Racial Difference in Retail Prices Paid*

R. Andrew Butters†

Daniel W. Sacks‡

Boyoung Seo§

September 30, 2022

Abstract

We document that Black and Hispanic households pay 1-2 percent higher prices than white households for physically identical products. This difference suggest that conventional measures of racial income differences understate real racial income inequality. The racial price gap is not explained by differences in income, demographics, or education. Instead it is entirely explained by three factors: black and Hispanic households buy smaller packages with higher unit prices, benefit less from coupons, and live in places where prices tend to be high. The place-based price differences appear driven not by supermarket presence but by differences in carrying and transportation costs.

JEL codes: J15, D63
Key words: racial disparities, real income inequality, retail prices

*We are grateful to Rebecca Diamond and Angus Deaton for conference discussion; Alberto Ortega, Justin Sydnor, Trevon Logan, and participants at the Discrimination and Disparities Online Seminar and the NBER Working Group on Race and Stratification Spring Meeting for helpful feedback; and the Kilts Center for Marketing at The University of Chicago Booth School of Business for granting access to the Nielsen Retail Scan and Homescan data. Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein. Information on availability and access to the data is available at http://research.chicagobooth.edu/nielsen. We acknowledge the Indiana University Pervasive Technology Institute for providing HPC resources that have contributed to the research results reported in this paper (`https://pti.iu.edu`). This project was partially funded by the Vice Provost for Research through the Research Equipment Fund. Boyoung Seo appreciates the feedback, discussion, and financial support from the visiting scholar program in the Research Department of the Federal Reserve Bank of Atlanta's Early Career Visitor Program.

†Business Economics & Public Policy, Kelley School of Business, Indiana University (rabutter@indiana.edu).

‡Risk & Insurance, Wisconsin School of Business, University of Wisconsin - Madison (dan.sacks@wisc.edu).

§Business Economics & Public Policy, Kelley School of Business, Indiana University (seob@indiana.edu).

1 Introduction

In 2020, median income among non-Hispanic white households was $75,000, and among black households it was $47,000 (U.S. Department of Commerce, Bureau of the Census, 2020). This racial income gap has narrowed slowly across generations (Margo, 2016), much more slowly than would be expected given the intergenerational mobility of white Americans (Chetty et al., 2020). Yet as large and as persistent as this gap is, we show that, as typically measured, the income gap between black and white Americans understates racial differences in real income, because black households pay higher retail prices for physically identical goods.

Racial differences in prices paid can arise for multiple reasons. In some contexts such as housing and employment, explicit discrimination contributes to racially disparate outcomes (e.g., Yinger (1986); Christensen and Timmins (2018); Bertrand and Mullainathan (2004); Pena (2018); Kline et al. (2021)). Our context is retail stores with posted prices, where explicit discrimination is unlikely, at least for a given product in a given store. Instead differences in prices can arise for at least two reasons. First, retailers differ substantially in the prices they charge for a given product (Hitsch et al., 2017), and households differ in their access to low-priced supermarkets (e.g., Baker et al. (2006); Allcott et al. (2019)). Second, retailers commonly offer both quantity-based discounts and temporary promotional prices, and households may differ in their ability to take advantage of these low prices, because of differences in wealth, liquidity, or the ability to hold or carry inventory (e.g., Hamilton and Darity Jr (2009); Kuhn et al. (2020); Ganong et al. (2020)).

We investigate racial differences in retail prices paid using transaction-level data from 175,000 households in the Nielsen Consumer Panel. These data record household demographic characteristics as well as the prices paid and quantities purchased at the individual bar code level for all retail purchases. We develop a household-specific relative price measuring how much the household paid for its groceries, relative to what it would pay if it faced average prices but kept its composition the same. The relative price index has a welfare interpretation: by what percent would a household be willing to reduce its grocery spending to pay the average prices, relative to the prices it actually paid for the

products it purchased. It can be aggregated and compared across households, and we use it to measure racial differences in retail prices paid.

We find that non-Hispanic black households pay prices about 2 percent higher than non-Hispanic white households, and Hispanic households pay prices about 0.8 percent higher than non-Hispanic white households. Thus, if black households paid the same prices as white households, but otherwise did not alter the products they purchased, they could purchase 2 percent more groceries. Income and demographics (education, age, family composition) explain little of the racial price gap. A proximate cause is that black households shop at higher price retailers. Overall three factors fully explain it: package size (which explains nearly half the black-white gap), coupon use (a quarter), and location (a quarter). Package size is important because we show that within-product elasticity of price with respect to size of -0.3, and (within-product), black households are more likely than white households to buy smaller packages. White households are also more likely to use coupons than black households.

While location is important for the racial price gap, geographic proximity to supermarkets or convenience stores is not a key factor. Instead, local rates of home and car ownership are especially important for prices paid and the racial price gap. These factors matter not only because transportation and storage is important for bulk purchasing, but especially because they permit access to low price retailers, which likely places downward price pressure on *all* neighborhood retailers, reducing prices paid even by households without car or home ownership. We find that neighborhood car- and home-ownership matter because they are associated with both bulk purchases and shopping at low-price retailers.

We provide suggestive evidence on the mechanisms underlying the importance of package size and location. While location matters for the racial price gap, this is primarily not because of access to supermarkets. Controlling for access to supermarkets makes little difference. Instead, local home and car ownership rates are particularly important for explaining the racial price gap. We interpret these results as evidence that carrying costs and storage costs may be driving forces, because they allow households to stock up on low-price goods, and because local home and car ownership rates place downward

pressure on local prices.

Our results suggest that conventional estimates of the black-white income gap, which do not adjust for differences in prices paid, understate the real income gap, because at any income, the higher prices paid by black households imply lower purchasing power. Deriving the quantitative implications of our results requires an assumption on how representative our products are of the broader universe of retail and non-retail price gaps. Extrapolating to all expenditures, our results imply that real income gaps are six percent larger than conventionally measured gaps. Extrapolating only to non-automobile retail purchases, our results imply real income gaps about 1.7 percent larger than conventionally measured.

Our results primarily contribute to the large literature on racial inequality. This literature has established the magnitude and persistence of racial income inequality (e.g., Margo (2016); Bayer and Charles (2018); Chetty et al. (2020)), and investigated many potentially contributing factors, including among others labor market discrimination (e.g., Darity and Mason (1998); Altonji and Blank (1999); Bertrand and Mullainathan (2004); Pena (2018); Kline et al. (2021)) parental income (e.g., Rothstein and Wozny (2013)), education quality (e.g., Card and Krueger (1992); Neal and Johnson (1996)), neighborhood quality and segregation (e.g., Wilson (2012); Chetty et al. (2020)), tax and transfer policies (e.g., Sullivan and Ziegert (2021)), and cultural factors (e.g., Austen-Smith and Fryer Jr (2005)). Brouillette et al. (2021) develop a measure of relative welfare of black and white Americans that accounts for differences in life expectancy, consumption, leisure, and inequality, but do not account for differential prices.

