Generative AI and Price Discrimination in the Housing Market

Jitsama Tanlamai

Desautels Faculty of Management, McGill University
jitsama.tanlamai@mail.mcgill.ca

Warut Khern-am-nuai

Desautels Faculty of Management, McGill University
warut.khern-am-nuai@mcgill.ca

Maxime C. Cohen

Desautels Faculty of Management, McGill University
maxime.cohen@mcgill.ca

Abstract

Housing discrimination has been recognized as an important societal issue for decades. While this issue can manifest in multiple ways, one of the most observed avenues is price discrimination, where houses in white-dominant neighborhoods are worth more than houses in minority-dominant neighborhoods that are otherwise similar. Prior studies have empirically documented such pricing discrimination and attributed it to human biases. In addition, recent studies have shown that issues of this kind are unlikely to be addressed by traditional AI models, even those specifically designed to address discrimination. In this paper, we first compare AI-generated versus human-generated housing prices using a sample of 284,749 U.S. properties. We then study the impact of generative AI in the context of price discrimination in the housing market and find that it can help alleviate this issue. Our mechanism exploration provides empirical evidence regarding underlying mechanisms that drive such a counter-intuitive result. Practical and policy implications are also discussed.

Keywords: generative AI, housing discrimination, price discrimination, housing market

1. Introduction

The issue of housing discrimination, which describes prejudicial treatment affecting minority groups' ability to acquire accommodation, is well documented in the literature (Ross and Turner 2005). Housing discrimination can manifest in various ways, such as discrepancies in mortgage approval rates (Quillian et al. 2020) and mortgage costs between applicants (Cheng et al. 2015). Discrimination in housing prices is another manifestation, and one that prompts serious concerns among individuals, firms, and policymakers (Ihlanfeldt and Mayock 2009; Schafer 2017). For instance, several studies have demonstrated that the price of a house in white-dominant neighborhoods is significantly higher than that of a similar house located in minority-dominant (e.g., Black or Hispanic) neighborhoods (Myers 2004). This issue is critical because studies have shown that minorities already have a higher cost to acquire a house (Apgar and Calder 2005; Foggo and Villasenor 2020). Taken together, these issues negatively contribute to the concern of fair housing practices and affordable housing and further affect the wealth inequality between groups of individuals (Kuebler 2013). More importantly, this issue has been further exacerbated with the rise in housing prices in the last decade. For example, on average, homes in neighborhoods with a majority of white residents are appraised at $371,000 more than similar homes in neighborhoods where most residents are people of color (Kamin 2022).

Similar to other bias and discrimination issues (e.g., Gunarathne et al. 2022; Tang et al. 2023), the issue of housing price discrimination is highly controversial. The common argument of opponents of this idea, which usually posit that housing price discrimination is a not an issue, is that white neighborhoods are more desirable than minority-dominant neighborhoods in certain aspects (e.g., education, safety), so that houses in white neighborhoods are naturally more expensive due to market forces (Guerrieri et al. 2013). Meanwhile, proponents of this idea argue that the differences in aspects such as education and safety between white- and minority-dominant neighborhoods are partially driven by housing price discrimination, and that interventions are required to level the playing field and ensure equality in the long run (Bayer et al. 2018). Our study does not aim to support or argue against these opposing views on housing price discrimination. Rather, we explore how the housing price discrimination issue is affected by the emergence of artificial intelligence (AI), specifically generative artificial intelligence (GenAI). Note also that our study does not aim to capture fundamental biases that are embedded in the generation process of housing prices as our study only compares biases present in housing prices generated by humans versus biases present in housing prices generated by GenAI.

With the recent advancement of AI, entities such as mortgage lenders and realtors are in the process of routinely adopting AI to assist their decision-making processes. As such, researchers have begun to investigate the role of AI in housing discrimination issues, especially regarding discrimination and fairness in AI prediction results (So et al. 2022). Sadly, several studies have empirically demonstrated that the use of AI in the housing context can amplify the bias historically observed in the market (e.g., Lambrecht and Tucker 2015). Moreover, AI models designed to address discrimination issues, typically called "fair AI," do not appear to address the housing discrimination issue in the short term (Zou and Khern-am-nuai 2023). In this study, we complement previous work in the field of AI and housing discrimination by examining how the use of GenAI, which is an emerging class of AI, affects the housing discrimination issue.

The GenAI market has grown drastically, and its value is expected to increase from $40 billion in 2022 to $1.3 trillion by 2032 (Bloomberg 2023; Chui et al. 2023). GenAI has many applications, with the most famous ones being generating text answers based on prompts (e.g., ChatGPT) and generating pictures based on input descriptions (e.g., Midjourney). For commercial purposes, GenAI has been used for several decision-support tasks (Dencik et al. 2023), and a potential application in this area is product or service pricing (Fowler et al. 2023; KPMG 2023). In this paper, we focus on the use case of leveraging GenAI to generate selling prices for houses. Specifically, we consider the situation where a real estate agent uses GenAI to set the selling prices of newly listed properties on a real estate aggregator platform. We first compare the AI-generated prices to human-generated ones and then examine how GenAI prices would affect the price discrimination issue in the housing market. Specifically, we propose the following research questions:

RQ1: Does the use of GenAI to set the selling prices of houses result in prices that are higher or lower than the human-generated prices?

RQ2: Do housing prices generated by GenAI amplify or alleviate price discrimination in the housing market?

The focal issue of our study is highly relevant to both research and practice. Our first research question relates to comparing human decisions to AI-generated decisions, which is a topic of growing interest in our society and highly relevant to the literature on pricing (Cohen et al. 2022; Gao et al. 2022). Our second research question aims to fill a gap in the existing literature on the impact of AI (e.g., Möhlmann et al. 2021; Nair and Saenz 2024; Wang et al. 2023b), specifically on fairness and AI biases. For instance, the literature has discussed potential biases that GenAI may induce due to its utilization of historical data (Fui-Hoon Nah et al. 2023; Sætra 2023). However, such arguments are usually developed conceptually, and empirical evidence is still limited. In addition, concerns regarding the existence of biases in the housing market, where prejudice is historically and consistently observed, can also be extended to the direction of bias where it exists. As such, identifying and quantifying existing biases is crucial from a policymaking perspective, given that humans are generally inclined to make judgments toward reference points (Kahneman et al. 1982). In that regard, the empirical evidence uncovered in our study helps resolve lingering questions regarding fairness and biases surrounding the use of GenAI.

To empirically study our proposed research questions, we collected the prices of 284,924 properties listed on a well-known U.S. real estate aggregator platform from August to September 2023. We aggregated the listings based on the characteristics of the houses and their locations at the ZIP code level, and we used their prices as the baseline compared to the price generated by GenAI for each listed house. We employed ChatGPT as the GenAI model due to its outstanding popularity and trustworthiness compared to the other GenAI models (Desk 2023; Liu et al. 2023). We used a prompt that asked ChatGPT to assume the role of a real estate agent to generate precise housing selling prices based on the information provided in the listings. We first compared the AI-generated prices to the human-generated ones and find that the prices generated by AI were 16% lower, on average. We then investigated our second research question by considering the discrimination factor that is commonly observed to impact the housing market: race in the neighborhood where each house is located. Our results suggested that while the housing prices generated by GenAI are not entirely discrimination-free, the price discrepancy gap is reduced by 15.6% when race is the discrimination factor. As such, the issue of price discrimination in the housing market is partially alleviated when prices are generated by GenAI, as opposed to some of the traditional AI models, which have been found to amplify the issue (Zou and Khern-am-nuai 2023). Furthermore, our mechanism explorations demonstrate that the results observed in our study are not solely driven by the central tendency behavior that is commonly observed in traditional AI models. Rather, the reduction in the pricing gap generated by GenAI appears to be driven by the model training process, which utilizes content available on the Internet. We also undertake a formal comparison between the prices generated by GenAI, traditional AI models, and humans in terms of housing price discrimination. Finally, we conduct several robustness tests with alternative specifications, different LLMs (GPT3.5, Google Gemini 1.5 Flash, and Claude 3.5 Sonnet), and alternative discrimination factors. The fact that our results are robust across four different LLMs strengthen the validity and generalizability of our findings.

Our findings contribute to both research and practice. First, this study is among the first to provide empirical evidence on how GenAI impacts price discrimination in the housing market, which is relevant to societal issues such as fair housing practices and affordable housing. Second, our findings contrast the belief in the literature regarding the potential biases that GenAI may demonstrate based on historical data. In that regard, our work provides a new perspective for future studies on the role of GenAI in fairness-related issues, which differs from that of traditional AI models. At the broader level, our study contributes to the emerging and growing literature at the intersection of AI and business management by providing insights that allow business managers to optimize and improve their decisions that have strong implications on firm performance (Dai and Tayur 2022; Terwiesch et al. 2020). From a practical perspective, our findings shed light on the benefits and concerns regarding the use of GenAI to set housing prices, which is likely to become a prevalent application of GenAI in the near future (Chui et al. 2023). Lastly, with the recent calls for meaningful AI regulations, our findings provide input for policymakers to help them develop AI regulations where social welfare is at stake.

The rest of the paper is organized as follows. In the next section, we review the prior literature relevant to our study. Following that, we describe our data and descriptive statistics in Section 3. Section 4 presents our empirical analyses and results. Finally, we discuss the implications of our work, along with its limitations and avenues for future research in Section 5.

2. Literature Review

In this section, we survey the literature related to our study. Specifically, we review prior works that have empirically studied housing discrimination. We then survey papers in information systems that discuss bias, discrimination, and fairness issues related to AI. Following that, we discuss technological or other factors that may contribute to bias mitigation. Lastly, we briefly review the current literature on generative AI.

2.1 Housing Discrimination

The issue of housing discrimination has been discussed in the literature for decades (Hanson and Hawley 2011). Classic literature, such as Becker (1957), discusses price discrimination from economic perspectives, providing theoretical foundations for subsequent studies that specifically focus on housing price discrimination (Yinger 1978). This discrimination can manifest in several ways, such as the behavior of real estate agents (Christensen and Timmins 2022), the behavior of landlords (Rosen et al. 2021), differences in the loan approval rates and mortgage costs (Quillian et al. 2020), and differences in housing prices (Ihlanfeldt and Mayock 2009; Schafer 2017). Discrimination factors include race (Christensen and Timmins 2022; Cui et al. 2020), gender (Flage 2018), and occupation (Auspurg et al. 2017). Housing discrimination is recognized as an important societal issue because it directly affects fair housing practices and the availability of affordable housing, which eventually leads to wealth inequality between different groups in our society. As such, this issue also attracts attention from legal scholars (e.g., deVise 1984; Schwemm and Taren 2010) and public policy entities (e.g., Neal et al. 2024; Neal et al. 2023).

While issues related to housing discrimination are widely discussed, both in the literature and in practice, they have recently been in the spotlight due to the increasing popularity of AI, which has been used by several organizations to perform tasks related to housing discrimination. For example, Rosen et al. (2021) argue that the use of AI for rental application screening among landlords with large portfolios can contribute to the gap in rental costs and rental approval rates between racial majority and minority applicants. In the same vein, Zou and Khern-am-nuai (2023) argue that the use of AI in mortgage approval can amplify the bias that historically exists in the market, and the use of fair AI models, if done blindly, may not yield benefits that the models are designed to provide.

In this study, we investigate the intersection of housing price discrimination and GenAI. In that regard, we next survey the literature on AI and issues of discrimination and fairness.

2.2 AI and Fairness Issues

The issues regarding AI and fairness have been discussed extensively in the computer science literature over the last decade. See Richardson and Gilbert (2021) for a systematic review of this important topic. In this subsection, we focus on prior studies in business and management related to AI and fairness.

The pioneering works in this area typically discuss issues related to algorithmic bias. For example, Lee and Hosanagar (2019) utilize a randomized field experiment to show that a recommendation algorithm commonly used in the market tends to create the "rich-get-richer" concentration bias. Similarly, the simulations in Zhou et al. (2023) demonstrate that preference bias in recommendation algorithms, which occurs because the algorithms are trained on preference ratings that users generated in the previous iterations, has significant effects on both the systems and the users. On the system side, preference bias can impair the prediction accuracy of the recommendations. Meanwhile, users' consumption outcomes, including consumption relevance and consumption diversity, can be negatively impacted by preference bias. When these algorithms are used and biases are generated, users' fairness perceptions and technology-related behavior, including technology acceptance, appreciation, and adoption, can be significantly affected (Kordzadeh and Ghasemaghaei 2022).

Since fairness-related issues in AI have become a heated discussion topic, several recent papers propose models or approaches to address this issue. For instance, Ahsen et al. (2019) developed a bias-aware classification AI model that can eliminate the negative impacts of bias in breast cancer diagnosis. Relatedly, Shimao et al. (2025) propose a framework that can be used to study the fairness properties of AI models in both prediction results and behavioral differences between majority and minority prediction subjects. Nevertheless, several studies have shown that these "fair" AI models can negatively impact welfare metrics for the relevant stakeholders and society in general if they are used inappropriately (e.g., Fu et al. 2022; Rhue 2024).

