence. *Agrawal. A, Gans, J. and Goldfarb, A.*

Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, *11*, 685–725.

Avraham, R., Logue, K. D., & Schwarcz, D. (2013). Understanding insurance antidiscrimination law. *S. Cal. L. Rev.*, *87*, 195.

Bartik, A., & Nelson, S. (2016). Deleting a signal: Evidence from pre-employment credit checks.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014a). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, *28*(2), 29–50.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014b). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, *81*(2), 608–650.

Brown, Z. Y., & MacKay, A. (2021). *Competition in pricing algorithms* (tech. rep.). National Bureau of Economic Research.

Calvano, E., Calzolari, G., Denicolo, V., & Pastorello, S. (2020). Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review*, *110*(10), 3267–97.

Chalfin, A., Danieli, O., Hillis, A., Jelveh, Z., Luca, M., Ludwig, J., & Mullainathan, S. (2016). Productivity and selection of human capital with machine learning. *American Economic Review*, *106*(5), 124–27.

Chandler, D., Levitt, S. D., & List, J. A. (2011). Predicting and preventing shootings among at-risk youth. *American Economic Review*, *101*(3), 288–92.

Chandra, A., Gulati, S., & Sallee, J. M. (2017). Who loses when prices are negotiated? an analysis of the new car market. *The Journal of Industrial Economics*, *65*(2), 235–274.

Cohen, A. (2005). Asymmetric information and learning: Evidence from the automobile insurance market. *Review of Economics and statistics*, *87*(2), 197–207.

Council, I. R. (2021). *Auto insurance affordability: Countrywide trends and state comparisons* (tech. rep.). IRC.

Cowgill, B. (2018). Bias and productivity in humans and algorithms: Theory and evidence from resume screening. *Columbia Business School, Columbia University*, *29*.

Cowgill, B., & Tucker, C. E. (2020). Algorithmic fairness and economics. *Columbia Business School Research Paper*.

Crocker, K. J., & Snow, A. (1986). The efficiency effects of categorical discrimination in the insurance industry. *Journal of Political Economy*, *94*(2), 321–344.

Cunningham, S., & Shah, M. (2018). Decriminalizing indoor prostitution: Implications for sexual violence and public health. *The Review of Economic Studies*, *85*(3), 1683–1715.

Datta, A., Tschantz, M. C., & Datta, A. (2014). Automated experiments on ad privacy settings: A tale of opacity, choice, and discrimination. *arXiv preprint arXiv:1408.6491*.

Demyanyk, Y., & Van Hemert, O. (2011). Understanding the subprime mortgage crisis. *The review of financial studies*, *24*(6), 1848–1880.

Doleac, J. L., & Hansen, B. (2020). The unintended consequences of "ban the box": Statistical discrimination and employment outcomes when criminal histories are hidden. *Journal of Labor Economics*, *38*(2), 321–374.

Elul, R., Souleles, N. S., Chomsisengphet, S., Glennon, D., & Hunt, R. (2010). What" triggers" mortgage default? *American Economic Review*, *100*(2), 490–94.

Finkelstein, A., & Poterba, J. (2004). Adverse selection in insurance markets: Policyholder evidence from the uk annuity market. *Journal of political economy*, *112*(1), 183–208.

Finkelstein, A., Poterba, J., & Rothschild, C. (2009). Redistribution by insurance market regulation: Analyzing a ban on gender-based retirement annuities. *Journal of financial economics*, *91*(1), 38–58.

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as data. *Journal of Economic Literature*, *57*(3), 535–74.

Gruber, J. (1994). The incidence of mandated maternity benefits. *The American economic review*, 622–641.

Hansen, B., & McNichols, D. (2020). *Information and the persistence of the gender wage gap: Early evidence from california's salary history ban* (tech. rep.). National Bureau of Economic Research.

Hastie, T., Tibshirani, R., & Tibshirani, R. J. (2017). Extended comparisons of best subset selection, forward stepwise selection, and the lasso. *arXiv preprint arXiv:1707.08692*.

Hay, J., & Leahy, M. J. (1982). Physician-induced demand: An empirical analysis of the consumer information gap. *Journal of Health Economics*, *1*(3), 231–244.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67.

Huang, S., & Salm, M. (2020). The effect of a ban on gender-based pricing on risk selection in the german health insurance market. *Health economics*, *29*(1), 3–17.

Huh, J., & Reif, J. (2021). Teenage driving, mortality, and risky behaviors. *American Economic Review: Insights*, *3*(4), 523–39.

Hunter, J. R., Feltner, T., & Heller, D. (2013). What works: A review of auto insurance rate regulation in america and how best practices save billions of dollars. *Consumer Federation of America, http://www. consumerfed. org/pdfs/whatworks-report_nov2013_hunter-feltner-heller. pdf*.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.

Klein, T. (2021). Autonomous algorithmic collusion: Q-learning under sequential pricing. *The RAND Journal of Economics*, *52*(3), 538–558.

Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction policy problems. *American Economic Review*, *105*(5), 491–95.

Kleinberg, J., Ludwig, J., Mullainathan, S., & Rambachan, A. (2018). Algorithmic fairness. *Aea papers and proceedings*, *108*, 22–27.

Kleinberg, J., Ludwig, J., Mullainathan, S., & Sunstein, C. R. (2020). Algorithms as discrimination detectors. *Proceedings of the National Academy of Sciences*, *117*(48), 30096–30100.

Lambrecht, A., & Tucker, C. (2019). Algorithmic bias? an empirical study of apparent gender-based discrimination in the display of stem career ads. *Management science*, *65*(7), 2966–2981.

Lundberg, S. J., & Startz, R. (1983). Private discrimination and social intervention in competitive labor market. *The American economic review*, *73*(3), 340–347.

Moore, T. J., & Morris, T. (2021). *Shaping the habits of teen drivers* (tech. rep.). National Bureau of Economic Research.

Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87–106.

Peysakhovich, A., & Eckles, D. (2018). Learning causal effects from many randomized experiments using regularized instrumental variables. *Proceedings of the 2018 World Wide Web Conference*, 699–707.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review*, *62*(4), 659–661.

Polborn, M. K., Hoy, M., & Sadanand, A. (2006). Advantageous effects of regulatory adverse selection in the life insurance market. *The Economic Journal*, *116*(508), 327–354.

Pope, D. G., & Sydnor, J. R. (2011). Implementing anti-discrimination policies in statistical profiling models. *American Economic Journal: Economic Policy*, *3*(3), 206–31.

Puelz, R., & Snow, A. (1994). Evidence on adverse selection: Equilibrium signaling and cross-subsidization in the insurance market. *Journal of Political Economy*, *102*(2), 236–257.

Rea, S. A. (1987). The market response to the elimination of sex-based annuities. *Southern Economic Journal*, 55–63.

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso.

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, *28*(2), 3–28.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, *101*(476), 1418–1429.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, *67*(2), 301–320.

Zou, H., & Zhang, H. H. (2009). On the adaptive elastic-net with a diverging number of parameters. *Annals of statistics*, *37*(4), 1733.

Appendices

A Model

First This sections presents the proofs of the theoretical predictions presented in Section 3.3.

Prediction 1: Expected risk between two groups will be less in blind case if firms cannot perfectly predict group characteristics. Let us assume group 1 is riskier than group 2 on average without loss of generality, i.e. $(p_1 > p_2)$.

$$E^{Non-blind}(x_i|y_i, i \in g_1) - E^{Non-blind}(x_i|y_i, i \in g_2) > E^{Blind}(x_i|y_i, i \in g_1) - E^{Blind}(x_i|y_i, i \in g_2)$$

Proof:

$$E^{Non-blind}(x_i|y_i, i \in g_1) - E^{Non-blind}(x_i|y_i, i \in g_2) > E^{Blind}(x_i|y_i, i \in g_1) - E^{Blind}(x_i|y_i, i \in g_2)$$

$$\gamma y_i + (1-\gamma)p_1 - \gamma y_i - (1-\gamma)p_2 > \gamma y_i + (1-\gamma)E(p_1|y_i) - \gamma y_i - (1-\gamma)E(p_2|y_i)$$

$$(1-\gamma)(p_1 - p_2) > (1-\gamma)[E(p_1|y_i) - E(p_2|y_i)]$$

$$(p_1 - p_2) > P(i \in g_1|y_i)p_1 + P(i \in g_2|y_i)p_2 - P(i \in g_1|y_i)p_1 - P(i \in g_2|y_i)p_2$$

$$(\mathrm{p}_1 - \mathrm{p}_2) > 0 \qquad \blacksquare$$

Prediction 2: Signals $\mathrm{y_i}$ positively correlated with group characteristics will gain more weight in risk prediction.

$$\mathrm{E}^{\mathrm{Non-blind}}(\mathrm{x_i}|\mathrm{y_i}) = \gamma_0 y_i + (1-\gamma_0)p_i$$

and

$$\mathrm{E}^{\mathrm{Blind}}(\mathrm{x_i}|\mathrm{y_i}) = \gamma_1 y_i + (1-\gamma_1)\mathrm{E}(\mathrm{p_i})$$

Suggesting that, $\gamma_1 \geq \gamma_0$.