Within this large literature, our work is especially closely related to recent, innovative work by Avenancio-Leon and Howard (2019)—who document that black and Hispanic households pay higher property tax rates than white households for the same public services, because black houses are assessed at relatively high values—and by Dorsey and Wolfson (2021), who document racial differences in prices paid for solar panel installation, a context where personalized pricing and discrimination are in principle possible. Our findings of differences in retail prices paid for physically identical goods echo their findings, although our context is wholly different, as is our mechanism.

Our work also complements a literature investigating inflation heterogeneity in urban economics and macroeconomics. This literature has investigated how prices vary across income and location (Diamond, 2016; Handbury and Weinstein, 2015; Handbury, 2019), and how inflation varies by income (Kaplan and Schulhofer-Wohl, 2017; Jaravel, 2019, 2021). Our work contributes to this literature in two ways. First, this literature primarily documents price and inflation heterogeneity between higher and lower income households, whereas we focus explicitly on race, and the racial price inequality we document is distinct from income-based price inequality. Second, this literature constructs price indices for different groups to reflect their different consumption patterns, finding that heterogeneous inflation is driven by different inflation rates for different products. In our context, the price inequality we document is driven by different prices paid for *identical* products.

2 The Nielsen Consumer Panel

We use Nielsen Consumer Panel (HMS) data from 2006 to 2018. Panelists use scanners to record all of their retail purchases, transaction by transaction. In 2006 the Panel consisted of about 40,000 households, and since 2007 it has included about 61,000 households annually. The sample is refreshed annually. Panelists are selected according to a stratified random sampling scheme so that, properly weighted, the panel is representative at the National level, and also representative of each of 52 major markets (defined by Nielsen).

The HMS data include transactions in 10 broad categories which Nielsen calls departments. The departments include food, non-food grocery, as well as the non-grocery departments "Health and Beauty Aids" and "General Merchandise." The information from each transaction includes the exact barcode purchased (i.e., the UPC), quantity, coupon value, and store identifier. Nielsen does not release store level identifiers, but does provide encrypted store and chain identifiers, as well as information on retail channel (example categories include mass merchandise, grocery, convenience, bodega, and warehouse club). Prices are recorded in two ways. If the transaction occurs at a store that provides data to Nielsen's Retail Scanner database, Nielsen imputes the price paid as the weekly

average price paid for that UPC. If the store is not part of the Retail Scanner database, panelists are asked to report the price they paid. We work with prices net of any discounts reported by panelists. These discounts include discounts applied at the register (such as loyalty cards or store-run "buy one get one free promotions") as well as other discounts (such as manufacturer's coupons).

In most of our analysis we focus on the price per unit of commonly purchased products. Following Hitsch et al. (2017), we define a product by its physical features. Specifically, we group UPCs into products that share identical brand and UPC descriptions. Within a product and across UPCs, all variation is in package size (6 vs. 36 pack) or package material (e.g., can vs. bottle). For example, a 6 pack of Scott Unscented Toilet Paper and a 36 pack of Scott Toilet Unscented Toilet Paper would be the same product, but different UPCs with manufacturers (Charmin) or descriptions (Scott's Comfort Plus) would be different products.[1] To accommodate private label brands (which share product descriptions across retails), we define private label products as descriptions-by-retailers. Our results are robust to omitting private labels. Nielsen provides standard units at our product level (e.g., counts of rolls of toilet paper), and for each transaction, we obtain the price per unit by dividing by unit size.

Households report demographic information annually, including information on race/ethnicity, as well as income, age, education, and household composition. We use this information to define four mutually exclusive and exhaustive race/ethnicity categories: non-Hispanic Black, non-Hispanic White, Hispanic, and all other (including unreported). For simplicity we sometimes refer to these categories as "race" and we sometimes refer to white and black (omitting the "non-Hispanic" qualifier). Income, age, and education are reported in binned categories. When multiple heads are present, we take the maximum age and education reported for each.

Our analysis sample consists of all participating households, except for a small number of households who never reported any grocery products. Our final sample consists of 523 million product-level purchases made by 175,440 households. Summary statis-

[1] Hitsch et al. (2017) call this grouping "brand," but we call it "product" to emphasize that the grouping pools physically identical products.

tics are given in Appendix Table A.1. The racial composition of our sample is similar to that of the US as a whole (U.S. Department of Commerce, Bureau of the Census, 2017), although White non-Hispanic Americans are somewhat overrepresented, and Hispanic Americans somewhat unrepresented. Total spending per year on our focal products is just over $3,500 per household, meaning in aggregate the households in our sample spent about $184 million annually on the products we study.

3 Measuring racial price differences

We develop a microfounded, household-level relative price that summarizes price difference paid by different households and adjusts for detailed differences in the composition of purchases. The relative price measures the prices a given household pays for its groceries, relative to the average price paid for those groceries in all transactions.

3.1 Model

Consider a household purchasing with income y facing prices p and purchasing the utility maximizing bundle of goods x. Let dp_j be the difference between the price the household pays for product j, p_j, and the average price paid, $\bar{p}_j$. If dp_j is not too large, then by the envelope theorem, the change in utility is $\lambda x_j dp_j$ (where λ is the marginal utility of income) and the willingness to pay for such a price change, WTP_j, is

$$WTP_j = x_j dp_j = x_j p_j \frac{dp_j}{p_j} \approx x_j d\ln p_j. \tag{1}$$

Aggregating across a vector of small price changes dp, the willingness to pay is

$$WTP(dp) \approx \sum_j x_j p_j d\ln p_j. \tag{2}$$

Scaling by total expenditures e, we have:

$$\frac{WTP}{e} \approx \sum_j \frac{x_j p_j}{e_i} = \sum_j s_j d\ln p_j \equiv w(dp) \tag{3}$$

In words, $w(dp)$ equals the expenditure-share weighted average change in log prices induced by dp. It is approximately the household's willingness to pay (as a fraction of its expenditures) to face the average prices paid instead of the prices it actually paid. We call $w(dp)$ the relative price. This price varies in the population for two reasons: the willingness to pay for any set of price changes depends on preferences, and dp varies with actual prices paid. $w(dp)$ is a money metric, so we can aggregate across households and compare at the individual or group level.[2]

3.2 Operationalization

We measure the relative price $w(dp)$ at the household-year level. Doing so requires that we measure, for each household i, product j, and year y, expenditure shares s_{ijy} and log price deviations $d\ln p_{ijy}$. We measure expenditures shares as the spending by i on j in y as a share of its total expenditures among our commonly purchased products.