The topic of fairness in the context of pricing discrimination has also been studied in the literature (Cohen et al. 2022; Cohen et al. 2021). One of the studies in this research stream that is closely relevant to our work is Fu et al. (2023). The authors examine the use of an algorithm called Zestimate, which is an AI model that consumes historical housing data to predict the market value of houses, on Zillow, which is one of the most popular U.S. real estate aggregator platforms. The key difference between this work and our paper is that Zestimate relies on a traditional AI model, which is known to be prone to bias inheritance, especially in the housing market (Zou and Khern-am-nuai 2023). In our paper, however, we study the bias implications of GenAI models, an aspect that is currently underexplored in the literature. Our work aims to provide empirical evidence regarding whether GenAI and traditional AI models are the same in terms of inheriting biases from the training data. Next, we survey the literature that discusses technological or other factors that may contribute to bias mitigation in Section 2.3. Then, in Section 2.4, we briefly review the literature on GenAI and how it is different from traditional AI models.

2.3 Bias Mitigation

In the previous subsection, several studies have theoretically and empirically demonstrated bias, discrimination, and fairness issues driven by AI. In response, a stream of research is dedicated to identifying factors that may contribute to bias mitigation.

The factor that receives the most attention in bias mitigation is technology. In that regard, Shimao et al. (2025) discuss several fairness constraints that are commonly used to mitigate bias in AI predictions and how different constraints affect the behavior of prediction subjects. Relatedly, Fu et al. (2021) propose a fair AI model that mitigate bias in prediction by removing the correlation between the sensitive variable and the predictors and show that although the de-bias model suffers from a lower accuracy, it still improves the decision quality in the context of crowd lending. Several other studies have also proposed and discussed the use of technological factors to mitigate bias in AI predictions, such as the design principle of bias-free AI models, the evaluation metrics specifically developed for bias mitigation, and an approach to automatically detect and mitigate potential biases (Abbasi et al. 2024; Guo et al. 2022; Lalor et al. 2024).

Other than technological factors, several alternative factors have been explored. For instance, one of the factors that have gained traction in recent years is the human factor. In this regard, Rhue (2024) demonstrates how human emotions can be leveraged for bias mitigation. Other strategies that are proven to be effective in mitigating bias in human-AI collaboration include time allocation (Rastogi et al. 2022), promoting transparency (Ferrara 2023), and improving business processes (Roselli et al. 2019). In the meantime, legal and regulation factors are also explored, given that several countries have initiated the discussion on legal frameworks to govern AI to ensure fairness (Nazer et al. 2023). However, the effectiveness of legal and regulatory frameworks in mitigating bias so far was found to be limited (Bai et al. 2022). Lastly, a few recent studies have discussed an emerging factor that relies on education to help mitigate AI-induced biases through practices such as knowledge sharing (Soleimani et al. 2022) and raising awareness (Wang et al. 2023a).

2.4 Generative AI

The key distinctive feature of GenAI models is their ability to generate new outputs that aim to be similar to the data they observed in the training process (Banh and Strobel 2023). This process is vastly different from traditional AI models, which focus on learning the relationship between the input variables and the output. GenAI, on the other hand, learns the underlying data structure and generation processes. As such, GenAI is generally considered a new class of AI models, and it is unclear whether the properties of traditional AI models apply to GenAI models (Ooi et al. 2023). We refer to Susarla et al. (2023) for more details about GenAI models and their impact on the business and management fields.

Several recent studies have shown that GenAI can create new opportunities in various fields, including education (Van Slyke et al. 2023), creative tasks (Benbya et al. 2024), and knowledge management (Alavi et al. 2024). Indeed, GenAI allows for human-AI augmented intelligence tasks, which could significantly enhance workforce productivity (Eloundou et al. 2023; Lou et al. 2023; Noy and Zhang 2023). At the same time, this technology, if used irresponsibly, can be a severe threat to society (Sabherwal and Grover 2024), particularly because of potential behavioral biases that GenAI can produce (Chen et al. 2023). As such, several ethical concerns have been raised regarding the use of GenAI in business (Davison et al. 2024).

Our study complements the prior literature on the societal impact of GenAI by empirically examining how the use of GenAI to generate housing prices, an application that is commonly adopted, impacts the price discrepancy that is commonly observed in the market. In that regard, our study is at the crossroads of GenAI, fairness issues in AI, and housing discrimination issues, all of which are important and relevant to society. We address this topic by leveraging observational data collected from a well-known U.S. real estate aggregator platform and data generated by a widely used GenAI model. We describe our data in the next section.

3. Research Context and Data

3.1 Generative AI and Real Estate Agents

GenAI has increasingly become an essential tool in the real estate industry (Abouzakhar 2024). More specifically, GenAI is now routinely used by many real estate agents to improve their efficiency in various tasks, including how agents market properties, interact with clients, and streamline operations (Fitzpatrick et al. 2023). For example, GenAI tools can be used to generate property descriptions, promotional materials, and personalized emails at scale (Dudhat 2024). Among these features, the function that inspired our research is the use of GenAI for recommending prices for newly listed properties, an application of GenAI that is becoming common in practice (e.g., Finno 2023; Lanoue 2024) but is underexplored in the literature.

To understand how real estate agents use GenAI for pricing purposes, we engaged in a conversation with a prominent real estate agency in North America. We have leveraged this collaboration to learn the factors that impact real estate agents' pricing decisions. These factors include zip code (aka "location location location"), city, state, housing age, number of bedrooms, number of bathrooms, house size, garage space, property type, school rating, flood risk, wildfire risk, heat risk, wind risk, air pollution risk, and whether it was recently renovated. We have also learned that real estate brokers have increasingly relied on GenAI tools to help set the listed price of the properties they handle. The key attractiveness of GenAI tools (as opposed to traditional AI models) is the ease of use and trust in the technology since the vast majority of agents already use GenAI tools in their daily lives. As such, the use of GenAI to generate housing prices is often viewed as a data-driven approach that leverages historical data and state-of-the-art technologies. Generally, the real estate agents feed the details of the properties, including the property and neighborhood characteristics listed above, to the GenAI tool and ask the tool to estimate (or recommend) the price of the property. The generated prices are then reviewed and adjusted as necessary. Our conversation revealed that many real estate agents view housing prices generated by GenAI tools as being reasonably accurate (i.e., in line with the prices that they would have generated themselves), except for extreme outlier cases. As such, real estate agents tend to view GenAI as a productivity tool that allows them to price tens of properties in a single click with limited technical skills. Our separate conversations with several independent real estate agents also revealed that they tend to rely heavily on GenAI to price properties in areas that they are not familiar with. Next, we discuss the data sources and how we utilize GenAI to generate housing prices in our study.

3.2 Data Sources

Our study utilizes three data sources. First, we collected house listings from a leading real estate aggregator website in the United States between August and September 2023. We considered listings for five main property types: single-family houses, condos, multi-family houses, townhomes, and duplexes/triplexes. For each listing, we collected the price, the location at the ZIP code level, and the characteristics of the property, including its age (in years), the number of bedrooms, the number of bathrooms, its size (in square feet), and the amount of garage space (i.e., the number of cars that can park at the property). We note that the prices collected here are the prices set by the selling party (e.g., homeowners or real estate agents representing the sellers) and not the final selling price of the properties. We note also that the majority of the listings is handled by real estate agents. We also collected each listing's description provided and employed keyword detection to identify renovated properties. The used keywords are "renovate," "repair," "makeover," "upgrade," "revamp," "overhaul," "redesign," "remodel." Furthermore, we gathered each listing's neighborhood information based on its geographic coordination. The information includes the average rating of the nearby schools and environmental risk factors, including flood, wildfire, heat, wind, and air pollution. These scores range from 1 to 10, with the higher score presenting a better school rating or greater potential for the specified hazard. After the data collection, we cleaned the dataset by excluding records with missing information, extreme numbers of bedrooms or bathrooms (i.e., 10 or more), or extreme prices (e.g., less than $1,000). In total, our dataset consisted of 284,924 records. The listing prices were treated as the baseline in our empirical analyses. Second, we retrieved the area demographics from the U.S. Census Bureau at the ZIP code level. For each ZIP code, we collected the population, the percentage of white residents, and the mean household income. We dropped the ZIP codes with missing information and merged the collected data with the listing dataset previously described. After this step, 284,749 records remained. Lastly, we used GenAI to generate an estimated price for each house in our dataset. Here, we used GPT-4o-2024-08-06 as the primary GenAI model and relied on the application program interface (API) provided by OpenAI to collect the data. The Generative pre-trained transformer (GPT) model is a pre-trained large-scale language model (LLM) that allows users to ask questions on various topics. The model utilizes extensive training data to generate an original answer for each inquiry. The model version used in this study is trained using data publicly available on the Internet, data licensed by third parties, and input from OpenAI workers as of October 2023 (OpenAI 2024b; Schade 2023). We chose GPT-4o as the primary GenAI for our analysis because it was the most advanced GPT model as of the data collection date (October 2024) and it is claimed to be the most intelligent GenAI in the market across several tasks at that time (OpenAI 2024a).

Our GenAI prompts assumed the role of a realtor who estimates a house price based on the house type, location, characteristics, and neighborhood information. Details of the prompt are available in Section A of the online appendix. The key analysis of our study compared the house prices generated by GenAI to the human-generated prices obtained from the real estate aggregator website for houses in the advantaged group and houses in the disadvantaged group. By comparing AI- and human-generated prices, we could examine the effect of GenAI on housing price discrimination. We highlight that our target variable is the asking price (and not the closing price), listed by realtors or homeowners when selling a property. We also conduct an analysis on the variation of human-generated prices and GPT-generated prices and provide the results in Section B.1 of the online appendix.

Table 1: Data description (1)

N = 284,749		Mean, Median (Std)
Price	Human-generated price	553,453.85, 379,500.00 (879,317.65)
	GPT-generated price	425,577.21, 325,000.00 (444,222.56)
Property characteristics	Number of beds	3.27, 3.00 (1.03)
	Number of bathrooms	2.27, 2.00 (0.93)
	Size (sqft.)	2,025.87, 1,773.00 (1218.49)
	House age (year)	46.44, 41.00 (33.41)
	Garage space	1.56, 2.00 (6.61)
Neighborhood information	Average school rating	5.22, 5.14 (1.50)
	Flood risk	2.00, 1.00 (2.21)
	Wildfire risk	2.71, 2.00 (1.99)
	Heat risk	5.34, 5.00 (2.53)
	Wind risk	3.87, 2.00 (2.97)
	Air pollution risk	3.18, 3.00 (2.13)
Property location	Total population	30,034.18, 27,815.00 (19,481.71)
	Percentage of white residents	0.71, 0.76 (0.21)

Table 2: Data description (2)

N = 284,749		Number of Observations, Percentage
Property characteristics	Renovated house:	
	No	221,147, 77.66%
	Yes	63,602, 22.34%
	Property Type:	
	Condos	27,292, 9.58%
	Duplex/Triplex	213, 0.07%
	Multi-family	5,958, 2.09%
	Single-family	233,213, 81.90%
	Townhomes	18,073, 6.35%
	Market position (by median age):	
	Old property	140,122, 49.21%
	New property	144,627, 50.79%
	Market position (by median supply):	
	Low-supply area	26,948, 9.46%
	High-supply area	257,801, 90.54%
Property locations	Location characteristics:	
	White neighborhoods	239,595, 84.14%
	None-white neighborhoods	45,154, 15.86%

To facilitate our analyses, we categorized properties in our dataset according to their market position by splitting properties in our dataset based on the median value of the number of houses in the ZIP code and the house ages. The former variable captures the supply of houses in the market and the latter captures the maturity of the properties, both of which have strong implications in the literature (e.g., Bellamy et al. 2020; Wang et al. 2022). For the discrimination factor, we focused on the percentage of white residents in the neighborhood (i.e., we label the neighborhood white-dominant if this value exceeds 50%). Details of our data description are provided in Tables 1 and 2.

4. Empirical Analyses and Results

We utilized fixed-effect linear regression models to examine the implications of using GenAI to generate housing prices. Recall that we use GenAI to generate housing prices based on the type, location, and characteristics of each house. We begin by examining the nature of GenAI pricing versus the actual prices on the real estate aggregator platform and the nature of price discrimination in the housing market.

4.1 GenAI Pricing vs. Human Pricing and Price Discrimination

The specification of our first regression model is presented in Equation (1):

$$\log(Price_{ijk}) = \alpha AI_{ijk} + \beta_1 MarketHighSupply_{ijk} + \beta_2 MarketNewHouse_{ijk} \tag{1}$$
$$+ \beta_3 Underprivileged_{ijk} + \boldsymbol{\varpi C_{ijk}} + \theta_j + \zeta_k + \varepsilon_{ijk},$$

where $Price_{ijk}$ is the price of property i, which has type j at k location (at the city-state level). AI_{ijk} is the dummy variable that takes the value 1 if $Price_{ijk}$ is generated by GenAI and 0 if it is the actual price listed on the real estate aggregator platform. $MarketHighSupply_{ijk}$ and $MarketNewHouse_{ijk}$ are the dummy variables for market positions, indicating whether the listing falls into the high-supply housing or new houses, respectively. $Underprivileged_{ijk}$ is a binary variable, representing the discrimination factor, indicating whether the listing is in a minority-dominant area (i.e., an underprivileged group). $\boldsymbol{C_{ijk}}$ is the set of control covariates, including the number of beds (Bed_{ijk}), the number of bathrooms ($Bath_{ijk}$), the size ($Size_{ijk}$), the garage space ($Garage_{ijk}$), the indication of being renovated ($Renovate_{ijk}$), the average rating of nearby schools ($School_{ijk}$), the flood risk score ($Flood_{ijk}$), the wildfire risk score ($Fire_{ijk}$), the heat risk score ($Heat_{ijk}$), the wind risk score ($Wind_{ijk}$), the air pollution risk score (Air_{ijk}), and the population in the corresponding ZIP code ($Population_{ijk}$). The terms θ_j and ζ_k represent the house type and location fixed effects, respectively, which control for house-type-specific and location-specific characteristics. ε_{ijk} represents the error term. For numerical variables that are power law distributed, we transform them by applying the natural logarithm transformation with log(1+x) where x is the transformed variable to ensure that x can take the value zero.