Proof:

$$\mathrm{E}^{\mathrm{Blind}}(\mathrm{x_i}|\mathrm{y_i}) = \gamma_1 y_i + (1-\gamma_1)\mathrm{E}(\mathrm{p_i})$$

$$\mathrm{E}^{\mathrm{Blind}}(\mathrm{x_i}|\mathrm{y_i}) = \gamma_1 y_i + (1-\gamma_1)[\mathrm{P}(\mathrm{i} \in \mathrm{g_1}|\mathrm{y_i})\mathrm{p_1} + \mathrm{P}(\mathrm{i} \in \mathrm{g_2}|\mathrm{y_i})\mathrm{p_2}]$$

$$\mathrm{E}^{\mathrm{Blind}}(\mathrm{x_i}|\mathrm{y_i}) = \gamma_1 y_i + (1-\gamma_1)[(\hat{\alpha_0} + \hat{\alpha_1}\mathrm{y_i})\mathrm{p_1} + (\hat{\beta_0} + \hat{\beta_1}_{\mathrm{i}})\mathrm{p_2}]$$

$$\mathrm{E}^{\mathrm{Blind}}(\mathrm{x_i}|\mathrm{y_i}) = \gamma_1 y_i + (1-\gamma_1)[(\alpha_0\mathrm{p_1} + \beta_0\mathrm{p_2}) + (\alpha_1\mathrm{p_1} + \beta_1\mathrm{p_2})\mathrm{y_i}]$$

Let us call the constant term $\alpha_0\mathrm{p_1} + \beta_0\mathrm{p_2}$ as c.

$$\mathrm{E}^{\mathrm{Blind}}(\mathrm{x_i}|\mathrm{y_i}) = \gamma_1 y_i + (1-\gamma_1)[\mathrm{c} + (\alpha_1\mathrm{p_1} + \beta_1\mathrm{p_2})\mathrm{y_i}]$$

$$\mathrm{E}^{\mathrm{Blind}}(\mathrm{x_i}|\mathrm{y_i}) = \underbrace{[\gamma_1 + (1-\gamma_1)(\alpha_1\mathrm{p_1} + \beta_1\mathrm{p_2})]}_{\text{Weights given to signal}}\mathrm{y_i} + \mathrm{c}$$

Whereas the expected risk $\mathrm{x_i}$ given signal $\mathrm{y_i}$ in non-blind case is

$$\mathrm{E}^{\mathrm{Non-Blind}}(\mathrm{x_i}|\mathrm{y_i}) = \gamma_1 y_i + (1-\gamma_1)\mathrm{p_i}$$

By definition $\gamma > 0$ and $0 \leq \mathrm{p_i} \leq 1$.

Also, β_1 is non-negative by definition as it is the coefficient in the regressions of signals $\mathrm{y_i}$

that are positively correlated with group characteristics.

Then, we can infer that

$$\gamma_1 + (1 - \gamma_1)(\alpha_1 \mathrm{p}_1 + \beta_1 \mathrm{p}_2) > \gamma_0 \qquad \blacksquare$$

B Data

B.1 Tables

Table B1: Extended summary statistics for consumer characteristics dataset

	Full Sample		Young males	
	Mean	St Dev	Mean	St Dev
Vehicle year	2002.53	8.39	2001.2	7.93
Vehicle age	9.56	7.46	10.87	6.88
Annual miles	9,660.48	11,035.35	10,612	13,676.83
Odometer reading	83,504.98	67,419.88	106,162.50	68,192.57
Gas	0.94	0.22	0.97	0.16
Car	0.49	0.49	0.59	0.45
Van	0.07	0.23	0.02	0.11
Sport utility vehicle	0.21	0.38	0.13	0.25
Pickup	0.16	0.36	0.21	0.31
Age	53.83	18.38	18.65	3.47
Male	0.46	0.49	N/A	N/A
White	0.84	0.36	0.78	0.41
Black	0.06	0.24	0.07	0.26
Asian	0.02	0.16	0.04	0.2
Household size	2.57	1.31	3.82	1.41
Middle school	0.07	0.25	0.24	0.42
High school	0.22	0.41	0.19	0.39
College	0.25	0.43	0.2	0.4
Bachelor degree	0.19	0.39	0.07	0.26
Graduate degree	0.15	0.36	0.01	0.011
Sales job	0.11	0.31	0.19	0.39
Admin job	0.04	0.21	0.02	0.14
Manufacturing job	0.07	0.26	0.11	0.31
Managerial job	0.22	0.39	0.08	0.27
MSA over one million population	0.16	0.37	0.18	0.39
MSA with 500k-1m population	0.28	0.45	0.3	0.45
MSA with 200k-500k population	0.35	0.47	0.34	0.47
Housing units per sq mile	1,580.62	2,983.88	1,708.23	3,073.41
Population density per sq mile	3,577.73	5,169.39	4,118.48	5,694.11
Observations	676,574		29,731	

Notes: This table summarizes the main statistics for the NHTS dataset for the entire sample used in this study and only for young males under 25 separately. Car, sport utility vehicle, and pickup are dummy variables taking values of 0 or 1 and representing different vehicle categories. MSA over 1M population refers to Metropolitan Statistical Areas with at least 1 million.