To measure $d\ln p_{ijy}$, we start at the transaction level with $\ln p_{iujty}$, the log price per unit paid by household i for UPC u and product j in transaction t and year y. Here u refers to the most detailed product classification (including packaging and size), whereas product j refers to our coarser product definition which aggregates over packaging and size. We work with prices net of coupon discounts to accommodate price differences from coupons. We define

$$d\ln p_{iujty} = \ln p_{iujty} - \overline{\ln p}_{jy},$$

where $\overline{\ln p}_{jy}$ is the quantity-weighted average log unit price paid for product j across all transactions in year y for that product. So $d\ln p_{ijty}$ measures how much more or less i paid per unit of j in transaction t, relative to the average payment in y. We aggregate across transactions to obtain the household-product price deviation $\overline{d\ln p}_{ijy}$, defined as the quantity-weighted average across transactions (within household-product-year) of $d\ln p_{iujty}$. We then aggregate across products, following the model, to obtain our baseline

[2]When we aggregate across households, we typically weight each household equally, although it would be more appropriate to weight them by their expenditures, and we show robustness to expenditure weighting.

price w_{iy}:

$$w_{iy} = \sum_{jt} s_{ijy} \overline{d \ln p}_{ijy}. \tag{4}$$

w_{iy} is equal to the expenditure share weighted average (log) difference in prices paid by i and average prices paid for the same products in year t. It is also the percent increase in expenditures of household i and year y from the price it paid, relative to average prices, at fixed purchases. Under utility maximization w_{iy} is approximately household i's willingness to pay to face average prices instead of paid prices, as a share of expenditures.

Our baseline measure, w_{iy}, measures household-year level prices relative to product-year averages of price net of coupon discounts. To investigate heterogeneity across departments (i.e. broad product categories), we also construct department-specific price indices, in which case we define expenditure shares within department. In some analyses we also measure prices relative to product-*size*-year averages, to show the importance of package size; in others we look at prices gross of coupon discounts.

A natural concern with comparing prices paid across households is that prices can differ with quality, and we might expect higher-income households to select higher quality varieties within broad product categories. To the extent that race is correlated with income, this would imply that price paid would would differ systematically across races. This concern is unlikely to be a problem in our context, for two reasons. First, our price index is based differences in prices paid for physically identical products, and so avoids comparing prices across different quality levels. Second, if present, the bias from differential quality would imply that we find lower prices for black households, opposite to our actual result.

While quality differences across household do not lead to variation in our price, several other factors do. We highlight four potentially important factors. First, we show below that households pay lower unit prices if they buy larger package sizes. Second is retail channel: super markets, drug stores, convenience stores, and mass merchandise stores differ in both package size availability and prices of a given size potentially vary with retail channel. Third, households may differ in their tendency to use coupons or loyalty cards. Fourth, for a given store and package size, prices vary over time as promo-

tional pricing becomes available. Households may differ in their tendency to stock up on products during promotional pricing periods. w_{iy} captures all these differences.

Savings from package size, coupons, and retailers: We measure household savings or excess spending from package size, coupons, and retailer choice. To measure package size savings, we start by re-defining w_{iy} at the product-*size* level. Specifically at the transaction level we define

$$d\ln p^{size}_{iujty} = \ln p_{iujty} - \overline{\ln p}_{sjy},$$

where $\ln p_{iujty}$ is the log price paid (as before) and $\overline{\ln p}_{sjy}$ is the average log price paid, averaging over not all UPCs in the product but over all UPC-sizes in the product-size. We aggregate across the households purchase to obtain $\overline{d\ln p}^{size}_{ijy}$, and then we then define w^{size}_{iy} as

$$w^{size}_{iy} = \sum_{jt} s_{ijy}\overline{d\ln p}^{size}_{ijy}. \tag{5}$$

w^{size}_{iy} measures, in percent terms, how much more or less i paid than the average households, for the product-size combinations it actually purchased. If a household pays low price only because it buys large package sizes, then w_{iy} will be negative but w^{size}_{iy} will be positive. The savings from package size is therefore $w^{size}_{iy} - w_{iy}$.

We take similar steps to measure the savings or excess spending from coupon use and retailer choice. We define w^{gross}_{iy} identical to w_{iy}, except instead of using actual prices we use prices gross of coupon discounts. Then the savings from coupons is $w_{iy} - w^{gross}_{iy}$. Likewise we define $w^{retailer}_{iy}$ identical to w_{iy} but netting out the product-year-*retailer* mean price. Our size-, coupon- and retailer-based discount measures are not additive. If certain retailers offer large package sizes and low prices at a given package size, then households who use those retailers will exhibit substantial size- and retailer based discounts, but their overall savings relative to the average household will be less than the sum of the retailer and size-based savings. We use a decomposition analysis to account for the overlap across these savings modes.

Measuring racial price gaps: To measure racial price gaps, we use regression to aggregate w_{iy} across households and to adjust for observed differences. Specifically, we

estimate

$$w_{iy} = \beta_0 + \beta_1 Black + \beta_2 Hispanic + \beta_3 OtherRace + X_{iy}\theta + \epsilon_{iy}. \quad (6)$$

The omitted race-ethnicity is non-Hispanic white, and we define Black as non-Hispanic. Our interest is in β_1 and β_2, which we refer to as the black-white and Hispanic-white price gaps. Our baseline measure of these gaps omits all controls and pools all years. In some specifications we stratify on year, income, department, or state. In others we adjust for covariates X_{iy}: income (indictors for each level of self-reported income), demographics (indicators for age, number of adults in the household, number of children, marital status, and educational attainment), area fixed effects, or area characteristics.

Discussion: The construction of our relative price index balances two competing objectives. On the one hand, our model implies that we should focus on price differences between products that are perfect substitutes in consumption. Thus we measure prices paid at the detailed product level. This level of detail is important because compositional differences can spuriously generate price differences. These composition differences can arise for example from the fact higher income households buy higher quality goods, even in narrowly defined product categories (Broda et al., 2009; Handbury and Weinstein, 2015). On the other hand, compositional differences across households could reflect differences not in preferences but in prices or availability. Adjusting too extensively for composition effects might therefore mask some price differences. Our measure balances these concerns in a conservative way. In particular, we control for very detailed product characteristics, leaving primarily store, package size, and discount-pricing related factors as the main sources of price variation for a given product-year. We think it is unlikely that preferences for store, packaging, or size differ systematically with race. While discrimination and racial profiling may differ across stores (Gabbidon and Higgins, 2007; Pittman, 2020), the underlying preference to avoid being profiled while shopping is likely common across race/ethnicity, so racial differences in store preference driven by the desire to avoid profiling reflect differences in choice sets rather than preferences. It is possible that some of the observed differences in the composition of purchased products reflects differences in prices (or availability) rather than differences in preferences. If so,

our estimates understate racial price differences.