Table 1: Gen AI vs. human pricing and price discrimination

	$\log(Price_{ijk})$
AI_{ijk} baseline: human-generated prices	-0.135*** (0.0007)
$MarketHighSupply_{ijk}$ baseline: low-supply areas	0.055*** (0.006)
$MarketNewHouse_{ijk}$ baseline: old houses	0.061*** (0.001)
$Underprivileged_{ijk}$ baseline: privileged areas	-0.186*** (0.002)
Control variables	$\log(Bed_{ijk}), \log(Bath_{ijk}), \log(Size_{ijk}), \log(Garage_{ijk})$ $Renovate_{ijk}, \log(School_{ijk}), \log(Flood_{ijk}), \log(Fire_{ijk})$ $\log(Heat_{ijk}), \log(Wind_{ijk}), \log(Air_{ijk}), \log(Population_{ijk})$
Fixed effects	House type, Location
N	569,498
R^2	0.8647

Note. Standard errors in parentheses are robust and clustered by property. * < 0.1, ** < 0.05, *** < 0.01.

Table 3 reports the results from this regression analysis. We found that housing prices generated by GenAI are lower than the actual listed prices by an average of 13.5%. In addition, we found that houses in the high-supply areas are about 5.5% more expensive than houses in the low-supply areas that have similar type, characteristics, neighborhood, and location. New houses are about 6.1% more expensive than old houses with similar type, characteristics, neighborhood, and location. Furthermore, consistent with findings in the literature, our analysis confirmed that houses in white-dominant neighborhoods are significantly more expensive than those with similar type, characteristics, neighborhood, and location that are located in minorities-dominant neighborhoods (18.6% more expensive, on average).

Thus, the answer to our first research question is that GenAI generates housing prices that are 13.5% lower than human-generate prices, on average. The natural question that follows is whether this finding applies to all properties. To examine this, we performed an additional analysis in which we separated the data into four segments based on the property price quartile listings fall into. For each group, we re-ran our analysis; the results are reported in Table 4.

Table 4: Gen AI vs. human pricing and price discrimination (grouped by housing prices)

	$\log(Price_{ijk})$ 1st quartile	$\log(Price_{ijk})$ 2nd quartile	$\log(Price_{ijk})$ 3rd quartile	$\log(Price_{ijk})$ 4th quartile
AI_{ijk} baseline: human-generated prices	0.107*** (0.001)	-0.089*** (0.001)	-0.180*** (0.001)	-0.382*** (0.001)
$MarketHighSupply_{ijk}$ baseline: low-supply areas	-0.0007 (0.012)	0.017*** (0.005)	0.036*** (0.005)	0.095*** (0.009)
$MarketNewHouse_{ijk}$ baseline: old houses	0.099*** (0.002)	0.046*** (0.0008)	0.025*** (0.001)	-0.031*** (0.002)
$Underprivileged_{ijk}$ baseline: privileged areas	-0.190*** (0.004)	-0.072*** (0.002)	-0.061*** (0.002)	-0.126*** (0.005)
Control variables	$\log(Bed_{ijk}), \log(Bath_{ijk}), \log(Size_{ijk}), \log(Garage_{ijk})$ $Renovate_{ijk}, \log(School_{ijk}), \log(Flood_{ijk}), \log(Fire_{ijk})$ $\log(Heat_{ijk}), \log(Wind_{ijk}), \log(Air_{ijk}), \log(Population_{ijk})$			
Fixed effects	House type, Location			
N	144,262	140,494	143,336	141,406
R^2	0.6084	0.5681	0.6217	0.7813

Note. Standard errors in parentheses are robust and clustered by property. * < 0.1, ** < 0.05, *** < 0.01.

Interestingly, we find that GenAI does not always price houses lower than humans do. It turns out that GenAI prices for cheaper houses are higher than those determined by humans (10.7% higher for houses in the first quartile), but the effect changes when houses become more expensive (8.9% lower for houses in the second quartile, 18% lower for houses in the third quartile, and 38.2% lower for houses in the fourth quartile). Meanwhile, the price discrimination effect in the market (e.g., houses in white-dominant areas are more expensive than similar houses in minority-dominant areas) is consistent for all segments. Next, we investigated the impact of GenAI pricing on housing price discrimination.

4.2. Implications of GenAI Pricing on Housing Price Discrimination

Next, we shifted our focus to our second proposed research question. Particularly, we investigated whether housing prices generated by GenAI can help alleviate the pricing discrimination phenomenon observed earlier. For this analysis, we added interaction terms between AI_{ijk} and the dummy variable that represent the discrimination factor (i.e., $Underprivileged_{ijk}$) to the regression specification in Equation (1). The new model specification is presented in Equation (2).

$$\log(Price_{ijk}) = \alpha AI_{ijk} + \beta_1 MarketHighSupply_{ijk} + \beta_2 MarketNewHouse_{ijk} + \beta_3 Underprivileged_{ijk} + \beta_4 (AI_{ijk} \times Underprivileged_{ijk}) + \boldsymbol{\varpi C_{ijk}} + \theta_j + \zeta_k + \varepsilon_{ijk}. \quad (2)$$

The results from this specification are reported in Table 5. We observed that ChatGPT continued to generate housing prices that are, on average, lower than human-generated prices. Similarly, the coefficients of market position and discrimination factors are consistent with those reported in Table 3.

Table 5: Impact of GenAI pricing on housing price discrimination

	$\log(Price_{ijk})$
AI_{ijk} baseline: human-generated prices	-0.149*** (0.001)
$MarketHighSupply_{ijk}$ baseline: low-supply areas	0.055*** (0.006)
$MarketNewHouse_{ijk}$ baseline: old houses	0.061*** (0.001)
$Underprivileged_{ijk}$ baseline: privileged areas	-0.229*** (0.003)
$AI_{ijk} \times Underprivileged_{ijk}$	0.087*** (0.002)
Control variables	$\log(Bed_{ijk}), \log(Bath_{ijk}), \log(Size_{ijk}), \log(Garage_{ijk})$ $Renovate_{ijk}, \log(School_{ijk}), \log(Flood_{ijk}), \log(Fire_{ijk})$ $\log(Heat_{ijk}), \log(Wind_{ijk}), \log(Air_{ijk}), \log(Population_{ijk})$
Fixed effects	House type, Location
N	569,498
R^2	0.8652

Note. Standard errors in parentheses are robust and clustered by property. * < 0.1, ** < 0.05, *** < 0.01.

As for our main results, we found that the coefficient of $AI_{ijk} \times Underprivileged_{ijk}$ is positive and statistically significant, indicating that GenAI tends to generate higher prices in minority-dominant neighborhoods relative to human-generated prices. To interpret the coefficient of interest, recall that the baseline for our analysis is the human-generated price for houses in white-dominated areas. We found that human-generated prices when houses are in minority-dominated areas are 22.9% lower. At the same time, GenAI-generated prices are 14.9% lower for houses in white-dominated areas and 29.1% lower when the houses are in minority-dominated areas. As a result, while the prices of houses located in minority-dominated areas are lower by 22.9% when humans generate the price, the reduction is only 14.2% when GenAI generates the price. Therefore, we conclude that GenAI reduces price discrimination, on average, by 8.7% in minority-dominant neighborhoods.

Table 6: Impact of GenAI pricing on housing price discrimination

	$\log(Price_{ijk})$ 1st quartile	$\log(Price_{ijk})$ 2nd quartile	$\log(Price_{ijk})$ 3rd quartile	$\log(Price_{ijk})$ 4th quartile
AI_{ijk} baseline: human-generated prices	0.100*** (0.002)	-0.090*** (0.001)	-0.185*** (0.001)	-0.408*** (0.002)
$MarketHighSupply_{ijk}$ baseline: low-supply areas	-0.001 (0.012)	0.017*** (0.005)	0.036*** (0.005)	0.095*** (0.009)
$MarketNewHouse_{ijk}$ baseline: old houses	0.099*** (0.002)	0.046*** (0.0008)	0.025*** (0.001)	-0.031*** (0.002)
$Underprivileged_{ijk}$ baseline: privileged areas	-0.208*** (0.005)	-0.077*** (0.002)	-0.080*** (0.002)	-0.214*** (0.005)
$AI_{ijk} \times Underprivileged_{ijk}$	0.037*** (0.004)	0.010*** (0.002)	0.038*** (0.002)	0.175*** (0.003)
Control variables	$\log(Bed_{ijk}), \log(Bath_{ijk}), \log(Size_{ijk}), \log(Garage_{ijk})$ $Renovate_{ijk}, \log(School_{ijk}), \log(Flood_{ijk}), \log(Fire_{ijk})$ $\log(Heat_{ijk}), \log(Wind_{ijk}), \log(Air_{ijk}), \log(Population_{ijk})$			
Fixed effects	House type, Location			
N	144,262	140,494	143,336	141,406
R^2	0.6087	0.5681	0.6226	0.7843

Note. Standard errors in parentheses are robust and clustered by property. * < 0.1, ** < 0.05, *** < 0.01.

Next, to examine the heterogeneity of our findings, we separated the data into four groups comprising listings in the first, second, third, and fourth quartiles of the property price. The results are reported in Table 6. Overall, our heterogeneity analysis yields results that are consistent with our main analysis. Namely, houses in white-dominant areas are significantly more expensive than similar houses in minority-dominant areas. At the same time, GenAI can alleviate housing price discrimination, regardless of the price groups. Next, we explored potential mechanisms that drive GenAI pricing in the housing market.

To showcase the robustness of our results, we conducted several robustness tests with alternative specifications, different LLMs (GPT3.5, Google Gemini 1.5 Flash, and Claude 3.5 Sonnet), and alternative discrimination factors. The results of tests are reported in Section C of the online appendix.

4.3 Exploring Housing Price-Generating Mechanisms

Our results empirically demonstrate that, unlike the case of traditional AI models, housing prices generated by GenAI reduce price gaps based on discrimination factors, such as the price between houses in white-dominant neighborhoods and those that are similar but located in minority-dominant neighborhoods. This finding raises a question regarding the mechanisms that lead GenAI to generate housing prices that reduce such housing price discrimination. In this section, we attempt to discern the underlying mechanism that leads GenAI to generate counter-intuitive housing prices observed in our study.

4.3.1 Central Tendency

The first potential mechanism that emerges in our context originates from the classic bias vs. variance trade-off in traditional AI models (Larose 2015). Many models, especially those that utilize simple supervised machine learning techniques, tend to avoid the overfitting issue by lowering the variance at the cost of increasing the bias. As a result, these models tend to predict outcomes that are close to the mean, a behavior often referred to as central tendency (Qayyum et al. 2023). It is worth noting that the trade-off between bias and variance is largely the model developer's choice. Since the development of GenAI is mostly black box, we take an exploratory approach to examine whether the reduced pricing discrimination observed from our main results could simply be attributed to the central tendency behavior (i.e., GenAI reduces price discrimination because it tends to generate prices that are closer to the mean).

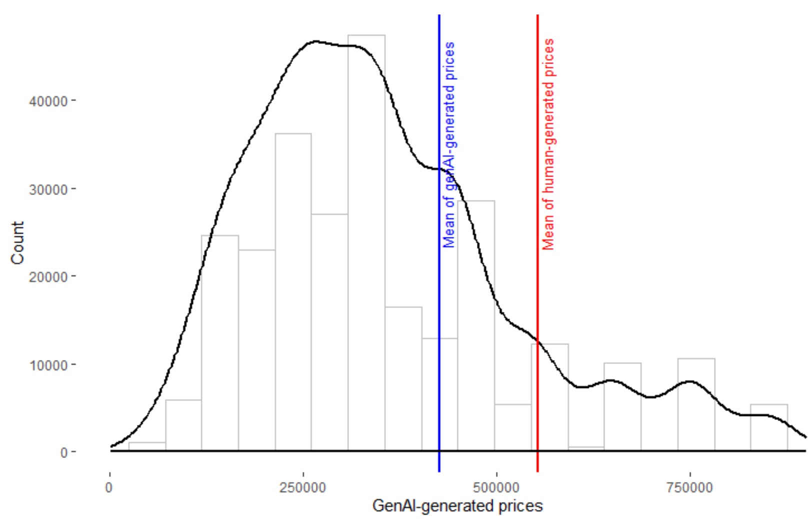

Figure 1: Distribution of the GenAI-generated prices

Figure 1 plots the distribution of GenAI-generated prices. If central tendency is the underlying mechanism, we should observe the GenAI-generated prices centered around the mean of human-generated prices. However, as observed in Figure 1, the GenAI-generated prices are not distributed around the mean of human-generated prices and they appear to be non-monotonic, indicating that central tendency may not apply to GenAI in housing price generation. In addition, we conducted another analysis that aims to capture the heterogeneity effect of GenAI pricing on price discrimination with respect to market position of each property. Here, we added three-way interaction terms between AI_{ijk}, market position factors, and discrimination factors.

Table 7: Exploring mechanisms in GenAI pricing.