B.2 Figures

Relativities for Each Rating Factor

Marital Status / Gender / Years Driving Experience	BI	PD	UMBI	COMP	COLL	MED	UMPD	LOAN
MF00	1.10	1.12	1.19	1.34	1.33	1.19	1.33	1.34
MF01	1.10	1.12	1.19	1.34	1.33	1.19	1.33	1.34
MF02	1.10	1.04	1.19	1.34	1.17	1.08	1.17	1.34
MF03	1.00	0.89	1.06	1.32	1.17	1.03	1.17	1.32
MF04	0.84	0.89	0.85	1.29	1.00	0.95	1.00	1.29
MF05	0.74	0.88	0.76	1.29	0.95	0.91	0.95	1.29
MF06	0.66	0.82	0.69	1.19	0.87	0.90	0.87	1.19
MF07	0.65	0.78	0.64	1.11	0.84	0.87	0.84	1.11
MF08	0.65	0.75	0.64	1.02	0.77	0.82	0.77	1.02
MF09	0.64	0.75	0.64	1.00	0.77	0.82	0.77	1.00
MF10	0.64	0.75	0.64	1.00	0.77	0.82	0.77	1.00
MF11	0.59	0.67	0.64	0.98	0.75	0.72	0.75	0.98
MF12	0.59	0.67	0.56	0.93	0.75	0.66	0.75	0.93
MF13	0.59	0.67	0.56	0.84	0.74	0.66	0.74	0.84
MF14	0.63	0.72	0.64	0.84	0.74	0.73	0.74	0.84
MF15-18	0.63	0.72	0.64	0.84	0.74	0.73	0.74	0.84
MF19-23	0.76	0.82	0.70	0.79	0.79	0.78	0.79	0.79
MF24-28	0.88	0.86	0.86	0.78	0.86	0.82	0.86	0.78
MF29-33	0.88	0.89	0.86	0.78	0.86	0.87	0.86	0.78
MF34-38	0.87	0.89	0.86	0.69	0.85	0.87	0.85	0.69
MF39-43	0.87	0.80	0.82	0.61	0.74	0.87	0.74	0.61
MF44-48	0.77	0.77	0.76	0.61	0.61	0.78	0.61	0.61
MF49-53	0.77	0.80	0.78	0.52	0.61	0.81	0.61	0.52
MF54-58	0.85	0.93	0.85	0.56	0.67	0.88	0.67	0.56
MF59+	0.85	0.93	0.85	0.56	0.67	0.88	0.67	0.56
SF00	1.40	1.47	1.36	1.03	1.32	1.49	1.32	1.03
SF01	1.40	1.47	1.36	1.03	1.25	1.49	1.25	1.03
SF02	1.40	1.47	1.36	1.01	1.23	1.44	1.23	1.01
SF03	1.38	1.43	1.29	1.00	1.23	1.44	1.23	1.00
SF04	1.36	1.32	1.27	0.99	1.21	1.44	1.21	0.99
SF05	1.21	1.25	1.18	0.97	1.19	1.26	1.19	0.97
SF06	1.12	1.15	1.05	0.96	1.10	1.17	1.10	0.96
SF07	1.07	1.08	1.05	0.96	1.06	1.12	1.06	0.96
SF08	0.90	0.92	0.94	0.92	0.95	0.93	0.95	0.92
SF09	0.90	0.88	0.90	0.92	0.92	0.92	0.92	0.92
SF10	0.89	0.86	0.90	0.92	0.92	0.92	0.92	0.92
SF11	0.89	0.86	0.90	0.92	0.92	0.92	0.92	0.92
SF12	0.88	0.85	0.85	0.92	0.89	0.92	0.89	0.92
SF13	0.88	0.85	0.83	0.92	0.89	0.92	0.89	0.92
SF14	0.88	0.82	0.82	0.92	0.89	0.94	0.89	0.92
SF15-18	0.88	0.86	0.90	0.92	0.89	0.94	0.89	0.92
SF19-23	1.00	0.93	1.01	0.92	0.89	1.01	0.89	0.92
SF24-28	1.00	1.00	1.04	0.88	0.90	1.10	0.90	0.88
SF29-33	1.01	1.00	1.04	0.88	0.90	1.10	0.90	0.88
SF34-38	1.01	1.00	1.04	0.86	0.90	1.07	0.90	0.86
SF39-43	1.08	0.93	1.04	0.76	0.90	1.01	0.90	0.76
SF44-48	1.08	0.93	1.05	0.82	0.90	1.01	0.90	0.82
SF49-53	1.08	1.05	1.05	0.82	0.95	1.08	0.95	0.82
SF54-58	1.08	1.05	1.05	0.82	0.99	1.08	0.99	0.82
SF59+	1.08	1.05	1.05	0.82	0.99	1.08	0.99	0.82

Figure B1: Risk score table based on gender, marital status and driving experience

Notes: This is an example from an insurer's filing on how insured characteristics translated into risk sores. Each column represents different insurance coverage.

Relativities for Each Rating Factor

Type of Vehicle / Symbol

Model Year	Make	Model	Style	BI	PD	UMBI	MED	COMP / LOAN	COLL / UMPD
1997	AC	CL	24	0.91	1.06	1.15	1.15	0.98	1.31
1998	AC	CL	24	0.91	1.06	1.18	1.18	0.96	1.32
1999	AC	CL	24	0.92	1.06	1.12	1.12	0.90	1.33
1997	AC	CL	26	1.22	1.13	1.00	1.00	1.10	1.51
1998	AC	CL	26	1.23	1.08	0.87	0.87	1.07	1.50
1999	AC	CL	26	1.23	1.06	1.07	1.07	1.09	1.50
2001	AC	CL	26	1.29	1.20	0.92	0.92	1.32	1.73
2002	AC	CL	26	1.28	1.21	0.94	0.94	1.30	1.74
2003	AC	CL	26	1.27	1.21	1.15	1.15	1.32	1.73
2013 ... 2025	AC	HX	44	1.11	1.10	1.15	1.15	1.15	1.01
1981 ... 1989	AC	LE	26	1.02	1.00	1.08	1.08	1.69	1.34
1990	AC	LE	26	1.04	1.00	1.09	1.09	1.86	1.57
1991	AC	LE	26	1.08	1.00	1.04	1.04	2.01	2.06
1992	AC	LE	26	1.12	1.00	1.03	1.03	2.15	2.24
1993	AC	LE	26	1.12	1.00	1.02	1.02	2.18	2.20
1994	AC	LE	26	1.10	0.99	1.01	1.01	2.19	2.24
1995	AC	LE	26	1.10	0.99	1.01	1.01	2.21	2.24
1981 ... 1989	AC	LE	46	1.23	1.07	1.14	1.14	1.32	1.39
1990	AC	LE	46	1.24	1.07	1.13	1.13	1.57	1.42
1991	AC	LE	46	1.25	1.07	1.13	1.13	1.49	1.44
1992	AC	LE	46	1.26	1.07	1.12	1.12	1.52	1.51
1993	AC	LE	46	1.26	1.07	1.12	1.12	1.51	1.56
1994	AC	LE	46	1.26	1.07	1.11	1.11	1.51	1.59
1995	AC	LE	46	1.26	1.06	1.12	1.12	1.51	1.63
2013 ... 2025	AC	LX	44	1.22	1.22	1.15	1.15	1.27	0.77
2001	AC	MD	5K	0.87	1.18	1.01	1.01	0.80	1.07
2002	AC	MD	5K	0.86	1.17	0.99	0.99	0.86	1.04
2003	AC	MD	5K	0.85	1.17	0.97	0.97	0.93	1.02
2004	AC	MD	5K	0.83	1.17	0.94	0.94	0.99	0.99
2005	AC	MD	5K	0.84	1.17	0.94	0.94	1.00	0.95
2006	AC	MD	5K	0.84	1.17	0.93	0.93	0.98	1.00
2007	AC	MD	5K	0.84	1.17	0.93	0.93	0.97	1.02
2008	AC	MD	5K	0.84	1.17	0.92	0.92	0.95	1.04
2009	AC	MD	5K	0.84	1.17	0.93	0.93	0.94	1.04
2010	AC	MD	5K	0.84	1.17	0.90	0.90	0.92	1.04
2011	AC	MD	5K	0.90	1.17	0.90	0.90	0.93	1.04
2012	AC	MD	5K	0.90	1.17	0.90	0.90	0.96	1.04
2013 ... 2025	AC	MD	5K	0.90	1.07	0.90	0.90	0.94	0.99
1991	AC	NS	26	0.90	0.72	0.92	0.92	2.98	3.69
1992	AC	NS	26	0.89	0.71	0.92	0.92	3.02	3.47
1993	AC	NS	26	0.89	0.73	0.95	0.95	3.02	3.23
1994	AC	NS	26	0.94	0.74	0.96	0.96	3.01	3.30
1995	AC	NS	26	0.94	0.74	0.96	0.96	3.02	3.45
1996	AC	NS	26	0.94	0.74	0.97	0.97	3.19	3.45
1997	AC	NS	26	0.95	0.74	0.97	0.97	3.39	3.45

Figure B2: Risk score table based on vehicle models

Notes: This is an example from an insurer's filing on how insured characteristics translated into risk sores. Each column represents different insurance coverage.