4 Documenting and explaining racial differences in prices paid

4.1 Racial price differences

We report racial differences in prices paid in Table 1. The first column reports the average differential household price index for black and Hispanic households, relative to white households. Black households pay an average of 2 percent more for their products than they would pay if they faced the same prices as white households, and Hispanic households pay 0.8 percent higher prices.[3]

The black-white price gap is ubiquitous in our data, appearing across the income distribution, in all product categories, in every year, and in most states. The Hispanic-white price gap is common but not as widespread. In Figure 1 we report the average relative price in each income and race category. At each income level, black and Hispanic households pay higher prices than white households. The average within-income difference is 1.9 percent for black households and 0.8 percent for Hispanic households, about the same as the overall difference. The within-income racial price difference is similar to the overall difference, despite racial income differences, because prices are nonlinearly and only modestly associated with income.[4]

The racial price gap is ubiquitous. The black-white price gap is positive in each department and statistically significant in 9 of the 10 (Appendix Figure A.1. Importantly, we estimate a fairly similar price gap in non-food categories (general merchandise, non-food grocery, health & beauty) as in food categories, implying that the package size effects we document below are not limited to food products where package size might affect con-

[3]Here we pool race/ethnic groups for parsimony and power. Appendix Table A.2 reports differences across the most detailed race-ethnicity cells. The Hispanic and Black effects are roughly additive.

[4]Figure 1 shows that prices fall as income rises form the lowest levels in the data. In contrast, Broda et al. (2009) find that prices rise with income. The difference between our findings is that Broda et al. (2009) condition on package size, whereas we do not, and the higher income customers enjoy larger package size discounts.

sumption utility. For Hispanic households the racial price gap is smaller and not always positive; it is positive and significant in seven of 10 departments, and negative (and significant) in three. The racial price gap appears in each year of our data (Appendix Figure A.2), and the black-white gap appears in most states (Appendix Figure A.3; the Hispanic-white gap is not evidence in all states). It also appears when we exclude private label products, or warehouse club purchases (Appendix Table A.3, column (2) and (3)); it even appears within retail *channel* (such as discount store or grocery store; Appendix Table A.3, column (4)).

4.2 Explanations

We investigate possible explanations for this price gap in the remaining columns of Table 1. We begin by ruling out differences in income or demographic characteristics. In column (2) of Table 1 we present price differences that adjust income and demographics - indicators for each income level, as well as age of head, marital status, number of children, and households size, as well as indicators for each level of educational attainment. The age and household composition controls are meant to adjust for the fact that grocery demand and ability to buy in bulk likely differ with household characteristics. The educational controls address the fact that higher-educated shoppers are more likely to purchase generic brands (Bronnenberg et al., 2015) and, in general, may be more informed shoppers. These adjustments make little difference to the estimated racial difference in price pays: after controlling for age, family composition, and education (as well as income), the estimated black-white income gap falls to 1.8 percent, and the Hispanic-white price gap actually increases to 1.1 percent.

In the next column we show an important proximate cause of racial price gaps: differences in prices by retailer. Specifically, we control for "retailer discounts", defined as the amount of its overall relative price due to shopping at below- or above-average price retailers. Controlling for retailer discounts reduces the black-white gap to 1 percent and the Hispanic-white gap to 0.3 percent. We view these retailer discounts as proximate rather than ultiamte causes for two reasons. First, retailers offer savings in heterogeneous ways

(for example, through every day low prices, or through occasional large price reductions). Second, households' ability to take advantage of low-price retailers reflects a combination of transportation and housing market opportunities.

We probe these mechanisms more deeply in the remaining column, which drop the adjustment for retailer discounts. We show in particular that while income and demographic explain little of the racial price gap, three factors combine to explain the entirety of it: package size, couponing, and location. Package size discounts are a pervasive feature of retailing. For example, in our data we estimate an elasticity of unit price with respect to package size of -0.33, meaning for a fixed product, doubling the package size reduces the unit price by about a third, as we illustrate in Panel A of Figure 2. However, black households tend to buy smaller package sizes than white households, as can be seen from the race-specific distribution of package size (relative to product mean) plotted in Panel B of the figure.[5]

We show in column (4) of Table 1 the importance of package size for the black-white price gap. After adjusting for income and demographics, further controlling for package size-based discounts reduces the black-white price gap falls by half and the Hispanic-white price gap falls by about 20 percent (0.002/0.011). Controlling for coupon-based discounts further reduces both price gaps by about half.[6]

The final important factor is location. Location effects are potentially important because prices and economic circumstances—including retailer entry—vary systematically across places (e.g. Chetty et al. (2014); Handbury and Weinstein (2015); Diamond (2016)), and there are longstanding, strong racial sorting patterns in the United States, even conditional on income (e.g. Logan and Parman (2017); Bayer et al. (2021)). We show the importance of location in column (6) by adding a full set of zipcode fixed effects. Doing

[5]To construct this figure, we first calculate product level mean log size. Then for each barcode we calculate log size relative to its product-level mean. Pooling all bar codes, we calculate deciles of the residual size distribution, and we bin the data into these deciles. Then using the transaction level data we calculate the average log price (residualized net of product-year mean, as in our main analysis), and the race-specific transaction share, in each decile of residual package size.

[6]Although our discount measures are mechanically correlated with our dependent variable, they need not be conditionally correlated with race, and hence it is not guaranteed that including them reduces the racial price gap. Alternative approaches—looking at net-of-coupon relative or within-size relative as the dependent variable—yield similar results (Appendix Table A.3, columns (4) and (5)). We focus on this approach because our decomposition requires changing regressors rather than dependent variables.

so reduces both price gaps to essentially zero.

Zip code fixed effects potentially reflect a large bundle of attributes including access to stores, wealth, and ability to carry and store groceries, all of which could potentially influence retail prices paid. To investigate the importance of these individual attributes, we re-estimate the racial price gap, but adjusting for zip code-level characteristics rather than fixed effects, individually and then jointly. We report these results in Table 2. Our observables are car ownership rates, home ownership rates, supermarket and drug stores per 1,000 residents, zip code median income (in $10,000s), and zip code home prices ($100,000s). Each of these characteristics is strongly associated with race, even after adjusting for income (Appendix B).