		$\log(Price_{ijk})$
AI_{ijk} baseline: human-generated prices		-0.160*** (0.003)
$MarketHighSupply_{ijk}$ baseline: low-supply areas		0.020*** (0.007)
$MarketNewHouse_{ijk}$ baseline: old houses		0.117*** (0.002)
$Underprivileged_{ijk}$ baseline: privileged areas		-0.214*** (0.016)
$AI_{ijk} \times Underprivileged_{ijk}$		0.068*** (0.011)
$AI_{ijk} \times Underprivileged_{ijk} \times MarketHighSupply_{ijk}$		-0.015 (0.011)
$AI_{ijk} \times Underprivileged_{ijk} \times MarketNewHouse_{ijk}$		0.038*** (0.003)
Control variables	$\log(Bed_{ijk}), \log(Bath_{ijk}), \log(Size_{ijk}), \log(Garage_{ijk})$ $Renovate_{ijk}, \log(School_{ijk}), \log(Flood_{ijk}), \log(Fire_{ijk})$ $\log(Heat_{ijk}), \log(Wind_{ijk}), \log(Air_{ijk}), \log(Population_{ijk})$	
Fixed effects		House type, Location
N		569,498
R^2		0.8669

Note. Standard errors in parentheses are robust and clustered by property. * < 0.1, ** < 0.05, *** < 0.01. Other two-way interaction terms are included in the specification but omitted in this table for exposition purposes.

The results are presented in Table 7. If the GenAI-generated prices are driven by the central tendency behavior, then GenAI should reduce the pricing gap in a monotonic fashion across all market positions. However, as observed in Table 7, we find that GenAI reduced price discrimination differently based on the market position. While GenAI pricing reduced discrimination more for new houses, the level of housing supply in the area did not significantly affect the extent of this reduction. As such, central tendency does not appear to be the underlying mechanism that drives the ability of GenAI to reduce price discrimination in the housing market that we observe.

4.3.2 Bias Embedded in the Training Data

In the previous subsection, we show that our finding that GenAI-generated prices tend to reduce pricing discrimination does not seem to be driven by the common behavior in traditional AI models where the variance is lower at the cost of bias. In this subsection, we explore an alternative potential mechanism whereby the results observed in our study could simply mimic the training data (i.e., the housing prices from the real estate aggregator platform). As such, we conduct an analysis that uses the housing prices generated from traditional AI models utilizing supervised learning techniques. This analysis allows us to identify how the existing dataset may contribute to the bias (and bias reduction) of housing prices generated by the AI models. Here, we consider two sets of machine learning algorithms: off-the-shelf algorithms and state-of-the-art algorithms. In essence, the analysis in this section formally compares the prices generated by GenAI, traditional AI models, and humans in terms of housing price discrimination.

Table 8: Impact of Traditional AI pricing on housing price discrimination

	$\log(Price_{ijk})$ RF	$\log(Price_{ijk})$ XGB	$\log(Price_{ijk})$ CATBOOST	$\log(Price_{ijk})$ DNN
AI_{ijk} baseline: human-generated prices	0.214*** (0.001)	0.155*** (0.001)	0.195*** (0.001)	0.018*** (0.0010)
$MarketHighSupply_{ijk}$ baseline: low-supply areas	0.049*** (0.009)	0.053*** (0.009)	0.035*** (0.009)	0.028*** (0.008)
$MarketNewHouse_{ijk}$ baseline: old houses	0.042*** (0.001)	0.052*** (0.001)	0.044*** (0.001)	0.059*** (0.001)
$Underprivileged_{ijk}$ baseline: privileged areas	-0.126*** (0.003)	-0.131*** (0.003)	-0.118*** (0.003)	-0.136*** (0.003)
$AI_{ijk} \times Underprivileged_{ijk}$	-0.090*** (0.002)	-0.082*** (0.002)	-0.064*** (0.002)	-0.009*** (0.002)
Control variables	$\log(Bed_{ijk}), \log(Bath_{ijk}), \log(Size_{ijk}), \log(Garage_{ijk})$ $Renovate_{ijk}, \log(School_{ijk}), \log(Flood_{ijk}), \log(Fire_{ijk})$ $\log(Heat_{ijk}), \log(Wind_{ijk}), \log(Air_{ijk}), \log(Population_{ijk})$			
Fixed effects	House type, Location			
N	170,850	170,850	Yes	Yes
R^2	0.9351	0.9299	Yes	Yes

Note. Standard errors in parentheses are robust and clustered by property. * < 0.1, ** < 0.05, *** < 0.01.

Table 8 reports the results based on the Random Forest (RF) algorithm, XGBoost (XGB) algorithm, contemporary gradient-boosting (CATBOOST) algorithm, and deep neural network (DNN). In all cases, we find that the pricing gap is widened for all models, compared to prices generated by GenAI tools like ChatGPT. This finding indicates that the effects of GenAI on reducing the pricing gap are unlikely to be driven by the dataset. The full details of this analysis, including additional machine learning models, are available in Section B.3 of the online appendix.

4.3.3 Online Conversations about Housing Price Discrimination

Next, we explore the third potential mechanism, which is largely based on how GenAI was trained. While GenAI model training processes are typically opaque and proprietary to their owners, it is important to recall that these GenAI models are large language models (LLMs) trained on textual content available on the Internet. In that regard, there exists early evidence that GenAI models tend to produce results that reflect the conversations and the information that appear in the training data. For example, Kirk et al. (2021) demonstrate that GenAI models appear to pull distributions from the extremes in generating results. Particularly, in the context of occupational bias, the authors find that GenAI tends to produce outcomes resulting in a higher proportion of women in male-dominated jobs, and vice versa for female-dominated jobs. That is because there is more online content (e.g., news, online articles, discussions) advocating for women in male-dominated jobs and men in female-dominated jobs. In our context, we found in Table 7 that the effects of GenAI on reducing the pricing gap are stronger for newly built houses. The conjecture discussed in Kirk et al. (2021) suggests that such results could be driven by the online content on racial-based housing discrimination, which tends to surround newly built houses more than old houses.

To empirically test this potential mechanism, we collect statistics on the amount of online content related to housing discrimination concerning minorities and how they are affected by discrimination practices. Here, we use Google search with the following set of keywords that correspond to independent variables in our study: (1) housing markets, (2) race, and (3) house age. We limit the timeline to October 2023 (i.e., the month that GPT-4o was trained) so that the search results are aligned with the amount of online content that GPT-4o was trained upon. We found 23,600,000 entries of online content (on the Google search engine) regarding housing discrimination related to race. Of these entries, the vast majority of them (22,500,000 entries, i.e., 95.34% of the content) discuss racial-based housing discrimination for new houses. This finding preliminarily supports the conjecture discussed in Kirk et al. (2021) that the content of online discussion may be the underlying mechanism of the effects of interest in our study.

Table 9: Price generated by the constraint prompt

	$\log(Price_{ijk})$
$ConstrainedPrompt_{ijk}$ baseline: original prompt	-0.016*** (0.0005)
$MarketHighSupply_{ijk}$ baseline: low-supply areas	0.061*** (0.017)
$MarketNewHouse_{ijk}$ baseline: old houses	0.044*** (0.003)
$Underprivileged_{ijk}$ baseline: privileged areas	-0.166*** (0.006)
$ConstrainedPrompt_{ijk} \times Underprivileged_{ijk}$	-0.021*** (0.001)
Control variables	$\log(Bed_{ijk}), \log(Bath_{ijk}), \log(Size_{ijk}), \log(Garage_{ijk})$ $Renovate_{ijk}, \log(School_{ijk}), \log(Flood_{ijk}), \log(Fire_{ijk})$ $\log(Heat_{ijk}), \log(Wind_{ijk}), \log(Air_{ijk}), \log(Population_{ijk})$
Fixed effects	House type, Location
N	56,948
R^2	0.9638

Note. Standard errors in parentheses are robust and clustered by property. * < 0.1, ** < 0.05, *** < 0.01.

To further validate this mechanism, we conduct an additional exploratory analysis where we specifically ask ChatGPT to disregard conversations surrounding racial discrimination when estimating housing prices. For efficiency purposes, we randomly select 10% of properties from our dataset and re-estimate the prices with the new *constrained* prompt. Full details of this prompt are available in Section A of the online appendix. Table 9 reports the results of our analysis with the housing prices generated by ChatGPT with the constrained prompt. We find that the pricing gap is widened under the constrained prompt compared to the unconstrained prompt. This finding supports the conjecture that ChatGPT reduces the pricing gap in the housing market due to online conversations and information about racial discrimination in the housing market.

5. Discussion and Conclusions

GenAI has increasingly been adopted for various applications, including pricing. In this paper, we empirically compared GenAI prices to human-generated ones and examined the implications of GenAI pricing for housing discrimination. Housing price discrimination is the phenomenon in which the prices of houses located in certain areas (e.g., white-dominant areas) are higher than those of houses that are otherwise similar but are located in other areas (e.g., minority-dominant areas). Prior studies have demonstrated that such a pricing gap can be exacerbated when traditional AI models are used for pricing, leading us to investigate whether housing prices generated by GenAI amplify or alleviate housing price discrimination. For our analyses, we collected data on housing prices and property characteristics from a leading U.S. real estate aggregator platform.

5.1 Results Summary

Our initial analyses, which compared the housing prices collected from a real estate aggregator platform and the prices generated by GenAI, showed that GenAI tends to generate prices that are lower than human-generated prices. In addition, the prices of houses located in white-dominant areas and high-income areas are higher than those of houses that are otherwise similar but located in minority-dominant areas and low-income areas. Contrary to prices generated by traditional AI models, GenAI-generated housing prices substantially reduce the pricing gap that constitutes price discrimination. Particularly, we found that the price discrepancy gap is reduced by 8.7% when race is the discrimination factor. We also explored the price-generating mechanisms of GenAI and demonstrated that the reduced price discrimination in GenAI pricing is not solely driven by the central tendency behavior that is commonly seen in traditional AI models. Instead, GenAI models tend to generate higher prices for houses located in disadvantaged areas because a disproportionate amount of the Internet content used as training data relates to housing prices in disadvantaged areas compared to prices in advantaged areas. In other words, because there is a larger amount of online content about housing prices in disadvantaged areas, GenAI generates higher (lower) prices for houses in disadvantaged (advantaged) areas compared to the actual prices. Our results address growing concerns that the use of GenAI tools could amplify existing biases embedded in historical data. Nevertheless, we note that GenAI tools should not be used solely for bias mitigation purposes. Instead, fair machine learning models designed to provide fairness based on certain fairness criteria should be considered. We provide a comparison to one of the commonly used fair machine learning models in Section B.2 of the online appendix.

5.2 Theoretical, Managerial, and Policy Implications

Our study provides significant theoretical contributions to the literature on the impacts of GenAI, especially regarding issues of bias and fairness surrounding this emerging AI technology. By demonstrating that GenAI models can reduce the pricing gap associated with discrimination factors such as race and income level, our study challenges traditional perspectives that AI models tend to exacerbate historical biases in existing data. Instead, our findings suggest that by leveraging Internet content and training processes, GenAI can mitigate pricing disparities that have long plagued the housing market. This finding challenges the prevailing belief that AI models simply replicate or amplify historical biases and highlights the potential for AI to contribute positively to fairness and equity in pricing practices. This study opens up new avenues for research on the role of GenAI in addressing societal issues and underscores the importance of considering AI technologies as tools for promoting fairness and social welfare.

Similarly, the managerial and policy implications of this study on GenAI and pricing discrimination in the housing market are significant for real estate professionals and policymakers. The findings suggest that incorporating GenAI models such as ChatGPT in pricing decisions can help mitigate pricing disparities based on discriminatory factors such as race and income level. Real estate agents and companies can leverage GenAI to enhance pricing fairness and promote equitable practices in the housing market. Policymakers can also consider the potential of GenAI technologies in shaping regulations that promote fair housing practices. By understanding the role of GenAI in addressing price discrimination, managers in the real estate industry can make informed decisions that foster a more inclusive and equitable housing market.

5.3 Limitations and Future Research

Our study is not without limitations, which offer excellent avenues for future research. First, our study focused on the effects of GenAI pricing on the supply side of the housing market. Future research can investigate how consumers, whether end-users or real estate professionals, react to the prices generated by GenAI. Second, our research primarily examined the selling prices of houses located in the U.S. Future studies could extend our work to different housing markets, including those in other countries, as well as the rental sector. Such comparative studies would be invaluable for understanding the influence of GenAI in the housing market. Third, our study uses actual housing prices transacted on a real estate aggregation platform as the baseline in our comparison. Such prices may not reflect the true intrinsic prices of those houses. Future studies could develop a model that estimates the intrinsic housing prices based on their characteristics, which could help further identify fundamental biases in housing price generation.

References

Abbasi, A., Parsons, J., Pant, G., Sheng, O. R. L., and Sarker, S. 2024. "Pathways for Design Research on Artificial Intelligence," *Information Systems Research*).

Abouzakhar, N. 2024. "Artificial Intelligence-Based Solution Model for Real Estate Business and Entrepreneurial Operations: Case Study," *Proceedings of The 19th European Conference on Innovation and Entrepreneurship*: Academic Conferences International.

Ahsen, M. E., Ayvaci, M. U. S., and Raghunathan, S. 2019. "When Algorithmic Predictions Use Human-Generated Data: A Bias-Aware Classification Algorithm for Breast Cancer Diagnosis," *Information Systems Research* (30:1), pp. 97-116.

Alavi, M., Leidner, D. E., and Mousavi, R. 2024. "A Knowledge Management Perspective of Generative Artificial Intelligence," *Journal of the Association for Information Systems* (25:1), pp. 1-12.

Apgar, W. C., and Calder, A. 2005. "The Dual Mortgage Market: The Persistence of Discrimination in Mortgage Lending,").