Relativities for Each Rating Factor

Annual Mileage	BI	PD	UMBI	COMP	COLL	MED	UMPD
1000 ... 1499	0.56	0.60	0.59	0.55	0.50	0.51	0.52
1500 ... 1999	0.59	0.63	0.62	0.57	0.52	0.53	0.54
2000 ... 2499	0.62	0.65	0.65	0.59	0.54	0.55	0.56
2500 ... 2999	0.65	0.68	0.68	0.61	0.56	0.57	0.58
3000 ... 3499	0.68	0.72	0.71	0.63	0.58	0.59	0.60
3500 ... 3999	0.74	0.75	0.74	0.65	0.61	0.61	0.62
4000 ... 4499	0.78	0.79	0.79	0.67	0.63	0.63	0.64
4500 ... 4999	0.83	0.83	0.82	0.69	0.65	0.65	0.66
5000 ... 5499	0.88	0.87	0.85	0.70	0.67	0.67	0.68
5500 ... 5999	0.92	0.89	0.88	0.72	0.69	0.68	0.70
6000 ... 6499	0.95	0.93	0.91	0.77	0.72	0.74	0.72
6500 ... 6999	0.98	0.96	0.94	0.79	0.74	0.76	0.74
7000 ... 7499	1.01	0.98	0.98	0.81	0.76	0.78	0.76
7500 ... 7999	1.02	1.01	1.01	0.83	0.78	0.80	0.78
8000 ... 8499	1.04	1.03	1.04	0.84	0.79	0.83	0.79
8500 ... 8999	1.05	1.06	1.07	0.85	0.80	0.86	0.80
9000 ... 9499	1.07	1.08	1.10	0.86	0.82	0.89	0.81
9500 ... 9999	1.10	1.12	1.14	0.88	0.83	0.90	0.82
10000 ... 10499	1.13	1.15	1.18	0.90	0.84	0.91	0.83
10500 ... 10999	1.16	1.17	1.19	0.92	0.85	0.92	0.84
11000 ... 11499	1.17	1.20	1.21	0.93	0.86	0.93	0.85
11500 ... 11999	1.17	1.22	1.22	0.94	0.86	0.94	0.85
12000 ... 12499	1.19	1.24	1.22	0.95	0.87	0.95	0.86
12500 ... 12999	1.25	1.25	1.24	0.96	0.88	0.96	0.87
13000 ... 13499	1.31	1.26	1.25	0.97	0.89	0.97	0.88
13500 ... 13999	1.34	1.27	1.27	0.98	0.90	0.98	0.89
14000 ... 14499	1.37	1.30	1.28	0.99	0.93	0.99	0.91
14500 ... 14999	1.40	1.31	1.31	1.00	0.95	1.00	0.93
15000 ... 15499	1.43	1.32	1.35	1.01	0.96	1.01	0.94
15500 ... 15999	1.47	1.34	1.36	1.02	0.97	1.01	0.95
16000 ... 16499	1.50	1.35	1.38	1.03	0.98	1.02	0.96
16500 ... 17999	1.55	1.36	1.38	1.04	0.99	1.02	0.97
18000 ... 19499	1.55	1.37	1.39	1.07	1.00	1.03	0.98
19500 ... 20999	1.55	1.40	1.41	1.09	1.03	1.04	1.00
21000 ... 22499	1.56	1.43	1.42	1.12	1.05	1.05	1.02
22500 ... 23999	1.64	1.50	1.51	1.14	1.07	1.10	1.04
24000 ... 25499	1.73	1.56	1.51	1.15	1.08	1.10	1.05
25500 ... 26999	1.80	1.63	1.59	1.17	1.09	1.15	1.06
27000 ... 28499	1.80	1.63	1.59	1.18	1.09	1.15	1.06
28500 ... 29999	1.89	1.70	1.67	1.20	1.09	1.20	1.06
30000 ... 99999	1.89	1.70	1.67	1.21	1.11	1.20	1.08

Figure B3: Risk score table based on annual mileage

Notes: This is an example from an insurer's filing on how insured characteristics translated into risk sores. Each column represents different insurance coverage.

Zip Code Factors

Zip Code	BI	PD	Med	UM/UIM	Comp	Coll
90001	1.3181	1.0324	1.4620	1.3181	1.6125	1.2563
90002	1.4484	1.1385	1.6021	1.4484	1.4744	1.3023
90003	1.5334	1.0324	1.7442	1.5334	1.4006	1.3532
90004	2.0592	1.5398	1.9950	2.0592	1.3433	1.2928
90005	2.0592	1.6294	1.7998	2.0592	1.1470	1.3132
90006	2.0406	1.6061	2.0136	2.0406	1.2269	1.4081
90007	1.6769	1.1800	1.9100	1.6769	1.5087	1.3532
90008	1.6508	1.3438	1.7312	1.6508	1.4429	1.3775
90009	1.1556	1.0586	1.1696	1.1556	1.0326	1.1627
90010	2.0592	1.5240	1.7596	2.0592	1.3224	1.2713
90011	1.6039	1.3603	1.6319	1.6039	1.6332	1.3132
90012	1.8623	1.4260	2.2050	1.8623	1.2232	1.3023
90013	1.5686	1.3071	1.8171	1.5686	1.3023	1.2713
90014	1.5004	1.4159	2.0580	1.5004	1.3532	1.4572
90015	1.8623	1.4159	1.7442	1.8623	1.3862	1.3224
90016	1.7647	1.3438	1.6455	1.7647	1.3775	1.2402
90017	2.0119	1.4018	1.8419	2.0119	1.2761	1.3551
90018	1.7729	1.1800	1.7915	1.7729	1.3862	1.4319
90019	1.9936	1.4989	1.8060	1.9936	1.2232	1.2928
90020	2.0018	1.5450	1.7998	2.0018	1.1802	1.3745
90021	1.6039	1.3603	1.6455	1.6039	1.3132	1.2232
90022	1.1794	0.9252	1.2530	1.1794	1.1267	1.0057
90023	1.3855	0.9654	1.2928	1.3855	1.1802	1.0181
90024	1.5498	1.4018	1.3082	1.5498	0.9888	1.2058
90025	1.4570	1.4737	1.3082	1.4570	1.1627	1.1802
90026	2.0592	1.4159	1.9767	2.0592	1.2713	1.3745
90027	2.0592	1.4418	2.1521	2.0592	1.5749	1.3287
90028	1.8727	1.5264	1.6916	1.8727	1.3745	1.3287
90029	2.0119	1.5785	2.0774	2.0119	1.5200	1.4896
90031	1.5334	1.1705	1.5917	1.5334	1.0393	1.2269
90032	1.4218	1.2034	1.4620	1.4218	1.0484	1.1200
90033	1.5206	1.0891	1.6995	1.5206	1.0393	1.0702
90034	1.4418	1.2388	1.1696	1.4418	1.2563	1.1267
90035	1.6769	1.4018	1.5128	1.6769	1.2269	1.2402
90036	1.8623	1.4989	1.6916	1.8623	1.1470	1.1888
90037	1.6039	1.0754	1.9157	1.6039	1.6125	1.2761
90038	1.8623	1.4260	1.6995	1.8623	1.3775	1.2269
90039	1.8623	1.3295	1.9866	1.8623	1.2563	1.3862
90040	1.3020	1.0161	1.2357	1.3020	1.3224	1.0702
90041	1.6039	1.2205	1.8060	1.6039	1.1888	1.2525
90042	1.3110	1.1926	1.4410	1.3110	1.1128	1.1470
90043	1.5810	1.1340	1.5537	1.5810	1.4006	1.3132
90044	1.3643	1.0131	1.7780	1.3643	1.4896	1.2058
90045	1.2573	1.0586	1.1696	1.2573	1.0326	1.1627
90046	1.7507	1.4989	1.5008	1.7507	1.1200	1.2058
90047	1.4713	1.0161	1.5008	1.4713	1.4081	1.2269
90048	1.6926	1.4917	1.5738	1.6926	1.4006	1.2402
90049	1.6039	1.7144	1.5008	1.6039	1.0181	1.1627
90052	1.3181	1.0324	1.4620	1.3181	1.6125	1.2563
90056	1.3855	1.4418	1.4858	1.3855	1.1802	1.3132
90057	1.9936	1.4018	1.9502	1.9936	1.2232	1.3745
90058	1.3855	1.1499	1.5738	1.3855	1.4319	1.2928
90059	1.2573	0.9126	1.2975	1.2573	1.3532	1.2563
90061	1.5498	0.9654	1.7596	1.5498	1.6125	1.3532
90062	1.6287	1.0718	1.8171	1.6287	1.5837	1.4429
90063	1.3855	0.9545	1.3568	1.3855	1.0806	1.0626
90064	1.3806	1.2388	1.3270	1.3806	1.2525	1.1267
90065	1.4713	1.2836	1.5738	1.4713	1.2563	1.2563
90066	1.3454	1.2034	1.2051	1.3454	1.0057	1.0393
90067	1.5498	1.4917	1.4620	1.5498	1.1470	1.2232

Zip Code Page 1

Figure B4: Risk score table based on ZIP codes

Notes: This is an example from an insurer's filing on how insured characteristics translated into risk sores. Each column represents different insurance coverage.