We find a key role for car and home ownership, with store proximity less important. Columns (1) and (2) of Table 2 reproduce the final columns of Table 1, showing the importance of zip code fixed effects. Columns (3) and (4) show that controlling for zip-code car or home ownership alone also explains nearly all of the racial price gap, as these variables are strongly conditionally associated with race and prices. Controlling for supermarket and convenience store presence makes less of a difference because these variables are not as strongly conditionally associated with price. A 1-standard deviation increase in car ownership is associated with a 4-12 times larger fall in prices than is a 1-standard deviation in supermarket presence. Income and house prices are weakly, positively associated with price.

The results so far show that package size, coupons, and location—in particular, car and home ownership—collectively explain the racial price gap, but they do not show the individual importance of each factor. To do so, we use the method of Gelbach (2016), which decomposes the change in a regression coefficient from the addition of many regressors into parts due to each added regressor. We report in Table 3 the decomposition of the change in the racial price gap between its unadjusted value (column (1) of Table 1) and its value after adjusting for income, demographics, size, discounts, and location observables (column (8) of Table 2). We show decompositions not adjusting and adjusting for retailer-specific discounts. Not adjusting for discounts, almost half the black-white price gap is explained by package size discounts; coupons and location each explain about a

quarter. For the Hispanic-white price gap, size is less important, explaining 10 percent of the gap; coupon discounts explain 47 percent and location-level factors 79 percent.[7] Adjusting for retailer discounts, we find these discounts explain 37-40 percent of racial price gaps, and the importance of package size and, especially, location falls. Thus zip-code level observables—car and home ownership—matter in large part because they reflect the presence of low-price retailers offering bulk discounts.

5 Discussion and conclusions

Overall we document that black households pay prices 2.0 percent higher, and Hispanic households 0.8 percent higher, than white households, holding fixed the composition of products purchased. These higher prices are driven not by differences in income, education, or household size. Instead, the higher prices are explained by the facts that black and Hispanic households disproportionately buy small package sizes with high unit prices, benefit less from coupons, and live in areas where all people pay higher prices, primarily because they live in areas with low car and home ownership that end up with less access to low-price retailers.

None of these three factors—size-based discounts, coupon use, or location—is an ultimate explanation, since each ultimately reflects potentially endogenous choices. We find suggestive evidence that area-wide home and car ownership in particular, and likely storage and transportation costs in general, are important contributing factors to the racial price gap. Greater access to storage and transportation makes it easier to travel to—and stock up at—low price stores, as well as to buy in bulk and take advantage of coupons when available. These benefits can extend beyond individual car and home owners, and extend beyond coupon/bulk purchase discounts, because neighborhood-level home ownership and car ownership can have equilibrium effects, lowering prices for all households through greater retail competition. Our results on the importance of transportation and storage are however only suggestive because they rely on cross-sectional correlations.

[7] These factors explain more than 100 percent of the Hispanic-white price gap because the unadjusted gap is positive but the adjusted gap is negative, so the change is greater than 100 percent of the adjusted gap.

We view it as an important task for future work to use quasi-experimental variation to down the importance of cars, or other factors in general.[8]r

These price differences are likely welfare relevant. We say this, first, because it is unlikely that households differ in the utility they derive from package sizes; black households are unlikely to disproportionately prefer small packages. Second, it is also unlikely that the high retail prices experienced by black households are made up for by low prices elsewhere. We show directly that the black-white price gap is not driven by differential use of warehouse clubs. It is possible in principle that black households pay higher retail prices but lower housing prices in exchange. This, too, appears unlikely: relative to white households, black households have fewer location options, and hence less ability to tradeoff amenities for house prices. For example, as Bayer et al. (2021) show, in most US cities, it is difficulty "to choose a neighborhood that simultaneously provides even moderate levels of both median income and the share of Black neighbors" (p. 12). Ongoing discrimination in the housing market (as documented by Christensen and Timmins (2018, 2021)) makes it especially unlikely that higher retail prices paid by black households reflect different choices along the same amenity-housing price trade-off faced by white households. Further, this hypothesis is inconsistent with our finding that access to supermarkets, or retail channel more generally, statistically explains for little of the racial price gap.

This difference in prices paid exacerbates the substantial differences in household incomes by race in determining real racial income inequality. For example, in 2020, median income among non-Hispanic white households was $74,912, and among black households it was $46,600 (U.S. Department of Commerce, Bureau of the Census, 2020), a difference of 47 log points. If our estimate—which derives from commonly purchased retail products only—is reflective of *all* differences in prices paid across all goods and services, then differences in prices paid increase racial income inequality by about 4 percent. Extrapolating, more conservatively, to non-automobile retail spending in general,

[8]It does not have to be true that increases in transportation or storage reduce equilibrium prices. Eizenberg et al. (2021) show it is possible for local prices to rise when spatial frictions fall. The intuition is that when frictions fall, customers with elastic demand become less important for local retailers, and customers with inelastic demand more important, so the local retailers end up facing relatively less elastic demand.

which accounts for about a third of all household spending (calculated from U.S. Bureau of Labor Statistics (2021)), our price differences imply that conventional estimates of the black-white income gap understate real income inequality by about 1.3 percent.

References

Allcott, H., B. B. Lockwood, and D. Taubinsky (2019). Regressive sin taxes, with an application to the optimal soda tax. *The Quarterly Journal of Economics 134*, 1557–1626.

Altonji, J. G. and R. M. Blank (1999). Race and gender in the labor market. *Handbook of labor economics 3*, 3143–3259.

Austen-Smith, D. and R. G. Fryer Jr (2005). An economic analysis of "acting white". *The Quarterly Journal of Economics 120*(2), 551–583.

Avenancio-Leon, C. and T. Howard (2019). The assessment gap: Racial inequalities in property taxation. *Available at SSRN 3465010*.

Baker, E. A., M. Schootman, E. Barnidge, and C. Kelly (2006). The role of race and poverty in access to foods that enable individuals to adhere to dietary guidelines. *Preventing chronic disease 3*(3).

Bayer, P. and K. K. Charles (2018). Divergent paths: A new perspective on earnings differences between black and white men since 1940. *The Quarterly Journal of Economics 133*(3), 1459–1501.

Bayer, P., K. K. Charles, and J. Park (2021). Separate and unequal: Race and the geography of the american housing market.

Bertrand, M. and S. Mullainathan (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American economic review 94*(4), 991–1013.