Auspurg, K., Hinz, T., and Schmid, L. 2017. "Contexts and Conditions of Ethnic Discrimination: Evidence from a Field Experiment in a German Housing Market," *Journal of Housing Economics* (35), pp. 26-36.

Bai, B., Dai, H., Zhang, D. J., Zhang, F., and Hu, H. 2022. "The Impacts of Algorithmic Work Assignment on Fairness Perceptions and Productivity: Evidence from Field Experiments," *Manufacturing & Service Operations Management* (24:6), pp. 3060-3078.

Banh, L., and Strobel, G. 2023. "Generative Artificial Intelligence," *Electronic Markets* (33:1), pp. 1-17.

Bayer, P., Ferreira, F., and Ross, S. L. 2018. "What Drives Racial and Ethnic Differences in High-Cost Mortgages? The Role of High-Risk Lenders," *The Review of Financial Studies* (31:1), pp. 175-205.

Becker, G. S. 1957. *The Economics of Discrimination*. University of Chicago press.

Bellamy, M. A., Dhanorkar, S., and Subramanian, R. 2020. "Administrative Environmental Innovations, Supply Network Structure, and Environmental Disclosure." Wiley Online Library, pp. 895-932.

Benbya, H., Strich, F., and Tamm, T. 2024. "Navigating Generative Artificial Intelligence Promises and Perils for Knowledge and Creative Work," *Journal of the Association for Information Systems* (25:1), pp. 23-36.

Bloomberg. 2023. "Generative Ai to Become a $1.3 Trillion Market by 2032, Research Finds." Bloomberg.

Chen, Y., Andiappan, M., Jenkin, T., and Ovchinnikov, A. 2023. "A Manager and an Ai Walk into a Bar: Does Chatgpt Make Biased Decisions Like We Do?," *Available at SSRN 4380365*).

Cheng, P., Lin, Z., and Liu, Y. 2015. "Racial Discrepancy in Mortgage Interest Rates," *The Journal of Real Estate Finance and Economics* (51), pp. 101-120.

Christensen, P., and Timmins, C. 2022. "Sorting or Steering: The Effects of Housing Discrimination on Neighborhood Choice," *Journal of Political Economy* (130:8), pp. 2110-2163.

Chui, M., Yee, L., Hall, B., Singla, A., and Sukharevsky, A. 2023. "The State of Ai in 2023: Generative Ai's Breakout Year," QuantumBlack AI by McKensy.

Cohen, M. C., Elmachtoub, A. N., and Lei, X. 2022. "Price Discrimination with Fairness Constraints," *Management Science* (68:12), pp. 8536-8552.

Cohen, M. C., Miao, S., and Wang, Y. 2021. "Dynamic Pricing with Fairness Constraints," *Available at SSRN 3930622*).

Cui, R., Li, J., and Zhang, D. J. 2020. "Reducing Discrimination with Reviews in the Sharing Economy: Evidence from Field Experiments on Airbnb," *Management Science* (66:3), pp. 1071-1094.

Dai, T., and Tayur, S. 2022. "Designing Ai-Augmented Healthcare Delivery Systems for Physician Buy-in and Patient Acceptance," *Production and Operations Management* (31:12), pp. 4443-4451.

Davison, R. M., Chughtai, H., Nielsen, P., Marabelli, M., Iannacci, F., van Offenbeek, M., Tarafdar, M., Trenz, M., Techatassanasoontorn, A. A., and Díaz Andrade, A. 2024. "The Ethics of Using Generative Ai for Qualitative Data Analysis," *Information Systems Journal*).

Dencik, J., Goehring, B., and Marshall, A. 2023. "Managing the Emerging Role of Generative Ai in Next-Generation Business," *Strategy & Leadership* (51:6), pp. 30-36.

Desk, T. 2023. "Chatgpt Most Popular Ai Tool with 14.6 Billion Visits in a Year, Beats Google Bard & Midjourney." The Indian Express.

deVise, P. 1984. "Housing Discrimination in the Chicago Metropolitan Area: The Legacy of the Brown Decision," *DePaul L. Rev.* (34), p. 491.

Dudhat, L. 2024. "Exploring 20 Innovative Uses of Generative Ai in Real Estate." Retrieved 2025-01-25, 2025, from https://solguruz.com/blog/top-use-cases-of-generative-ai-in-the-real-estate-industry/

Eloundou, T., Manning, S., Mishkin, P., and Rock, D. 2023. "Gpts Are Gpts: An Early Look at the Labor Market Impact Potential of Large Language Models," *arXiv preprint arXiv:2303.10130*).

Ferrara, E. 2023. "Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies," *Sci* (6:1), p. 3.

Finno, A. 2023. "Building Price with Chatgpt." Retrieved 2025-01-25, 2025, from https://www.willowtreeapps.com/insights/building-price-with-chatgpt

Fitzpatrick, M., Gujral, V., Kapoor, A., and Wolkomir, A. 2023. "Real Estate Can Use Generative Ai to Turn the Industry's Data into Treasure in Seven Steps." Retrieved 2025-01-25, 2025, from https://www.mckinsey.com/industries/real-estate/our-insights/generative-ai-can-change-real-estate-but-the-industry-must-change-to-reap-the-benefits

Flage, A. 2018. "Ethnic and Gender Discrimination in the Rental Housing Market: Evidence from a Meta-Analysis of Correspondence Tests, 2006–2017," *Journal of Housing Economics* (41), pp. 251-273.

Foggo, V., and Villasenor, J. 2020. "Algorithms, Housing Discrimination, and the New Disparate Impact Rule," *COLuM. Sci. & TECH. L. REv.* (22), p. 1.

Fowler, A., Heider, C., Matar, M., and Stec, T. 2023. "Three Ways to Get More Value from Your Pricing Process with Generative Ai." Simon-Kucher.

Fu, R., Aseri, M., Singh, P. V., and Srinivasan, K. 2022. "“Un” Fair Machine Learning Algorithms," *Management Science* (68:6), pp. 4173-4195.

Fu, R., Huang, Y., Mehta, N., Singh, P. V., and Srinivasan, K. 2023. "Unequal Impact of Zestimate on the Housing Market." Available at SSRN 4480469.

Fu, R., Huang, Y., and Singh, P. V. 2021. "Crowds, Lending, Machine, and Bias," *Information Systems Research* (32:1), pp. 72-92.

Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., and Chen, L. 2023. "Generative Ai and Chatgpt: Applications, Challenges, and Ai-Human Collaboration." Taylor & Francis, pp. 277-304.

Gao, X., Jasin, S., Najafi, S., and Zhang, H. 2022. "Joint Learning and Optimization for Multi-Product Pricing (and Ranking) under a General Cascade Click Model," *Management Science* (68:10), pp. 7362-7382.

Guerrieri, V., Hartley, D., and Hurst, E. 2013. "Endogenous Gentrification and Housing Price Dynamics," *Journal of Public Economics* (100), pp. 45-60.

Gunarathne, P., Rui, H., and Seidmann, A. 2022. "Racial Bias in Customer Service: Evidence from Twitter," *Information Systems Research* (33:1), pp. 43-54.

Guo, Y., Yang, Y., and Abbasi, A. 2022. "Auto-Debias: Debiasing Masked Language Models with Automated Biased Prompts," *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1012-1023.

Hanson, A., and Hawley, Z. 2011. "Do Landlords Discriminate in the Rental Housing Market? Evidence from an Internet Field Experiment in Us Cities," *Journal of urban Economics* (70:2-3), pp. 99-114.

Ihlanfeldt, K., and Mayock, T. 2009. "Price Discrimination in the Housing Market," *Journal of Urban Economics* (66:2), pp. 125-140.
Kahneman, D., Slovic, P., and Tversky, A. 1982. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge university press.
Kamin, D. 2022. "Widespread Racial Bias Found in Home Appraisals," in: *The New York Times*.
Kirk, H. R., Jun, Y., Volpin, F., Iqbal, H., Benussi, E., Dreyer, F., Shtedritski, A., and Asano, Y. 2021. "Bias out-of-the-Box: An Empirical Analysis of Intersectional Occupational Biases in Popular Generative Language Models," *Advances in neural information processing systems* (34), pp. 2611-2624.
Kordzadeh, N., and Ghasemaghaei, M. 2022. "Algorithmic Bias: Review, Synthesis, and Future Research Directions," *European Journal of Information Systems* (31:3), pp. 388-409.
KPMG. 2023. "Unlocking the Value of Generative Ai in Pricing."
Kuebler, M. 2013. "Closing the Wealth Gap: A Review of Racial and Ethnic Inequalities in Homeownership," *Sociology Compass* (7:8), pp. 670-685.
Lalor, J. P., Abbasi, A., Oketch, K., Yang, Y., and Forsgren, N. 2024. "Should Fairness Be a Metric or a Model? A Model-Based Framework for Assessing Bias in Machine Learning Pipelines," *ACM Transactions on Information Systems* (42:4), pp. 1-41.
Lambrecht, A., and Tucker, C. E. 2015. "Can Big Data Protect a Firm from Competition?," *Available at SSRN 2705530*).
Lanoue, p. 2024. "How to Make a Pricing Spreadsheet Using Chatgpt." Retrieved 2025-01-25, 2025, from https://www.thebricks.com/resources/how-to-make-a-pricing-spreadsheet-using-chatgpt
Larose, D. T. 2015. *Data Mining and Predictive Analytics*. John Wiley & Sons.
Lee, D., and Hosanagar, K. 2019. "How Do Recommender Systems Affect Sales Diversity? A Cross-Category Investigation Via Randomized Field Experiment," *Information Systems Research* (30:1), pp. 239-259.
Liu, F., Budiu, R., Zhang, A., and Cionca, E. 2023. "Chatgpt, Bard, or Bing Chat? Differences among 3 Generative-Ai Bots." Nielsen Norman Group.
Lou, B., Sun, H., and Sun, T. 2023. "Gpts and Labor Markets in the Developing Economy: Evidence from China," *Available at SSRN 4426461*).
Möhlmann, M., Zalmanson, L., Henfridsson, O., and Gregory, R. W. 2021. "Algorithmic Management of Work on Online Labor Platforms: When Matching Meets Control," *MIS Quarterly* (45:4).
Myers, C. K. 2004. "Discrimination and Neighborhood Effects: Understanding Racial Differentials in Us Housing Prices," *Journal of urban economics* (56:2), pp. 279-302.
Nair, D., and Saenz, M. J. 2024. "Pair People and Ai for Better Product Demand Forecasting,").
Nazer, L. H., Zatarah, R., Waldrip, S., Ke, J. X. C., Moukheiber, M., Khanna, A. K., Hicklen, R. S., Moukheiber, L., Moukheiber, D., and Ma, H. 2023. "Bias in Artificial Intelligence Algorithms and Recommendations for Mitigation," *PLOS Digital Health* (2:6), p. e0000278.
Neal, M., Zhu, L., and Pruitt, M. 2024. "Subcommittee Hearing on Artificial Intelligence and Housing: Exploring Promise and Peril,").
Neal, M., Zhu, L., Young, C., Perry, V. G., and Pruitt, M. 2023. "Harnessing Artificial Intelligence for Equity in Mortgage Finance,").
Noy, S., and Zhang, W. 2023. "Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence," *Available at SSRN 4375283*).
Ooi, K.-B., Tan, G. W.-H., Al-Emran, M., Al-Sharafi, M. A., Capatina, A., Chakraborty, A., Dwivedi, Y. K., Huang, T.-L., Kar, A. K., and Lee, V.-H. 2023. "The Potential of Generative Artificial Intelligence across Disciplines: Perspectives and Future Directions," *Journal of Computer Information Systems*), pp. 1-32.
OpenAI. 2024a. "Hello Gpt-4o - Model Evaluations."
OpenAI. 2024b. "Models."
Qayyum, F., Khan, M. A., Kim, D.-H., Ko, H., and Ryu, G.-A. 2023. "Explainable Ai for Material Property Prediction Based on Energy Cloud: A Shapley-Driven Approach," *Materials* (16:23), p. 7322.
Quillian, L., Lee, J. J., and Honoré, B. 2020. "Racial Discrimination in the Us Housing and Mortgage Lending Markets: A Quantitative Review of Trends, 1976–2016," *Race and Social Problems* (12), pp. 13-28.
Rastogi, C., Zhang, Y., Wei, D., Varshney, K. R., Dhurandhar, A., and Tomsett, R. 2022. "Deciding Fast and Slow: The Role of Cognitive Biases in Ai-Assisted Decision-Making," *Proceedings of the ACM on Human-Computer Interaction* (6:CSCW1), pp. 1-22.

Rhue, L. 2024. "The Anchoring Effect, Algorithmic Fairness, and the Limits of Information Transparency for Emotion Artificial Intelligence," *Information Systems Research* (35:3), pp. 1479-1496.

Richardson, B., and Gilbert, J. E. 2021. "A Framework for Fairness: A Systematic Review of Existing Fair Ai Solutions," *arXiv preprint arXiv:2112.05700*).

Roselli, D., Matthews, J., and Talagala, N. 2019. "Managing Bias in Ai," *Companion proceedings of the 2019 world wide web conference*, pp. 539-544.

Rosen, E., Garboden, P. M., and Cossyleon, J. E. 2021. "Racial Discrimination in Housing: How Landlords Use Algorithms and Home Visits to Screen Tenants," *American Sociological Review* (86:5), pp. 787-822.

Ross, S. L., and Turner, M. A. 2005. "Housing Discrimination in Metropolitan America: Explaining Changes between 1989 and 2000," *Social Problems* (52:2), pp. 152-180.