Development of Rate Manual

	BI	PD	UMBI	UMPD	MedPay	Comp	Coll	Loan
Base Rate								
Super Good Driver Adj	x	x	x	x	x	x	x	
Safety Record	x	x	x	x	x	x	x	x
Driver Class	x	x	x	x	x	x	x	x
(1 - Recovery Device)						x		
Secondary Driver Characteristics	x	x	x	x	x	x	x	
Vehicle Type/Symbol	x	x	x	x	x	x	x	x
Vehicle Characteristics/[Model Year]Vehicle Age	x	x	x	x	x	x	x	x
Vehicle History [1]				x		x	x	x
Limit	x	x	x	x	x			x
Deductible [2]						x	x	
Stated [Value]Amount						x	x	
Claim Frequency	x	x	x	x	x	x	x	x
Claim Severity	x	x	x	x	x	x	x	x
(1 - Multi-Car Discount)	x	x	x	x	x	x	x	
(1 - Mature Driver Discount)	x	x	x	x	x	x	x	
Excess Vehicle Limit	x	x	x	x	x			x
Excess Vehicle Deductible						x	x	
Type of Vehicle Use	x	x	x	x	x	x	x	
Percent of Vehicle Use	x	x	x	x	x	x	x	
(1 - Multi-Policy Discount)	x	x	x	x	x	x	x	
Policy Term	x	x	x	x	x	x	x	
Annual Miles	x	x	x	x	x	x	x	
(1 - Good Driver Discount)	x	x	x	x	x	x	x	x
Round to nearest whole dollar								
Developed Premium [*][3]								

	CDW	ACPE	Roadside	Rental	Acq Exp
Base Rate					
Limit	x	x	x	x	
[Model Year]Vehicle Age			x	x	
Vehicle History [1]			x	x	
Claim Frequency			x	x	
Claim Severity			x	x	
(1 - Multi-Car Discount)			x	x	
(1 - Multi-Policy Discount)			x	x	
Policy Term	x	x	x	x	
(1 - Good Driver Discount)	x	x	x	x	x
Round to nearest whole dollar					
Developed Premium [*][3]					

[1] Vehicle History = (Vehicle History Damage Type 2 Factor / 100) * (Vehicle History Damage Type 1 Factor / 100) * (Vehicle History Theft Factor / 100) * (Vehicle History Title Issue Factor / 100) * (Vehicle History Interaction Factor / 100)

[2] Deductible = Deductible Factor * Deductible by Vehicle Age Factor * Deductible by BI Limit Facto

[*][3] Each selected coverage has a minimum premium of $1

Policy premium is determined by assigning the highest rated driver to the newest vehicle based on model year, second highest driver to second

- The highest rated driver is defined as the operator who has the highest combined BI Driver Class and Safety Record factor with consideration of the Good Driver and Super Good Driver Adjustment discounts.
- If two vehicles on the policy share the same model year, select the vehicle with the highest BI symbol factor. If the vehicles have the same BI symbol factor, the vehicles are assigned by the order of the drivers listed on the policy.
- If there are more vehicles than drivers, the additional vehicles will be rated using the excess vehicle rating factor.

Figure B5: Pricing algorithm

Notes: This is an example from an insurer's filing on how pricing rule work.Each column in the first panel represents different insurance coverage.

Rating Logic

Example Profile 1 - Zip code 90036 - Los Angeles

Rating information		Coverage Limits	
1 At Fault PD Accident		BI	15,000/30,000
15,000 Annual Mileage		PD	5,000
Licensed 2 Years		MP	2,000
Single Male		UMPD	3,500
2.4 GPA			
Commute to school 15 miles each way			
Vehicle: 2001 Ford Escort LX, 4 cyl automatic 4 dr sedan			

Item	BI	PD	MED	UMPD
Base Rate	233.78	289.45	25.63	23.69
Super Good Driver Adj	1.00	1.00	1.00	1.00
Driving Record Points Factor	1.45	1.38	1.33	1.24
Driver Class	1.52	1.54	1.72	1.51
Secondary Driver Characteristic	1.00	1.00	1.00	1.00
Annual Mileage Factor	1.43	1.32	1.01	0.94
Type of Vehicle Factor	0.98	0.84	1.38	0.64
Vehicle Characteristics - Mdl Yr	1.07	1.07	0.95	0.66
Excess Vehicle Limit Factor	1.00	1.00	1.00	N/A
Excess Vehicle Deductible Factor	N/A	N/A	N/A	1.00
Limit Factor	1.00	1.00	1.63	1.00
Deductible Factor	N/A	N/A	N/A	N/A
Claim Frequency Factor	1.48	1.37	1.28	1.34
Claim Severity Factor	0.92	0.92	0.83	1.04
Percent of Vehicle Use	1.00	1.00	1.00	1.00
(1 - Multi-car Discount)	1.00	1.00	1.00	1.00
(1 - Multi-policy Discount)	1.00	1.00	1.00	1.00
Policy Term Factor	0.50	0.50	0.50	0.50
(1 - Good Driver Discount)	1.00	1.00	1.00	1.00
Developed Premium *	$526	$460	$67	$12
Fixed Acquisition Expense Load	126.58			
(1 - Good Driver Discount)	1.00			
Fixed Acquisition Expense Load	$127			
Total (with CA Fraud Fee)	$1,192.88			

*Note: For each coverage, round to whole dollar after final computation.

Figure B6: Pricing example

Notes: This is an example from an insurer's filing which shows rating logic.

C Results

C.1 Tables

Table C1.Insurance gender gap among young people in California with different set of controls

Dependent variable: log(price)	(1) Young (California)	(2) Young (California)	(3) Young (California)	(4) Young (California)
Post * Male	-0.136** (0.063)	-0.177*** (0.064)	-0.177*** (0.064)	-0.212*** (0.051)
Demographic characteristics				✓
Location characteristics			✓	✓
Vehicle characteristics		✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	730	730	730	730
Adjusted R-squared	0.111	0.171	0.173	0.473

Notes: The logarithm of insurance price is regressed on male dummy and set of controls in each regression. Demographic controls are age, marital status, education level. Location characteristics are geographic region, population density, metropolitan statistical area dummy. Vehicle characteristics are vehicle type, vehicle year, vehicle brand, used status. The sample used in these regressions covers years between 2010 and 2020, only for young people under 25 in California. The heteroskedasticity-robust standard errors are reported in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table C2. The impact of the gender-neutral pricing policy in young drivers across states (all control states)

Dependent variable: log(price)	(1) Young (Control states)	(2) Young (Control states)	(3) Young (California)	(4) Young (All states)
Post * Male	-0.041	-0.049	-0.212***	
	(0.029)	(0.033)	(0.051)	
Post * Male * California				-0.180***
				(0.015)
Demographic characteristics		✓	✓	✓
Location characteristics		✓	✓	✓
Vehicle characteristics		✓	✓	✓
Year FE	✓	✓	✓	✓
State FE	✓	✓	✓	✓
State x Year FE	✓	✓	✓	✓
Observations	4,282	4,282	730	5,012
Adjusted R-squared	0.079	0.359	0.473	0.359
Empirical strategy	DID	DID	DID	Triple DID

Notes: The logarithm of insurance price is regressed on a male dummy and set of controls in Columns (1) to (3). Column (4) reports the coefficient for the triple interaction term in equation 3. Demographic controls are age, marital status, and education level. Location characteristics are geographic region, population density, and metropolitan statistical area dummy. Vehicle characteristics are vehicle type, year, brand, and used status. The sample used in these regressions covers the years between 2010 and 2020, only young people. Control states are all states that have gender-based insurance pricing. There are 41 control states included. Fixed effects for the year, state, and state*year are included in each regression. The heteroskedasticity-robust standard errors are clustered at the state level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table C3. List of consumer characteristics used in feature selection

Individual characteristics	**Vehicle characteristics**
Age	Car Brand
- Linear	- Dummy variables for 90 different car brands
- Six bins from 16 to 90*	Car Model
Education level	- Dummy variables for 661 different car models
- High school	Fuel type
- Bachelor	- Gas
- Graduate	- Diesel
Occupation	Vehicle type
- Sales	- Car
- Admin	- Van
- Manufacturing	- Sport utility vehicle
- Managerial	- Pickup
Homeownership	- Truck
- Ownership	Vehicle age
- Renting	- Linear
Number of vehicles in household	- 5-year bins from 0 to 40*
Number of drivers in household	Annual miles
	- Linear
Location characteristics	- Eight bins from 0 to 200,000 miles*
MSA category	Odometer reading
- Population with less than 1 million	- Linear
- Population with more than 1 million	- Eight bins from 0 to 500,000 miles*
MSA size	
- Population with less than 250,000	
- Population between 250,000 and 500,000	
- Population between 500,000 and 1 million	
- Population between 1 million and 3 million	
- Population with more than 3 million	
Urban area	
Population density per square mile	
Housing units per square mile	

Notes: This table lists all variables included in the prediction analysis for young male characteristics. It is based on National Household Travel Survey between 2001 and 2017. * denotes variables included in non-linear logit estimation only. The rest of the variables are included in lasso, ridge regression, elastic net, adaptive lasso, and linear logit estimations.