Broda, C., E. Leibtag, and D. E. Weinstein (2009). The role of prices in measuring the poor's living standards. *Journal of economic Perspectives 23*(2), 77–97.

Bronnenberg, B. J., J.-P. Dubé, M. Gentzkow, and J. M. Shapiro (2015). Do pharmacists buy bayer? informed shoppers and the brand premium. *The Quarterly Journal of Economics 130*(4), 1669–1726.

Brouillette, J.-F., C. I. Jones, and P. J. Klenow (2021). Race and economic well-being in the united states. Technical report, National Bureau of Economic Research.

Card, D. and A. B. Krueger (1992). School quality and black-white relative earnings: A direct assessment. *The quarterly journal of Economics 107*(1), 151–200.

Cattaneo, M. D., R. K. Crump, M. H. Farrell, and Y. Feng (2019). On binscatter. *arXiv preprint arXiv:1902.09608*.

Cattaneo, M. D., R. K. Crump, M. H. Farrell, Y. Feng, and M. Y. Feng (2021). Package 'binsreg'.

Chetty, R., N. Hendren, M. R. Jones, and S. R. Porter (2020). Race and economic opportunity in the united states: An intergenerational perspective. *The Quarterly Journal of Economics 135*(2), 711–783.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics 129*(4), 1553–1623.

Christensen, P. and C. Timmins (2018). Sorting or steering: Experimental evidence on the economic effects of housing discrimination. Technical report, National Bureau of Economic Research.

Christensen, P. and C. Timmins (2021). The damages and distortions from discrimination in the rental housing market. Technical report, National Bureau of Economic Research.

Darity, W. A. and P. L. Mason (1998). Evidence on discrimination in employment: Codes of color, codes of gender. *Journal of Economic Perspectives 12*(2), 63–90.

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review 106*(3), 479–524.

Dorsey, J. and D. Wolfson (2021). Income and racial disparities in markets for new energy technologies. Technical report, Mimeo, Indiana University.

Eizenberg, A., S. Lach, and M. Oren-Yiftach (2021). Retail prices in a city. *American Economic Journal: Economic Policy 13*(2), 175–206.

Gabbidon, S. L. and G. E. Higgins (2007). Consumer racial profiling and perceived victimization: A phone survey of philadelphia area residents. *American Journal of Criminal Justice 32*(1), 1–11.

Ganong, P., D. Jones, P. J. Noel, F. E. Greig, D. Farrell, and C. Wheat (2020). Wealth, race, and consumption smoothing of typical income shocks. Technical report, National Bureau of Economic Research.

Gelbach, J. B. (2016). When do covariates matter? and which ones, and how much? *Journal of Labor Economics 34*(2), 509–543.

Hamilton, D. and W. Darity Jr (2009). Race, wealth, and intergenerational poverty: There will never be a post-racial america if the wealth gap persists. *The American Prospect 20*(7), A10–A12.

Handbury, J. (2019). Are poor cities cheap for everyone? non-homotheticity and the cost of living across us cities. Technical report, National Bureau of Economic Research.

Handbury, J. and D. E. Weinstein (2015). Goods prices and availability in cities. *The Review of Economic Studies 82*(1), 258–296.

Hitsch, G. J., A. Hortaçsu, and X. Lin (2017). Prices and promotions in U.S. retail markets: Evidence from big data. Working Paper 17-18, Chicago Booth Research Paper.

Jaravel, X. (2019). The unequal gains from product innovations: Evidence from the us retail sector. *The Quarterly Journal of Economics 134*(2), 715–783.

Jaravel, X. (2021). Inflation inequality: Measurement, causes, and policy implications. *Annual Review of Economics 13*, 599–629.

Kaplan, G. and S. Schulhofer-Wohl (2017). Inflation at the household level. *Journal of Monetary Economics 91*, 19–38.

Kline, P. M., E. K. Rose, and C. R. Walters (2021). Systemic discrimination among large us employers. Technical report, National Bureau of Economic Research.

Kuhn, M., M. Schularick, and U. I. Steins (2020). Income and wealth inequality in america, 1949–2016. *Journal of Political Economy 128*(9), 3469–3519.

Logan, T. D. and J. M. Parman (2017). The national rise in residential segregation. *The Journal of Economic History 77*(1), 127–170.

Margo, R. A. (2016). Obama, katrina, and the persistence of racial inequality. *The Journal of Economic History 76*(2), 301–341.

Neal, D. A. and W. R. Johnson (1996). The role of premarket factors in black-white wage differences. *Journal of political Economy 104*(5), 869–895.

Pena, A. A. (2018). Skills and economic inequality across race and ethnicity in the united states: New evidence on wage discrimination using piaac. *The Review of Black Political Economy 45*(1), 40–68.

Pittman, C. (2020). "shopping while black": Black consumers' management of racial stigma and racial profiling in retail settings. *Journal of Consumer Culture 20*(1), 3–22.

Rothstein, J. and N. Wozny (2013). Permanent income and the black-white test score gap. *Journal of Human Resources 48*(3), 510–544.

Sullivan, D. H. and A. L. Ziegert (2021). Family poverty in black and white: Results from a new poverty measure. *The Review of Black Political Economy 48*(4), 410–454.

U.S. Bureau of Labor Statistics (2021). Average annual expenditures and characteristics of all consumer units, consumer expenditure surveys, 2013-2020. Technical report, National Bureau of Economic Research.

U.S. Department of Commerce, Bureau of the Census (2017). Table dp05: Acs demographic and housing estimates: 2017 acs 5-year estimates data profile. Technical report.

U.S. Department of Commerce, Bureau of the Census (2020). Historic income tables: Households; table h-5. race and hispanic origin of householder – households by median and mean income. Technical report.

Wilson, W. J. (2012). *The truly disadvantaged: The inner city, the underclass, and public policy*. University of Chicago Press.

Yinger, J. (1986). Measuring racial discrimination with fair housing audits: Caught in the act. *The American Economic Review*, 881–893.

Table 1: Documenting and explaining racial price differences

Controls	None (1)	+Demographics (2)	+Retailer (3)	+Size (4)	+Coupons (5)	+Zip FE (6)
Black	0.020	0.018	0.010	0.008	0.005	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	0.008	0.011	0.003	0.008	0.005	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
# Households	175,440	175,440	175,440	175,440	175,440	173,897
Controls						
Demographics		Yes	Yes	Yes	Yes	Yes
Retailer discount			Yes	No	No	No
Size discount			Yes	Yes	Yes	Yes
Coupon discount				Yes	Yes	Yes
Zip code FE						Yes

Notes: The dependent variable is the household-level relative price, relative to the product-year mean; see Section 3 for details. Table reports the coefficients on indicators for non-Hispanic black and Hispanic (mutually exclusive; the omitted category is non-Hispanic white, and we also include an "all other race/ethnicity" category). Demographic controls are indicators for binned income amounts,educational attainment, household size, married, age of head, and number of children; location controls are fixed effects as indicated . Robust standard errors, clustered on household, in parentheses.