Sabherwal, R., and Grover, V. 2024. "The Societal Impacts of Generative Artificial Intelligence: A Balanced Perspective," *Journal of the Association for Information Systems* (25:1), pp. 13-22.

Sætra, H. S. 2023. "Generative Ai: Here to Stay, but for Good?," *Technology in Society* (75), p. 102372.

Schade, M. 2023. "How Chatgpt and Our Language Models Are Developed." OpenAI.

Schafer, R. 2017. "Discrimination in Housing Prices and Mortgage Lending," *Housing Urban America*), pp. 294-308.

Schwemm, R. G., and Taren, J. L. 2010. "Discretionary Pricing, Mortgage Discrimination, and the Fair Housing Act," *Harv. CR-CLL Rev.* (45), p. 375.

Shimao, H., Khern-am-nuai, W., Kannan, K. N., and Cohen, M. C. 2025. "Strategic Best-Response Fairness Framework for Fair Machine Learning," *Information Systems Research* (Forthcoming).

So, W., Lohia, P., Pimplikar, R., Hosoi, A., and D'Ignazio, C. 2022. "Beyond Fairness: Reparative Algorithms to Address Historical Injustices of Housing Discrimination in the Us," *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, pp. 988-1004.

Soleimani, M., Intezari, A., and Pauleen, D. J. 2022. "Mitigating Cognitive Biases in Developing Ai-Assisted Recruitment Systems: A Knowledge-Sharing Approach," *International Journal of Knowledge Management (IJKM)* (18:1), pp. 1-18.

Susarla, A., Gopal, R., Thatcher, J. B., and Sarker, S. 2023. "The Janus Effect of Generative Ai: Charting the Path for Responsible Conduct of Scholarly Activities in Information Systems," *Information Systems Research*).

Tang, C., Li, S., Ding, Y., Gopal, R. D., and Zhang, G. 2023. "Racial Discrimination and Anti-Discrimination: The Covid-19 Pandemic's Impact on Chinese Restaurants in North America," *Information Systems Research*).

Terwiesch, C., Olivares, M., Staats, B. R., and Gaur, V. 2020. "Om Forum—a Review of Empirical Operations Management over the Last Two Decades," *Manufacturing & Service Operations Management* (22:4), pp. 656-668.

Van Slyke, C., Johnson, R. D., and Sarabadani, J. 2023. "Generative Artificial Intelligence in Information Systems Education: Challenges, Consequences, and Responses," *Communications of the Association for Information Systems* (53:1), p. 14.

Wang, D.-Y., Ding, J., Sun, A.-L., Liu, S.-G., Jiang, D., Li, N., and Yu, J.-K. 2023a. "Artificial Intelligence Suppression as a Strategy to Mitigate Artificial Intelligence Automation Bias," *Journal of the American Medical Informatics Association* (30:10), pp. 1684-1692.

Wang, W., Liu, X., Zhang, X., and Hong, Y. 2023b. "Knowledge Trap: Human Experts Distracted by Details When Teaming with Ai," *Available at SSRN*).

Wang, X., Tzeng, S.-Y., and Mardani, A. 2022. "Spatial Differentiation and Driving Mechanisms of Urban Household Waste Separation Behavior in Shanghai, China," *Technological Forecasting and Social Change* (181), p. 121753.

Yinger, J. 1978. "The Black-White Price Differential in Housing: Some Further Evidence," *Land Economics* (54:2), pp. 187-206.

Zhou, M., Zhang, J., and Adomavicius, G. 2023. "Longitudinal Impact of Preference Biases on Recommender Systems' Performance," *Information Systems Research*).

Zou, L., and Khern-am-nuai, W. 2023. "Ai and Housing Discrimination: The Case of Mortgage Applications," *AI and Ethics* (3:4), pp. 1271–1281.

Online Appendix to

Generative AI and Price Discrimination in the Housing Market

A. Prompt Examples

We set the model's system role as *"As a realtor, you estimate a house's price in the US using your prior knowledge of the neighborhood's demographics and characteristics, without accessing the internet"* and passed each listing's information through the user role as *"Here are the house details: Zip: [zip code], City: [city name], State: [state name], Housing Age: [number], Bed: [number], Bath: [number], Sqft: [number], Garage Space: [number], Property Type: [single-family houses, condos, multi-family houses, townhomes, or duplexes/triplexes], Mean School Rating out of 10: [1-10], Flood Risk out of 10: [1-10], Wildfire Risk out of 10: [1-10], Heat Risk out of 10: [1-10], Wind Risk out of 10: [1-10], Air Pollution Risk out of 10: [1-10], Renovated: [true or false]"*. We also control the model response to include only a specific price and not other information by employing as assistant role as *"You will provide the price as a number only and no other text. For example, if your answer is 'Based on my knowledge of similar markets and home values in the area, I estimate the price as XXX,XXX USD,' you'll only say 'XXXXXX'."* and set the model temperature at zero to receive a deterministic response (Alto 2024). An example code we used to ask ChatGPT for an estimated housing price through its OpenAI API is presented below:

```
from openai import OpenAI

SYSTEM = "As a realtor, you estimate a house's price in the US using your
          prior knowledge of the neighborhood's demographics and
          characteristics, without accessing the internet. "
USER = "Here are the house details: {'zip': 48152, 'city': 'Livonia', 'state':
       'MI', 'house_age': 88, 'bed': 3.0, 'bath': 1.0, 'sqft': 926.0,
       'garage_space': 1.0, 'property_type': 'single family',
       'mean_school_rate_score_out_of_10': 8.17, 'risk_flood_score_out_of_10':
       1.0, 'risk_wildfire_score_out_of_10': 1.0, 'risk_heat_score_out_of_10':
       3.0, 'risk_wind_score_out_of_10': 2.0, 'risk_air_score_out_of_10': 3.0,
       'renovated': False"
ASSISTANT = "You will provide the price as a number only and no other text.
            For example, if the your answer is 'Based on my knowledge of
            similar markets and home values in the area, I estimate the price
            as XXX,XXX USD', you'll only say 'XXXXXX'."
```

```
client = OpenAI(api_key='')
response = client.chat.completions.create(model = 'gpt-4o-2024-08-06',
                                          temperature = 0,
                                          messages = [{'role':'system',
                                                        'content': SYSTEM},
                                                       {'role':'user',
                                                        'content': USER},
                                                       {'role':'assistant',
                                                        'content': ASSISTANT}
                                                      ]
                                          )
```

After that, we extract the response from ChatGPT using the following code:

```
print(response.choices[0].message.content)
```

which result in as:

```
210000.0
```

In addition, in Section 4.3, we rely on a *constrained* prompt where we explicitly ask ChatGPT to ignore conversations related to racial discrimination when it estimates housing prices. The constraint we add is as follows "*When estimating the price, you do not concern about pricing discrimination by neighborhood's information. For example, if you think that the price is low because the house locates in a minority-dominated area, let it be.*"

B. Additional Analysis

In this appendix, we perform additional analyses to supplement the empirical analyses we conducted in the paper.

B.1 Pricing Variation Analysis

We explore how much pricing variations are accounted for by GPT-generated prices. Here, we run two regression models where the outcome variable of the first model is the housing prices generated by humans and that of the second model is the housing prices generated by ChatGPT. The explanatory variables of both models are the factors impacting real estate agents' pricing decisions. These factors are also used as input prompts to generate housing prices from ChatGPT. The results, reported in Table B.1 below, show that the R^2 value of the second model (i.e., the model with GPT-generated prices) is substantially higher than that of the first model (i.e., the model with human-generated prices). This finding suggests that pricing variations in general can be accounted for by ChatGPT more than by humans.

Table B.1: Human-generated prices vs. GPT-generated prices

	Human-generated prices $\log(Price_{ijk})$	GPT-generated prices $\log(Price_{ijk})$
$MarketHighSupply_{ijk}$	0.062*** (0.010)	0.049*** (0.004)
baseline: low-supply areas		
$MarketNewHouse_{ijk}$	0.076*** (0.002)	0.047*** (0.0007)
baseline: old houses		
$Underprivileged_{ijk}$	-0.204*** (0.003)	-0.168*** (0.002)
baseline: privileged areas		
$\log(Bed_{ijk})$	-0.085*** (0.004)	0.068*** (0.002)
$\log(Bath_{ijk})$	0.285*** (0.003)	0.190*** (0.001)
$\log(Size_{ijk})$	0.671*** (0.004)	0.493*** (0.002)
$\log(Garage_{ijk})$	0.081*** (0.002)	0.074*** (0.0007)
$Renovate_{ijk}$	0.022*** (0.002)	0.043*** (0.0006)
$\log(School_{ijk})$	0.231*** (0.004)	0.249*** (0.002)
$\log(Flood_{ijk})$	0.026*** (0.001)	-0.005*** (0.0005)
$\log(Fire_{ijk})$	-0.033*** (0.002)	-0.016*** (0.0006)
$\log(Heat_{ijk})$	-0.200*** (0.007)	-0.018*** (0.003)
$\log(Wind_{ijk})$	0.201*** (0.013)	0.048*** (0.005)
$\log(Air_{ijk})$	0.062*** (0.005)	0.017*** (0.002)
$\log(Population_{ijk})$	-0.075*** (0.003)	-0.039*** (0.001)
House type fixed effects	Yes	Yes
Location fixed effects	Yes	Yes
N	284,749	284,749
R^2	0.8350	0.9613

Note. Standard errors in parentheses are robust and clustered by property. * < 0.1, ** < 0.05, *** < 0.01.

B.2 Comparing with FairML Model

Our main results show that housing price discrimination is lower with GPT-generated prices compared with human-generated prices. In this appendix, we expand our analysis by considering machine learning models that are designed specifically to provide fairness in prediction results. Specifically, we utilize a similar set of information included in the ChatGPT prompt, encode the categorical information using the leave-one-out method, and employ Optuna (Akiba et al. 2019) with the tree-structured Parzen estimator (TPE) algorithm (Watanabe 2023) and 3-fold cross-validation approach to search for the optimal hyperparameter set for 200 trails.

The development of our fairML includes two stages, involving both pre-process and in-process methods. First, although we do not directly include the discrimination factor in an ML model (i.e., whether the property is in a white-dominated area), the model's fairness can still be violated due to the associations between the discrimination factor and the model's variables (i.e., property characteristics and neighborhood information). As such, our first stage aims to mitigate these correlations, and we employ the CorrelationRemover, a pre-processing algorithm from the Fairlearn package developed by Microsoft (Bird et al. 2020). This algorithm performs a linear transformation on the model's variables, aiming to attenuate the correlations while minimizing the least-squares error between the original and transformed information. The extent of the linear projection (alpha) corresponds to the degree of correlation removal, and it is a crucial hyperparameter, influencing the trade-off between the model's fairness and its performance.

The second stage involves developing an ML model to achieve the highest model fairness and performance. We train the model with the transformed variables and select Extreme Gradient Boosting (XGB) as the algorithm, as it demonstrated the highest performance, considering both training time and mean absolute percentage error (MAPE). We define fairness as the absolute difference between MAPE for each group (i.e., white-dominant areas and not) as it reflects the disparity in the cost of utilizing a model between groups (Corbett-Davies et al. 2023). A value equal to zero indicates a (perfectly) fair model. Because our training aims to balance fairness and performance, we combine the value with the model's overall MAPE, considering the combined value as the model loss. We then set the objective function during the hyperparameter search to minimize the mean of this loss value across 3-fold. The results, reported in Table B.2, show that the bias is reduced more when housing prices are generated by the fairML model. However, the R^2 value of the fairML model is significantly lower than that of GenAI. This trade-off between fairness and accuracy is commonly observed in the literature (e.g., Bai et al. 2022a; Fu et al. 2021).

Table B.2: Price generated by Fair ML and ChatGPT

	$\log(Price_{ijk})$ FairXGB	$\log(Price_{ijk})$ ChatGPT
AI_{ijk} baseline: human-generated prices	-0.650*** (0.002)	-0.149*** (0.001)
$MarketHighSupply_{ijk}$ baseline: low-supply areas	0.037*** (0.010)	0.061*** (0.011)
$MarketNewHouse_{ijk}$ baseline: old houses	0.041*** (0.002)	0.056*** (0.002)
$Underprivileged_{ijk}$ baseline: privileged areas	-0.170*** (0.005)	-0.227*** (0.005)
$\boldsymbol{AI_{ijk} \times Underprivileged_{ijk}}$	**0.124*** (0.006)**	**0.084*** (0.003)**
$\log(Bed_{ijk})$	-0.073*** (0.004)	-0.016*** (0.004)
$\log(Bath_{ijk})$	0.228*** (0.003)	0.232*** (0.004)
$\log(Size_{ijk})$	0.469*** (0.004)	0.597*** (0.004)
$\log(Garage_{ijk})$	0.043*** (0.002)	0.077*** (0.002)
$Renovate_{ijk}$	0.011*** (0.002)	0.033*** (0.002)
$log(School_{ijk})$	0.156*** (0.004)	0.237*** (0.005)
$log(Flood_{ijk})$	0.014*** (0.001)	0.011*** (0.001)
$\log(Fire_{ijk})$	-0.018*** (0.002)	-0.025*** (0.002)
$\log(Heat_{ijk})$	-0.103*** (0.007)	-0.103*** (0.008)
$\log(Wind_{ijk})$	0.114*** (0.012)	0.135*** (0.014)
$\log(Air_{ijk})$	0.035*** (0.004)	0.038*** (0.005)
$\log(Population_{ijk})$	-0.048*** (0.003)	-0.061*** (0.003)
House type fixed effects	Yes	Yes
Location fixed effects	Yes	Yes
N	170,850	170,850
R^2	0.7577	0.8737

B.3 Traditional AI Models

In this appendix, we develop several supervised ML models based on the exact information included in the ChatGPT prompt. For this analysis, we split 70% of the data into the training set and 30% into the test set.