Table C4: Prediction of being young male

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. variable: Young male	Lasso	Elastic Net	Ridge	Adaptive Lasso	Logit I	Logit II
Middle school graduate	0.0767	0.0767	0.0784	0.0765	1.0320**	0.6731***
High school graduate	0.0061	0.0061	0.0071	0.0062		
Homeownership at HH level	0.0083	0.0084	0.0085	0.0128		
Metropolitan Statistical Area	0.0085	0.0085	0.0083	0.0086		
Car	0.0083	0.0083	0.0038	0.0127		
Pickup	0.0068	0.0068	0.0079	0.0133	1.0120***	
Truck	0.0272	0.0272	0.0309			
Ford	0.0031	0.0031	0.0030			1.0576***
Buick	0.0116	0.0116	0.0018			
Lexus	0.0090	0.0089	0.0172			
Victory		0.0973	0.1141			
Ford Escort/EXP/ZX2	0.0734	0.0734	0.1014	0.0745		
Ford Crown Victoria	0.0548	0.0549	0.0829	0.0555		
Ford Focus	0.0221	0.022	0.0494	0.0233		
Ford Windstar	0.0377	0.0376	0.0633			
Ford Ranger	0.0101	0.0101	0.0324			
Mercury Mountaineer	0.1043	0.1043	0.0753	0.1071		
Buick Regal (2011 on)	0.2470	0.2470	0.2777	0.2568		
Cadillac Seville	0.2169	0.2169	0.2770	0.2144		
Chevrolet Chevelle	0.0894	0.0894	0.0656			
Chevrolet Nova	0.0513	0.0513	0.0273			
Chevrolet Cobalt	0.1183	0.1182	0.0938	0.1167		
Chevrolet Geo Metro/Metro	0.0479	0.0479	0.0242			
Chevrolet Monte Carlo (1995 on)	0.0513	0.0514	0.0278			
Chevrolet Colorado	0.0351	0.0351	0.0055			
Pontiac Catalina/Parisienne	0.0503	0.0503	0.0696			
Pontiac Grand Prix (FWD)	0.0468	0.0468	0.0661			
GMC Canyon	0.0961	0.0961	0.0743	0.0916		
Saturn SC	0.4084	0.4084	0.4139	0.4069		
Saturn ION	0.0953	0.0953	0.1011	0.0934		
Jeep YJ series/Wrangler	0.0303	0.0303	0.0504	0.0325		1.0772**
BMW Z3	0.0417	0.0417	0.0702			

Notes: Demographic controls are age, marital status, education level. Column (5) and (6) refers to logit model with linear and non-linear version, respectively. The dataset used in this analysis is National Household Travel Survey 2001-2017. Only positive coefficents are reported in Column (1)-(4), only positive and statistically significant at ($p < 0.05$) results are reported in Column (5) and (6). For presentation purposes, a covariate is included into the table only if it is selected by at least two methods. Gas is if a vehicle is using gas as opposed to diesel. Car, pickup and truck are vehicle type classifications in the dataset. Homeownership status is at household level. Metroppolitan statistical area, middle school graduate and high schoold graduate are dummy variables taking value 0 or 1. Remaining variables represents each car model name as it is stated in NHTS dataset. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

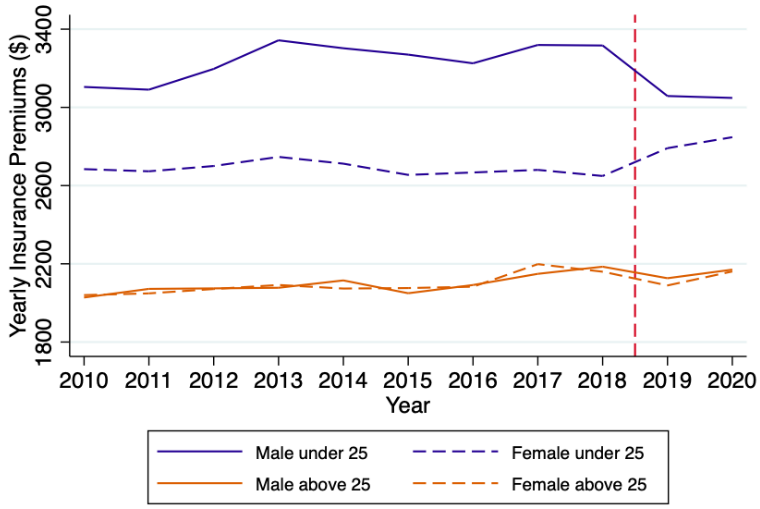
Table C4: Prediction of being young male (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. variable: Young male	Lasso	Elastic Net	Ridge	Adaptive Lasso	Logit I	Logit II
Nissan/Datsun Z-car, ZX	0.0876	0 .0876	0.1186	0.0853		
Nissan/Datsun 810/Maxima	0.0473	0.0473	0.0779	0.0451		
Nissan/Datsun 350Z/370Z	0.0739	0.0739	0.1045	0.0719		
Porsche 986/Boxster						1.0892**
Subaru Legacy/Outback	0.0365	0.0365	0.0565	0.0347		
Toyota Corolla	0.0118	0.0118	0.0171			
Toyota Celica	0.0419	0.0419	0.0474			
Toyota Camry	0.0092	0.0092	0.0142			
Toyota 4-Runner	0.0253	0.0253	0.0262	0.0278		
Toyota Tacoma	0.0094	0.0094	0.0093			
Volvo 240 series/DL/GL/GLT	0.1337	0.1337	0.1255	0.1316		
Volvo 70 Series (1998-2013)	0.0311	0.0311	0.0228			
Volvo 40 Series	0.0468	0.0468	0.1002			
Suzuki SX4/SX4 Crossover	0.1583	0.1583	0.1783	0.1579		
Suzuki Grand Vitara	0.1693	0.1693	0.1863	0.1714		
Acura Integra	0.0797	0.0797	0.0865	0.0778		
Acura Legend	0.1751	0.1752	0.1820	0.1731		
Acura TL	0.0230	0.023	0.0299			
Hyundai Azera	0.0909	0.0909	0.1175	0.0885		
Inifiniti M35/M37/M45/M56	0.2055	0.2055	0.2214	0.2034		
Lexus RX330/350/400h/450h					3.0702**	
Lexus NX					4.0566**	
Chrysler/Daimler Chrysler 200	0.0411	0.0412	0.0344			
Dodge Avenger (2008 on)						3.0780*
Dodge Dakota	0.0173	0.0174	0.0149			
Dodge Ram Pickup	0.0107	0.0107	0.0089			
Yamaha Motorcycle 125-349cc	0.0630	0.063	0.0624			
Yamaha Motorcycle (750cc)	0.0400	0.0401	0.0397			
Victory 750cc or greater	0.0973					
Other 450-749cc	0.0544	0.0545	0.0455			
Observations	674,853	674,853	674,853	674,853	571,382	453,740
R-squared	0.1455	0.1540	0.1541	0.1442	0.1370	0.1348
Out of Sample R-squared	0.1441	0.1415	0.1447	0.1447		
MSE	0.0171	0.0171	0.0170	0.0172		
Out of Sample MSE	0.0182	0.0183	0.0183	0.0182		
Mean Prediction Error	0.0174	0.0175	0.0176	0.0174		

Notes: Demographic controls are age, marital status, education level. Column (5) and (6) refers to logit model with linear and non-linear version, respectively. The dataset used in this analysis is National Household Travel Survey 2001-2017. Only positive coefficents are reported in Column (1)-(4), only positive and statistically significant at ($p < 0.05$) results are reported in Column (5) and (6). For presentation purposes, a covariate is included into the table only if it is selected by at least two methods. Gas is if a vehicle is using gas as opposed to diesel. Car, pickup and truck are vehicle type classifications in the dataset. Homeownership status is at household level. Metroppolitan statistical area, middle school graduate and high schoold graduate are dummy variables taking value 0 or 1. Remaining variables represents each car model name as it is stated in NHTS dataset. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