Table 2: Car and home ownership, more than supermarket proximity, are why location influences the racial price gap

Zip code Controls	None	Fixed Effects	Cars owned	Homes owned	Stores	Average Income	House Prices	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	0.005	0.000	-0.003	0.001	0.003	0.007	0.007	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	0.005	0.000	0.001	0.002	0.005	0.005	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Car ownership			-0.131					-0.113
			(0.004)					(0.004)
			[-0.014]					[-0.012]
Home ownership				-0.039				0.021
				(0.002)				(0.002)
				[-0.006]				[0.003]
Supermarkets					-0.039			-0.021
					(0.004)			(0.003)
					[-0.003]			[-0.001]
Convenience					0.010			0.005
					(0.001)			(0.001)
					[0.004]			[0.002]
Median income						0.002		-0.000
						(0.000)		(0.000)
						[0.005]		[-0.000]
Median house price							0.010	0.008
							(0.000)	(0.000)
							[0.014]	[0.012]
Adjusted R^2	0.459	0.579	0.484	0.464	0.462	0.462	0.486	0.502
# Households	171,122	171,122	170,825	170,832	170,830	170,819	170,326	170,325

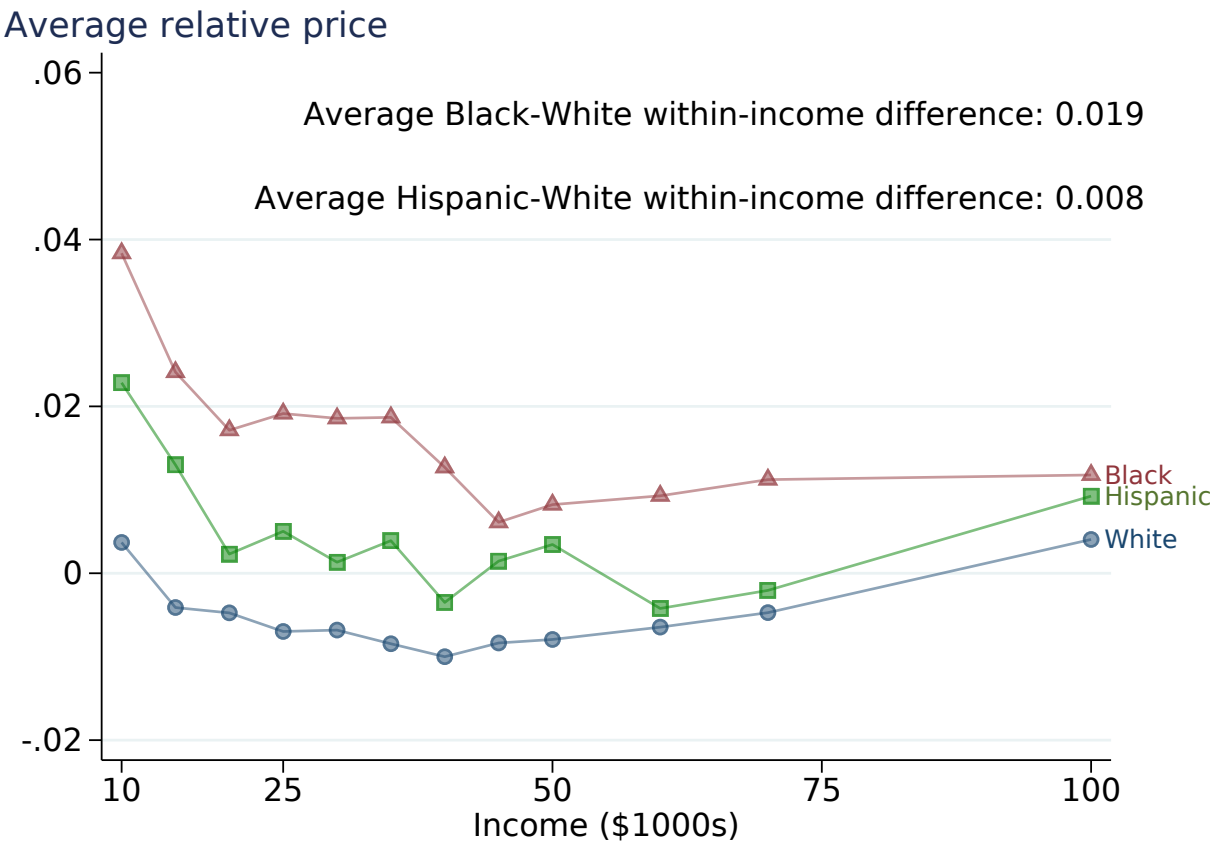
Notes: Table reports the coefficient on the indicated variable. The number in brackets is the implied effect of a 1-SD change in the indicated independent variable. The dependent variable is the household-level relative price, product-year mean; see Section 3 for details. All columns control for income (bins of income amount), demographics (indicators for educational attainment, household size, married, age of head, and number of children), coupon discounts, and size discounts. Column (2) controls for zip code fixed effects and columns (3)-(8) the indicated zip code level observables. Robust standard errors, clustered on household, in parentheses. The number in brackets is the implied marginal effect of a 1-SD increase in the indicated variables.

Table 3: Decomposition of the racial price gap

	No retailer adjustment		Retailer adjustment	
Gap:	Black-White	Hispanic-White	Black-White	Hispanic-White
	(1)	(2)	(3)	(4)
Income	-0.011	0.021	-0.006	0.028
	(0.002)	(0.008)	(0.001)	(0.004)
Demographics	0.003	-0.052	0.003	0.018
	(0.003)	(0.017)	(0.002)	(0.011)
Package size	0.477	0.100	0.203	0.043
	(0.020)	(0.053)	(0.011)	(0.022)
Coupon discounts	0.231	0.440	0.159	0.303
	(0.023)	(0.061)	(0.016)	(0.042)
Zip code characteristics	0.264	0.736	0.031	0.208
	(0.020)	(0.046)	(0.008)	(0.015)
Retailer			0.371	0.411
			(0.031)	(0.086)