We consider two sets of ML models. The first set includes commonly used ML models in the information systems literature (e.g., Liu et al. 2020; Wang et al. 2021; Xu et al. 2024; Yu et al. 2024) and the second comprises state-of-the-art algorithms. The first set consists of Stochastic Gradient Descent Regression

(SGDREG), K-Nearest Neighbors (KNN), Random Forest (RF), Extreme Gradient Boosting (XGB), Light Gradient-Boosting Machine (LGBM), and Neural Network (NN). SGDREG is an extension of linear regression that incorporates regularization techniques, in which the model is fitted by minimizing an empirical loss through stochastic gradient descent (Pedregosa et al. 2011). KNN is a distance-based algorithm that makes predictions by averaging the values of the nearest instances in the feature space (Rokach et al. 2023). RF is an ensemble learning method that employs the bagging technique, leveraging several Decision Trees (DTs) to mitigate model's overfitting and enhance its performance (Breiman 2001). Both XGB and LGBM are gradient-boosted DT algorithms designed to enhance model performance and efficiency. However, they employ distinct tree growth strategies. While XGB adopts a depth-wise approach, prioritizing balanced tree structures and facilitating overfitting control, LGBM utilizes a leaf-wise approach, prioritizing the most impactful leaves for splitting, leading to higher accuracy and reduced loss (Chen and Guestrin 2016; Ke et al. 2017). Lastly, NN is a non-linear model that can learn complex relationships within data (Murtagh 1991).

The second set consists of Hybrid Regression (HYBRID), stacked generalization (STACK), Categorical Boosting (CATBOOST), and Deep Neural Network (DNN). HYBRID is an ensemble technique that combines predictions from multiple models by equally weighting their outputs to achieve a final prediction (Zhang and Ma 2012). Our HYBRID model is constructed using the top three performing ML models from the first set. STACKING is another ensemble method that combines predictions from multiple models. However, based on their function, the models are categorized into two groups: base-learner, responsible for generating initial predictions, and meta-learner, which leverages these predictions as inputs to produce the final output (Zhang and Ma 2012). We employ the top performing ML model from the first set as the meta-learner. The remaining two models with the next best performance are the base models. CATBOOST is a contemporary gradient-boosting algorithm that employs decision trees designed to improve performance when handling categorical variables (Dorogush et al. 2018). Lastly, DNN is a class of neural networks distinguished by their deeper architectures, which excel at handling highly complex datasets. It is usually a backbone architecture in house price prediction systems, such as Zillow's Zestimates (Johnson 2023; Zillow 2023).

Before training the ML models, we transform the categorical features into numerical representations. Since some categorical variables (i.e., zip code, city, state) are high dimensional, we employ the leave-one-out encoding method for the data transformation. This approach allows us to avoid sparse data and the curse of dimensionality, which usually result from the one-hot-encoding approach (i.e., the commonly used categorical encoding technique) and suffer the model performance. Furthermore, the method mitigates overfitting and data leakage by excluding the current sample during encoding (Vasques 2024). During each model's development, we use Optuna (Akiba et al. 2019), a cutting-edge hyperparameter optimization framework with a sophisticated trial-based search approach, to find the optimal hyperparameter set and utilize the tree-structured Parzen estimator (TPE) algorithm (Watanabe 2023), a Bayesian optimization method, as the sampling algorithm. The search is conducted for 200 trials, with the objective function of minimizing the mean of mean absolute percentage error (MAPE) across 3-fold. However, we deviate from the K-fold cross-validation approach for the deep learning model to expedite the training process, in which the whole dataset is fed to train the model several times (i.e., epochs). Instead, we split the training set into 70 percent for the model training and utilize the remaining 30 percent to evaluate the model performance. The MAPE acquired from this holdout set serves as the input for the minimization objective function during the hyperparameter search.

The predicted values from the validation set are compared with the prices generated by ChatGPT. Tables B.3.1 and B.3.2 present the results, showing that the prices predicted by all ML models exhibit higher levels of discrimination compared to the prices generated by ChatGPT as the coefficient of $\boldsymbol{SupervisedML_{ijk}} \times \boldsymbol{Underprivileged_{ijk}}$ is negative.

Table B.3.1: Price generated by off-the-shelf ML models

	$\log(Price_{ijk})$ SGDREG	$\log(Price_{ijk})$ KNN	$\log(Price_{ijk})$ RF	$\log(Price_{ijk})$ XGB	$\log(Price_{ijk})$ LGBM	$\log(Price_{ijk})$ NN
$SupervisedML_{ijk}$ baseline: AI pricing	0.169*** (0.004)	0.174*** (0.001)	0.214*** (0.001)	0.155*** (0.001)	0.195*** (0.001)	0.134*** (0.004)
$MarketHighSupply_{ijk}$ baseline: low-supply areas	-0.029 (0.030)	0.040*** (0.008)	0.049*** (0.009)	0.053*** (0.009)	0.064*** (0.016)	0.014 (0.028)
$MarketNewHouse_{ijk}$ baseline: old houses	0.010*** (0.004)	0.052*** (0.001)	0.042*** (0.001)	0.052*** (0.001)	0.045*** (0.002)	0.032*** (0.004)
$Underprivileged_{ijk}$ baseline: privileged areas	-0.056*** (0.010)	-0.118*** (0.003)	-0.126*** (0.003)	-0.131*** (0.003)	-0.123*** (0.004)	-0.105*** (0.010)
$\boldsymbol{SupervisedML_{ijk} \times Underprivileged_{ijk}}$	**-0.177*** (0.015)**	**-0.041*** (0.002)**	**-0.090*** (0.002)**	**-0.082*** (0.002)**	**-0.078*** (0.004)**	**-0.200*** (0.013)**
$\log(Bed_{ijk})$	0.330*** (0.013)	0.059*** (0.003)	-0.0006 (0.003)	-0.008*** (0.003)	-0.039*** (0.004)	0.159*** (0.018)
$\log(Bath_{ijk})$	0.507*** (0.008)	0.288*** (0.003)	0.213*** (0.002)	0.243*** (0.003)	0.245*** (0.005)	0.312*** (0.017)
$\log(Size_{ijk})$	0.526*** (0.009)	0.471*** (0.003)	0.612*** (0.003)	0.586*** (0.003)	0.598*** (0.005)	0.521*** (0.019)
$\log(Garage_{ijk})$	0.030*** (0.005)	0.055*** (0.001)	0.057*** (0.001)	0.075*** (0.002)	0.068*** (0.003)	0.090*** (0.005)
$Renovate_{ijk}$	0.038*** (0.004)	0.039*** (0.001)	0.025*** (0.001)	0.035*** (0.001)	0.037*** (0.002)	0.039*** (0.004)
$log(School_{ijk})$	0.411*** (0.015)	0.253*** (0.003)	0.223*** (0.004)	0.238*** (0.004)	0.236*** (0.005)	0.365*** (0.015)
$log(Flood_{ijk})$	0.050*** (0.003)	0.014*** (0.001)	0.012*** (0.001)	0.010*** (0.001)	0.017*** (0.001)	0.010*** (0.004)
$\log(Fire_{ijk})$	-0.055*** (0.004)	-0.027*** (0.001)	-0.022*** (0.001)	-0.031*** (0.001)	-0.037*** (0.002)	-0.057*** (0.004)
$\log(Heat_{ijk})$	0.021 (0.019)	-0.067*** (0.005)	-0.046*** (0.005)	-0.071*** (0.005)	-0.073*** (0.007)	-0.117*** (0.019)
$\log(Wind_{ijk})$	0.081*** (0.029)	0.075*** (0.009)	0.070*** (0.009)	0.099*** (0.009)	0.112*** (0.013)	0.083*** (0.030)
$\log(Air_{ijk})$	0.068*** (0.014)	0.029*** (0.004)	0.030*** (0.004)	0.042*** (0.004)	0.037*** (0.005)	-0.016 (0.013)
$\log(Population_{ijk})$	-0.009 (0.007)	-0.048*** (0.002)	-0.045*** (0.002)	-0.047*** (0.002)	-0.048*** (0.003)	-0.022*** (0.008)
House type fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Location fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	170,850	170,850	170,850	170,850	170,850	170,850
R^2	0.5307	0.9263	0.9351	0.9299	0.8611	0.5073

Note. Standard errors in parentheses are robust and clustered by property. * < 0.1, ** < 0.05, *** < 0.01.

Table B.3.2: Price generated by advanced ML models

	$\log(Price_{ijk})$ HYBRID	$\log(Price_{ijk})$ STACK	$\log(Price_{ijk})$ CATBOOST	$\log(Price_{ijk})$ DNN
$SupervisedML_{ijk}$ baseline: GenAI	0.187*** (0.0009)	0.204*** (0.0009)	0.195*** (0.001)	0.018*** (0.0010)
$MarketHighSupply_{ijk}$ baseline: low-supply areas	0.046*** (0.008)	0.046*** (0.008)	0.035*** (0.009)	0.028*** (0.008)
$MarketNewHouse_{ijk}$ baseline: old houses	0.048*** (0.001)	0.044*** (0.001)	0.044*** (0.001)	0.059*** (0.001)
$Underprivileged_{ijk}$ baseline: privileged areas	-0.124*** (0.003)	-0.119*** (0.003)	-0.118*** (0.003)	-0.136*** (0.003)
$\boldsymbol{SupervisedML_{ijk} \times Underprivileged_{ijk}}$	**-0.071*** (0.002)**	**-0.080*** (0.002)**	**-0.064*** (0.002)**	**-0.009*** (0.002)**
$\log(Bed_{ijk})$	0.015*** (0.003)	0.023*** (0.003)	-0.036*** (0.003)	0.065*** (0.003)
$\log(Bath_{ijk})$	0.247*** (0.002)	0.267*** (0.003)	0.249*** (0.003)	0.250*** (0.003)
$\log(Size_{ijk})$	0.559*** (0.003)	0.558*** (0.003)	0.611*** (0.003)	0.494*** (0.004)
$\log(Garage_{ijk})$	0.062*** (0.001)	0.061*** (0.001)	0.069*** (0.002)	0.056*** (0.001)
$Renovate_{ijk}$	0.033*** (0.001)	0.029*** (0.001)	0.038*** (0.001)	0.047*** (0.001)
$log(School_{ijk})$	0.237*** (0.003)	0.240*** (0.003)	0.255*** (0.004)	0.231*** (0.003)
$log(Flood_{ijk})$	0.012*** (0.0009)	0.012*** (0.0009)	0.017*** (0.001)	0.004*** (0.0008)
$\log(Fire_{ijk})$	-0.027*** (0.001)	-0.023*** (0.001)	-0.033*** (0.001)	-0.035*** (0.001)
$\log(Heat_{ijk})$	-0.062*** (0.005)	-0.053*** (0.005)	-0.103*** (0.005)	-0.043*** (0.005)
$\log(Wind_{ijk})$	0.082*** (0.009)	0.070*** (0.008)	0.124*** (0.010)	0.066*** (0.008)
$\log(Air_{ijk})$	0.033*** (0.003)	0.030*** (0.003)	0.047*** (0.004)	0.048*** (0.003)
$\log(Population_{ijk})$	-0.047*** (0.002)	-0.046*** (0.002)	-0.054*** (0.002)	-0.027*** (0.002)
House type fixed effects	Yes	Yes	Yes	Yes
Location fixed effects	Yes	Yes	Yes	Yes
N	170,850	170,850	170,850	170,850
R^2	0.9393	0.9381	0.924	0.9252

Note. Standard errors in parentheses are robust and clustered by property. * < 0.1, ** < 0.05, *** < 0.01.

C. Robustness Tests

In this appendix, we consider alternative specifications of our empirical analyses to demonstrate the robustness of our results.

C.1 Model without log-transformations

In our main analysis, we applied log transformation to all numerical variables that are power law distributed. In this robustness exercise, we consider a specification without log transformations. The results, reported in Table C.1, are consistent with our main results.

Table C.1: Model without log transformations.

	$Price_{ijk}$
AI_{ijk} baseline: human-generated prices	-138,713.3*** (1,258.3)
$MarketHighSupply_{ijk}$ baseline: low-supply areas	102,586.3*** (11,947.1)
$MarketNewHouse_{ijk}$ baseline: old houses	-24,926.7*** (4,238.2)
$Underprivileged_{ijk}$ baseline: privileged areas	-128,143.1*** (4,596.6)
$AI_{ijk} \times Underprivileged_{ijk}$	68,337.8*** (2,137.9)
Bed_{ijk}	-15,645.5* (8,999.1)
$Bath_{ijk}$	169,913.6*** (24,685.5)
$Size_{ijk}$	140.9*** (36.0)
$Garage_{ijk}$	94.4 (211.8)
$Renovate_{ijk}$	2,913.5 (2,198.3)
$School_{ijk}$	16,789.7*** (1,720.7)
$Flood_{ijk}$	4,763.4*** (509.6)
$Fire_{ijk}$	-4,029.2*** (824.4)
$Heat_{ijk}$	-18,143.8*** (4,264.6)
$Wind_{ijk}$	27,746.2*** (2,665.6)
Air_{ijk}	5,940.8*** (2,156.9)
$Population_{ijk}$	-71,932.6*** (4,114.2)
House type fixed effects	Yes
Location fixed effects	Yes
N	569,498
R^2	0.5309

Notes. Standard errors in parentheses are robust and clustered by property. * < 0.1, ** < 0.05, *** < 0.01.