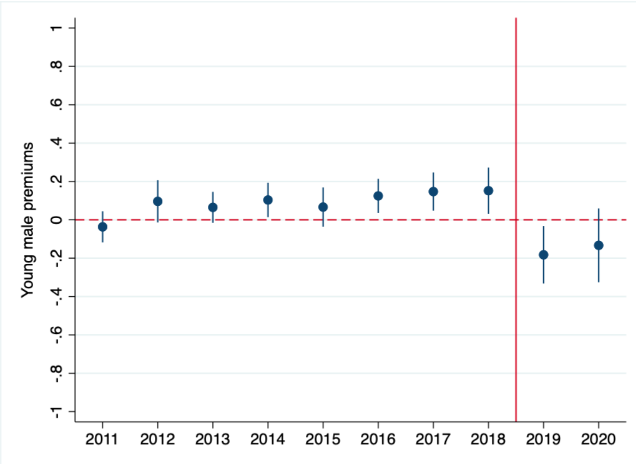
C.2 Figures

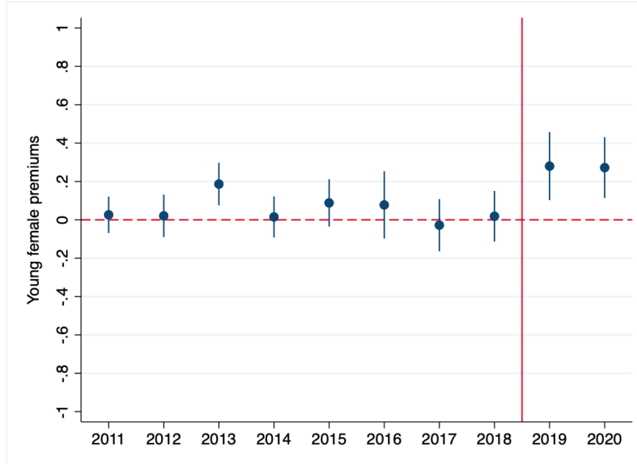
Figure C1. Premiums for young and old drivers by gender in California



Notes: This figure plots the raw trends in yearly insurance premiums paid by different groups in California. It uses the US Consumer Expenditure Survey dataset between 2010 and 2020.

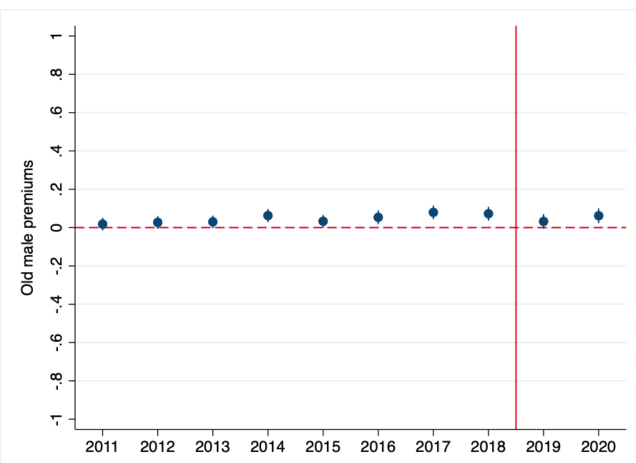
Figure C2. Insurance premiums for young males and females in California over time

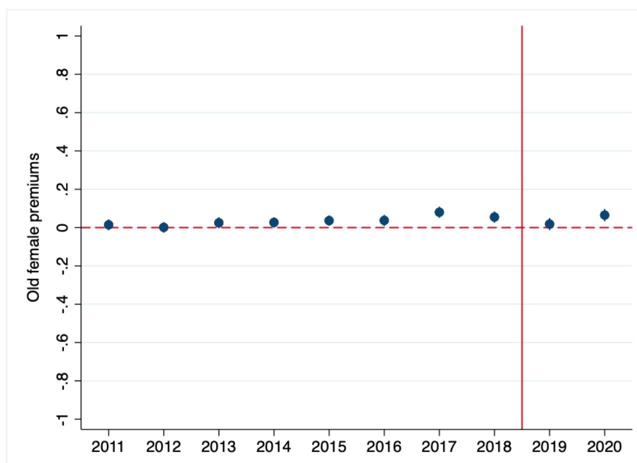




Notes: This figure presents the event study estimates for young males and females, respectively. Logarithm of insurance premiums is regressed on year dummies, demographic, location, vehicle and characteristics. Year coefficients are plotted for young males and young females by taking 2010 as the base year. It uses data from consumer expenditure data between 2010 and 2020. 90% confidence intervals are plotted. Heteroskedasticity-robust standard errors are used for each regression.

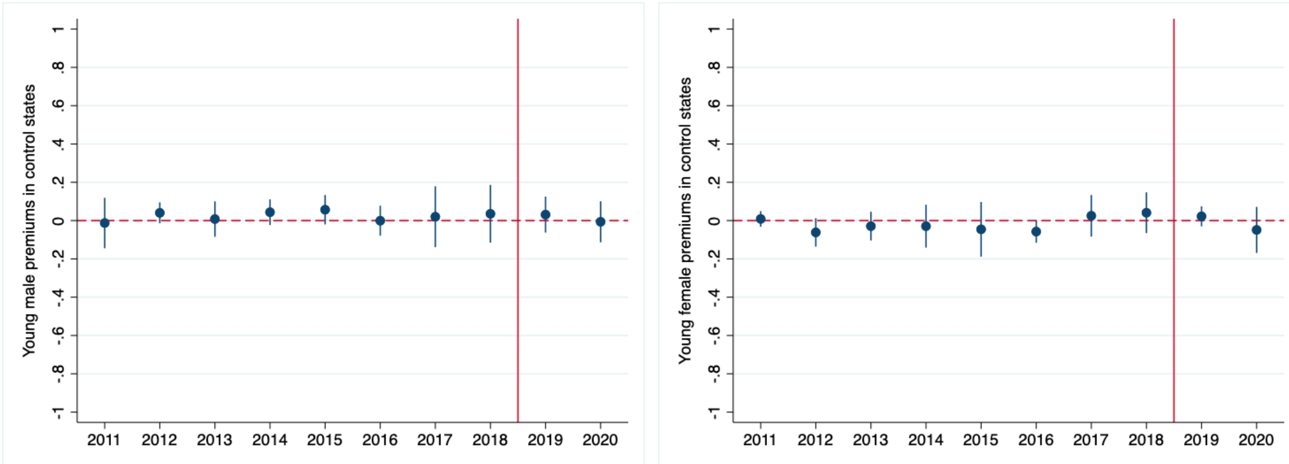
Figure C3. Insurance premiums for old males and females in California over time





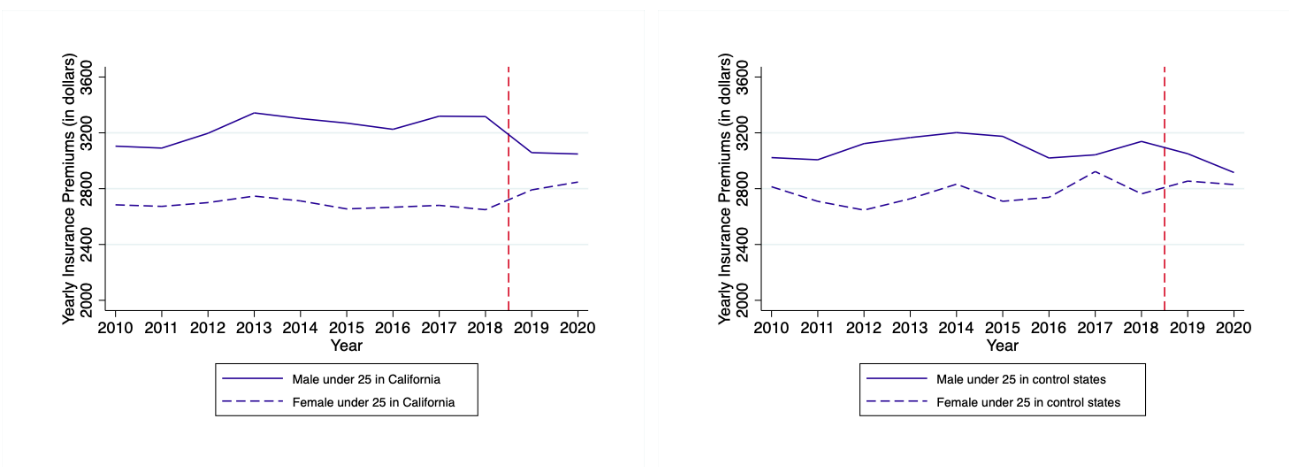
Notes: This figure presents the event study estimates for old males and females, respectively. Logarithm of insurance premiums is regressed on year dummies, demographic, location, vehicle and characteristics. Year coefficients are plotted for old males and old females by taking 2010 as the base year. It uses data from consumer expenditure data between 2010 and 2020. 90% confidence intervals are plotted. Heteroskedasticity-robust standard errors are used for each regression.