C.2 Alternative Generative AI Models

Our main analysis relied on housing prices generated from GPT-4o. In this appendix, we consider alternative GenAI models to showcase that our results are not solely driven by our choice of GenAI model.

C.2.1 GPT 3.5

One of the common concerns for using GPT-4o in our analysis is that the model may generate prices by referencing actual listings from its training data, given the overlap in the observational periods and the model training time (i.e., the listings used in our data are between August and September 2023, and GPT-4o was trained by the data until October 2023). To address this concern, we consider an alternative GenAI model where we use housing prices generated by GPT-3.5 instead.

For efficiency purposes, we conduct this robustness test by randomly selecting 10% of houses from our dataset (i.e., 28,474 properties). For each property, we repeat the same process of using GPT-4o to generate a price, but we use GPT-3.5-turbo-0125 instead for this robustness exercise. We highlight that GPT-3.5 is trained on data obtained until September 2021. Therefore, there is no overlap between the trained data and observations in our dataset. Following that, we also adjust the model-generated prices using the inflation rate of housing markets in the U.S. between 2021 and 2023 from the U.S. Bureau of Labor Statistics to further ensure that our findings are not influenced by inflation. We then repeat our main analysis using GPT-3.5 generated prices and the prices with the inflation adjustment.

Table C.2.1: Housing prices from GPT 3.5

	$\log(Price_{ijk})$	$\log(Price_{ijk})$
	GPT-3.5	GPT-3.5 with Inflation Adjustment
AI_{ijk} baseline: human-generated prices	-0.381*** (0.003)	-0.251*** (0.003)
$MarketHighSupply_{ijk}$ baseline: low-supply areas	0.053*** (0.020)	0.053*** (0.020)
$MarketNewHouse_{ijk}$ baseline: old houses	0.097*** (0.004)	0.097*** (0.004)
$Underprivileged_{ijk}$ baseline: privileged areas	-0.217*** (0.009)	-0.217*** (0.009)
$AI_{ijk} \times Underprivileged_{ijk}$	0.076*** (0.006)	0.076*** (0.006)

$\log(Bed_{ijk})$	0.016** (0.008)	0.016** (0.008)
$\log(Bath_{ijk})$	0.223*** (0.008)	0.223*** (0.008)
$\log(Size_{ijk})$	0.614*** (0.008)	0.614*** (0.008)
$\log(Garage_{ijk})$	0.078*** (0.004)	0.078*** (0.004)
$Renovate_{ijk}$	0.035*** (0.003)	0.035*** (0.003)
$log(School_{ijk})$	0.218*** (0.009)	0.218*** (0.009)
$log(Flood_{ijk})$	0.012*** (0.002)	0.012*** (0.002)
$\log(Fire_{ijk})$	-0.025*** (0.003)	-0.025*** (0.003)
$\log(Heat_{ijk})$	-0.047*** (0.015)	-0.047*** (0.015)
$\log(Wind_{ijk})$	0.171*** (0.026)	0.171*** (0.026)
$\log(Air_{ijk})$	0.035*** (0.010)	0.035*** (0.010)
$\log(Population_{ijk})$	-0.061*** (0.005)	-0.061*** (0.005)
House type fixed effects	Yes	Yes
Location fixed effects	Yes	Yes
N	56,948	56,948
R^2	0.8784	0.8736

Note. Standard errors in parentheses are robust and clustered by property. * < 0.1, ** < 0.05, *** < 0.01.

The results are reported in Table C.2.1. They are qualitatively similar to our main results.

C.2.2 Gemini and Claude

GenAI models in the market exhibit significant heterogeneity in their underlying architectures and training framework. These differences can lead to varying responses to an identical question, raising concerns about whether our findings can be applicable across a spectrum of GenAI models. To address this potential limitation, we conduct a robustness test following the methodology outlined in C.2.1. Specifically, we utilize the same randomly selected 10% of properties from our dataset. We then replicate the price generation process with GPT-4o but instead utilize two other prominent GenAI models: Gemini and Claude.

We choose Gemini 1.5 Flash (gemini-1.5-flash-002) and Claude 3.5 Sonnet (claude-3-5-sonnet-20241022), which are the latest models as of the data collection date (October 28, 2024). Gemini was developed by Google, prioritizing factual correctness in its generated responses (Gemini Team 2024). Meanwhile, Claude was developed by Anthropic, placing significant emphasis on ethical considerations, particularly mitigating bias and discrimination (Bai et al. 2022b). The results are reported in Table C.2.2, which shows qualitatively similar findings to our main analysis.

Table C.2.2: Housing prices from Gemini and Claude

	$\log(Price_{ijk})$ Gemini	$\log(Price_{ijk})$ Claude
AI_{ijk} baseline: human-generated prices	-0.154*** (0.003)	-0.178*** (0.002)
$MarketHighSupply_{ijk}$ baseline: low-supply areas	0.034* (0.019)	0.059*** (0.021)
$MarketNewHouse_{ijk}$ baseline: old houses	0.065*** (0.004)	0.057*** (0.004)
$Underprivileged_{ijk}$ baseline: privileged areas	-0.219*** (0.008)	-0.227*** (0.009)
$AI_{ijk} \times Underprivileged_{ijk}$	0.232*** (0.007)	0.077*** (0.006)
$\log(Bed_{ijk})$	-0.061*** (0.007)	0.010 (0.008)
$\log(Bath_{ijk})$	0.256*** (0.007)	0.243*** (0.008)
$\log(Size_{ijk})$	0.555*** (0.008)	0.594*** (0.009)
$\log(Garage_{ijk})$	0.054*** (0.004)	0.063*** (0.004)
$Renovate_{ijk}$	0.019*** (0.003)	0.034*** (0.004)
$log(School_{ijk})$	0.172*** (0.008)	0.252*** (0.009)
$log(Flood_{ijk})$	0.013*** (0.002)	0.016*** (0.002)
$\log(Fire_{ijk})$	-0.021*** (0.003)	-0.030*** (0.003)
$\log(Heat_{ijk})$	-0.051*** (0.014)	-0.064*** (0.016)
$\log(Wind_{ijk})$	0.143*** (0.024)	0.163*** (0.026)
$\log(Air_{ijk})$	0.033*** (0.009)	0.041*** (0.010)
$\log(Population_{ijk})$	-0.047*** (0.005)	-0.069*** (0.006)
House type fixed effects	Yes	Yes
Location fixed effects	Yes	Yes
N	56,948	56,948
R^2	0.8645	0.8789

Note. Standard errors in parentheses are robust and clustered by property. * < 0.1, ** < 0.05, *** < 0.01.

C.3 Alternative Discrimination Factors

Our primary analysis considers the majority racial composition within a neighborhood as a discriminatory factor and found that GenAI can mitigate housing price discrimination. To enhance the generalizability of our finding, this robustness test explores two alternative discriminatory factors. First, we examine the percentage of non-white residents relative to other racial groups within the neighborhood, identifying areas with a higher percentage as more underprivileged. Second, we consider household income within the neighborhood, classifying areas with an average income below the median income across all zip codes included in the dataset ($88,000) as the underprivileged group.

The results are presented in Table C.3, and they are qualitatively similar to our main result. Thus, we ensure that the conclusion of GenAI mitigating housing price discrimination does not solely rely on our choice of discrimination factor.

Table C.3: Impact of GenAI pricing on housing price discrimination

	Percentage of Minority	Income level
	$\log(Price_{ijk})$	$\log(Price_{ijk})$
AI_{ijk} baseline: human-generated prices	-0.196*** (0.001)	-0.155*** (0.0008)
$MarketHighSupply_{ijk}$ baseline: low-supply areas	0.036*** (0.006)	0.067*** (0.006)
$MarketNewHouse_{ijk}$ baseline: old houses	0.063*** (0.001)	0.054*** (0.001)
$Underprivileged_{ijk}$ baseline: privileged areas	-0.805*** (0.006)	-0.160*** (0.002)
$AI_{ijk} \times Underprivileged_{ijk}$	0.211*** (0.003)	0.053*** (0.001)
$\log(Bed_{ijk})$	0.0006 (0.002)	-0.012*** (0.002)
$\log(Bath_{ijk})$	0.232*** (0.002)	0.238*** (0.002)
$\log(Size_{ijk})$	0.575*** (0.002)	0.581*** (0.003)
$\log(Garage_{ijk})$	0.075*** (0.001)	0.077*** (0.001)
$Renovate_{ijk}$	0.033*** (0.0009)	0.033*** (0.0009)
$log(School_{ijk})$	0.170*** (0.003)	0.234*** (0.003)
$log(Flood_{ijk})$	0.010*** (0.0007)	0.010*** (0.0007)
$\log(Fire_{ijk})$	-0.023*** (0.0009)	-0.030*** (0.0009)
$\log(Heat_{ijk})$	-0.104*** (0.004)	-0.112*** (0.004)
$\log(Wind_{ijk})$	0.129*** (0.008)	0.127*** (0.008)
$\log(Air_{ijk})$	0.041*** (0.003)	0.037*** (0.003)
$\log(Population_{ijk})$	-0.042*** (0.002)	-0.063*** (0.002)
House type fixed effects	Yes	Yes
Location fixed effects	Yes	Yes
N	569,498	569,498
R^2	0.8690	0.8642

References

Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M. 2019. "Optuna: A Next-Generation Hyperparameter Optimization Framework," *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 2623-2631.

Alto, V. 2024. *Building Llm Powered Applications*. Birmingham, UK: Packt Publishing.

Bai, B., Dai, H., Zhang, D. J., Zhang, F., and Hu, H. 2022a. "The Impacts of Algorithmic Work Assignment on Fairness Perceptions and Productivity: Evidence from Field Experiments," *Manufacturing & Service Operations Management* (24:6), pp. 3060-3078.

Bai, Y., Kadavath, S., Kundu, S., Askell, A., Kernion, J., Jones, A., Chen, A., Goldie, A., Mirhoseini, A., and McKinnon, C. 2022b. "Constitutional Ai: Harmlessness from Ai Feedback," *arXiv preprint arXiv:2212.08073*).

Bird, S., Dudík, M., Edgar, R., Horn, B., Lutz, R., Milan, V., Sameki, M., Wallach, H., and Walker, K. 2020. "Fairlearn: A Toolkit for Assessing and Improving Fairness in Ai," *Microsoft, Tech. Rep. MSR-TR-2020-32*).

Breiman, L. 2001. "Random Forests," *Machine learning* (45), pp. 5-32.

Chen, T., and Guestrin, C. 2016. "Xgboost: A Scalable Tree Boosting System," *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785-794.

Corbett-Davies, S., Gaebler, J. D., Nilforoshan, H., Shroff, R., and Goel, S. 2023. "The Measure and Mismeasure of Fairness," *The Journal of Machine Learning Research* (24:1), pp. 14730-14846.

Dorogush, A. V., Ershov, V., and Gulin, A. 2018. "Catboost: Gradient Boosting with Categorical Features Support," *arXiv preprint arXiv:1810.11363*).

Fu, R., Huang, Y., and Singh, P. V. 2021. "Crowds, Lending, Machine, and Bias," *Information Systems Research* (32:1), pp. 72-92.

Gemini Team. 2024. "Gemini 1.5: Unlocking Multimodal Understanding across Millions of Tokens of Context," *arXiv preprint arXiv:2403.05530*).

Johnson, R. 2023. "Building the Neural Zestimate." *AI, Machine Learning & Research*, from https://www.zillow.com/tech/building-the-neural-zestimate/

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. 2017. "Lightgbm: A Highly Efficient Gradient Boosting Decision Tree," *Advances in neural information processing systems* (30).

Liu, Y., Pant, G., and Sheng, O. R. 2020. "Predicting Labor Market Competition: Leveraging Interfirm Network and Employee Skills," *Information Systems Research* (31:4), pp. 1443-1466.

Murtagh, F. 1991. "Multilayer Perceptrons for Classification and Regression," *Neurocomputing* (2:5-6), pp. 183-197.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., and Dubourg, V. 2011. "Scikit-Learn: Machine Learning in Python," *the Journal of machine Learning research* (12), pp. 2825-2830.

Rokach, L., Maimon, O., and Shmueli, E. 2023. *Machine Learning for Data Science Handbook*. Springer.

Vasques, X. 2024. "Feature Engineering Techniques in Machine Learning," in *Machine Learning Theory and Applications*. Wiley, p. 59.

Wang, G., Chen, G., Zhao, H., Zhang, F., Yang, S., and Lu, T. 2021. "Leveraging Multisource Heterogeneous Data for Financial Risk Prediction: A Novel Hybrid-Strategy-Based Self-Adaptive Method," *MIS Quarterly* (45:4).

Watanabe, S. 2023. "Tree-Structured Parzen Estimator: Understanding Its Algorithm Components and Their Roles for Better Empirical Performance," *arXiv preprint arXiv:2304.11127*).

Xu, Y., Ghose, A., and Xiao, B. 2024. "Mobile Payment Adoption: An Empirical Investigation of Alipay," *Information Systems Research* (35:2), pp. 807-828.

Yu, Y., Tan, X., and Tan, Y. 2024. "Understanding Volunteer Crowdsourcing from a Multiplex Perspective," *Information Systems Research*).

Zhang, C., and Ma, Y. 2012. *Ensemble Machine Learning: Methods and Applications*. Springer.
Zillow. 2023. "What Is a Zestimate?", from https://www.zillow.com/z/zestimate/