Figure C4. Insurance premiums for young males and females in control states over time



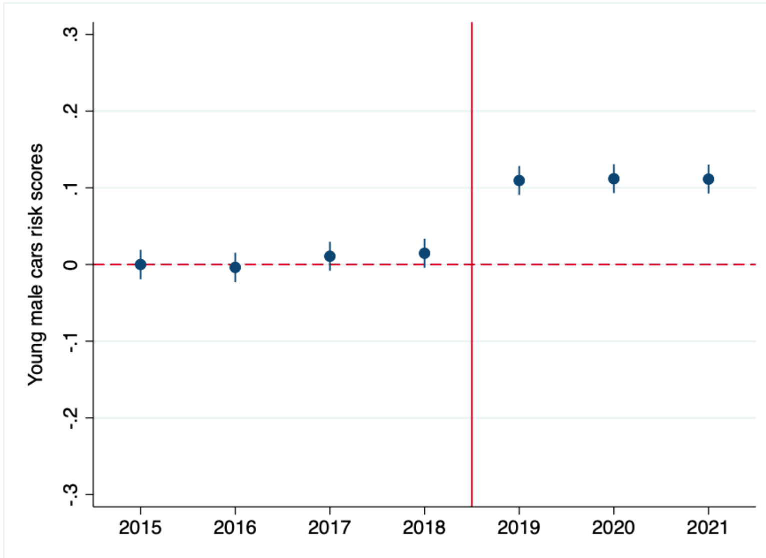
Notes: This figure presents the event study estimates for young males and females in control states. It uses data from consumer expenditure data between 2010 and 2020. The logarithm of insurance premiums is regressed on year dummies, demographic, location, vehicle, and characteristics. Fixed effects for the year, state, and state*year are included in each regression. Year coefficients are plotted for young males and females by taking 2010 as the base year. 90% confidence intervals are plotted. Heteroskedasticity-robust standard errors are clustered at the state level for each regression.

Figure C5. Premiums for young drivers by gender in California and control states



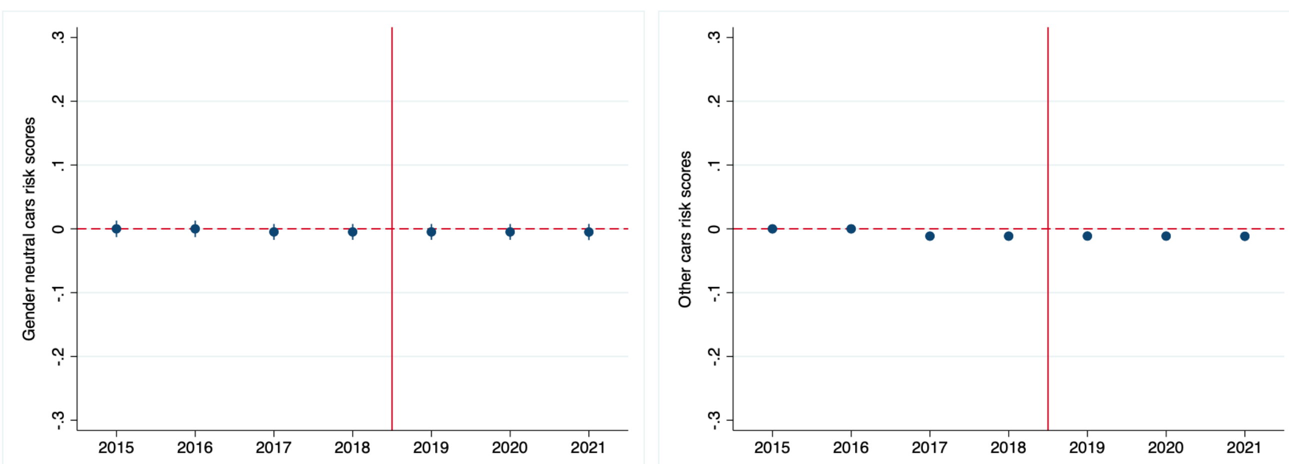
Notes: This figure plots the raw trends in yearly insurance premiums paid by young people under 25 in California and control states. It uses the US Consumer Expenditure Survey dataset between 2010 and 2020. Controls states included in this figure are all 36 states with gender-based insurance pricing.

Figure C6. Risk scores over time for young male cars



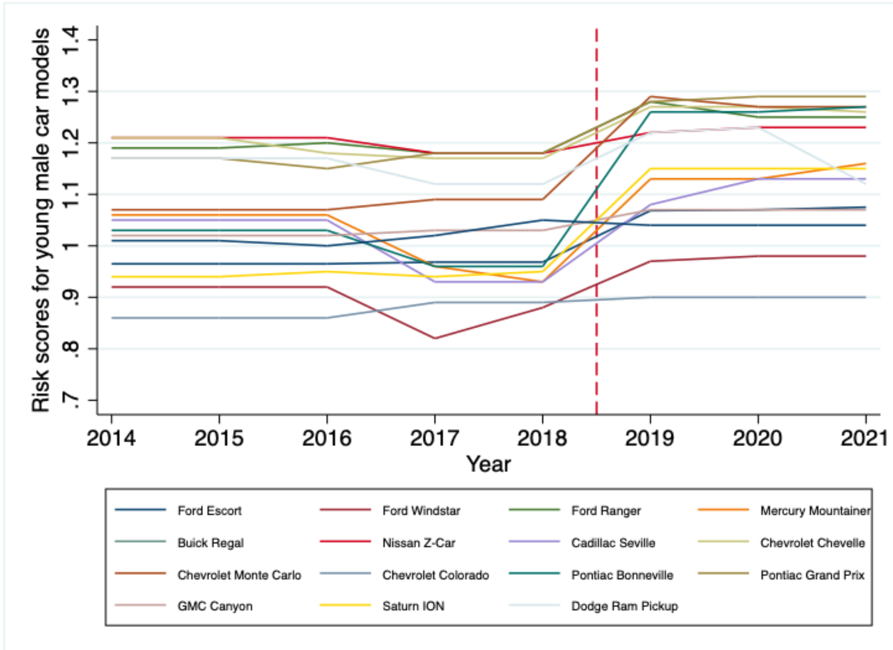
Notes: The analysis uses data from the pricing algorithm dataset compiled from insurance pricing filings of an insurance provider in California. The figure shows the event study estimates for young male cars only. The risk score for young male cars is regressed on year, car brand, and car model fixed effects for each regression. Year coefficients are plotted. It covers the years between 2014 and 2021. The base year 2014 is omitted in each figure. 95% confidence intervals are plotted. Heteroskedasticity-robust standard errors are used for each regression.

Figure C7. Risk scores over time for gender-neutral cars and other cars (control group)



Notes: The analysis uses data from the pricing algorithm dataset compiled from insurance pricing filings of an insurance provider in California. The figure shows the event study estimates for gender-neutral cars and other cars. Other cars are all the remaining car models that are not predicted by young male cars by the prediction analysis. The risk score for car models is regressed on year, car brand, and car model fixed effects for each regression. Year coefficients are plotted. It covers the years between 2014 and 2021. The base year 2014 is omitted in each figure. 95% confidence intervals are plotted. Heteroskedasticity-robust standard errors are used for each regression.

Figure C8. Risk scores over time for each young male car model



Notes: This figure shows time trends for risk scores of car models predicted as distinctively young male cars. It uses data from insurance filings of an insurance provider in California between 2014 and 2021.