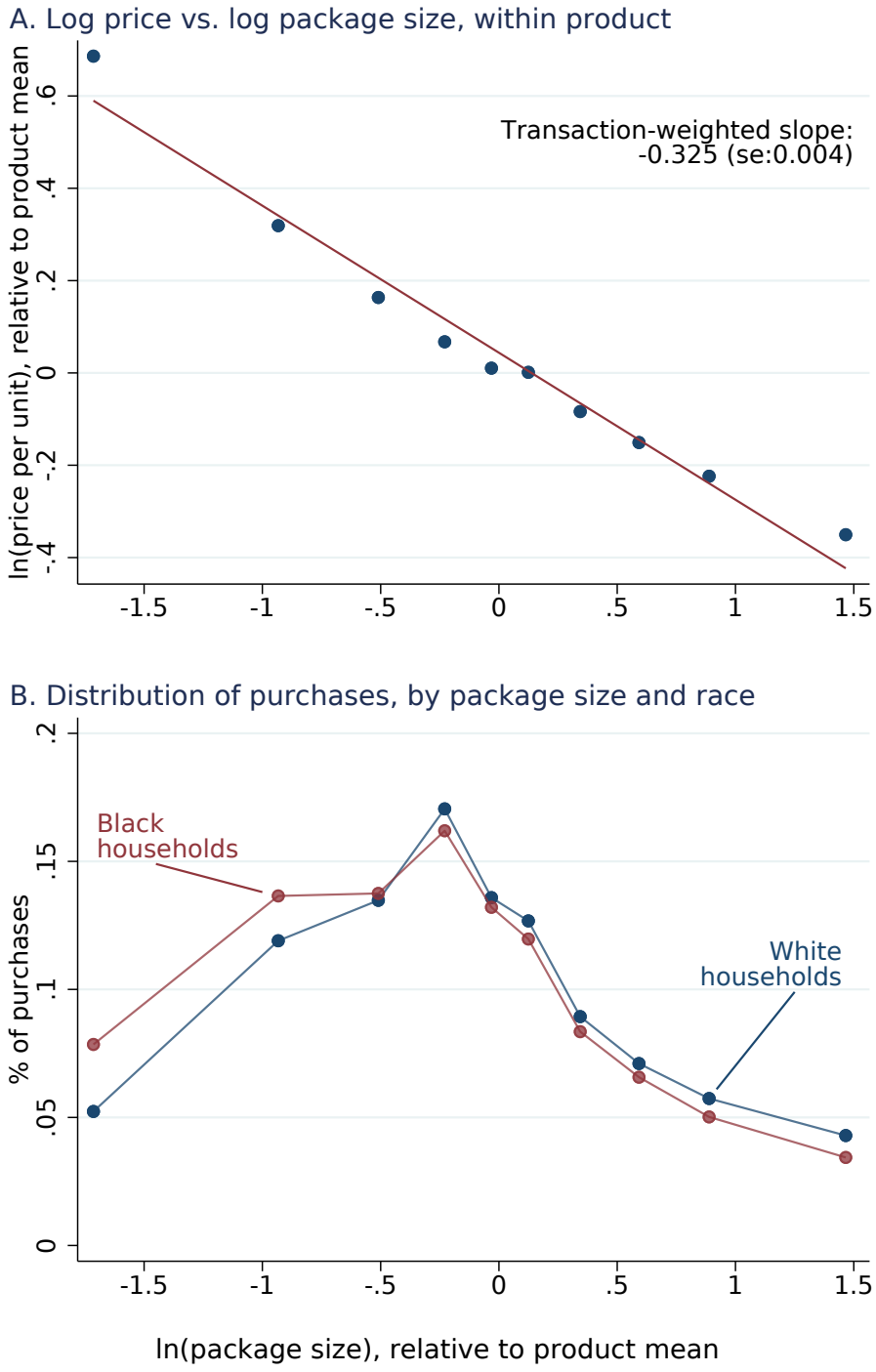
Notes: Table reports the share of each racial price gap explained by the indicated factor, using the decomposition of Gelbach (2016). Columns (1) and (2) decompose the price gap without using retailer identities; columns (3) and (4) include retailer-specific discounts. The shares add up to more than 1 because the racial price gap turns negative with the full set of controls. Robust standard errors, clustered on household, in parentheses.

Figure 1: Racial differences in prices paid, by income



Notes: Figure plots the average relative price index for the indicated income level and race. See Section 3 for details on the construction of the price index.

Figure 2: Price decreases with package size, but black households buy smaller packages



Notes: Panel A plots the average log price, residualized net of product-year fixed effects, in each bin of log package size, relative to product mean log package size. Panel B shows the distribution of purchases by package size, conditional on race, again relative to product mean log package size.

For online publication

A Appendix Exhibits

Table A.1: Summary statistics

Variable	Mean	Standard Deviation	Percentile 10	Percentile 50	Percentile 90
White	0.71				
Black	0.11				
Hispanic	0.12				
Other Race	0.06				
Less Than High School	0.03				
High School	0.27				
Some College	0.32				
College	0.25				
Post College	0.13				
Married	0.50				
Household size	2.55	1.46	1.00	2.00	5.00
# Adults	1.49	0.50	1.00	1.00	2.00
Age	48.72	12.09	30.00	50.00	65.00
Household income	52302	30650	15000	50000	100000
Expenditures on focal products	3515	2070	1365	3084	6215
# Observations	777,681				
# Households	175,440				

Notes: Table reports mean and, where appropriate, standard deviation and percentiles, for the indicated variables, at the household-year level.

Table A.2: Price differences by detailed race-ethnicity

Controls	None	Income	Demographics	Zip
	(1)	(2)	(3)	(4)
Black, non-Hispanic	0.020	0.020	0.018	0.011
	(0.001)	(0.001)	(0.001)	(0.001)
Asian, non-Hispanic	-0.002	-0.004	-0.004	-0.014
	(0.002)	(0.002)	(0.002)	(0.002)
Other, non-Hispanic	0.016	0.015	0.015	0.009
	(0.002)	(0.002)	(0.002)	(0.002)
White, Hispanic	0.003	0.003	0.006	0.003
	(0.001)	(0.001)	(0.001)	(0.001)
Black, Hispanic	0.029	0.029	0.033	0.019
	(0.004)	(0.004)	(0.004)	(0.004)
Asian, Hispanic	0.004	0.004	0.008	0.003
	(0.006)	(0.006)	(0.006)	(0.004)
Other, Hispanic	0.010	0.011	0.014	0.007
	(0.002)	(0.002)	(0.002)	(0.001)
# Households	175,440	175,440	175,440	171,122
Controls				
Income		Yes	Yes	Yes
Demographics			Yes	Yes
Zip code FE				Yes

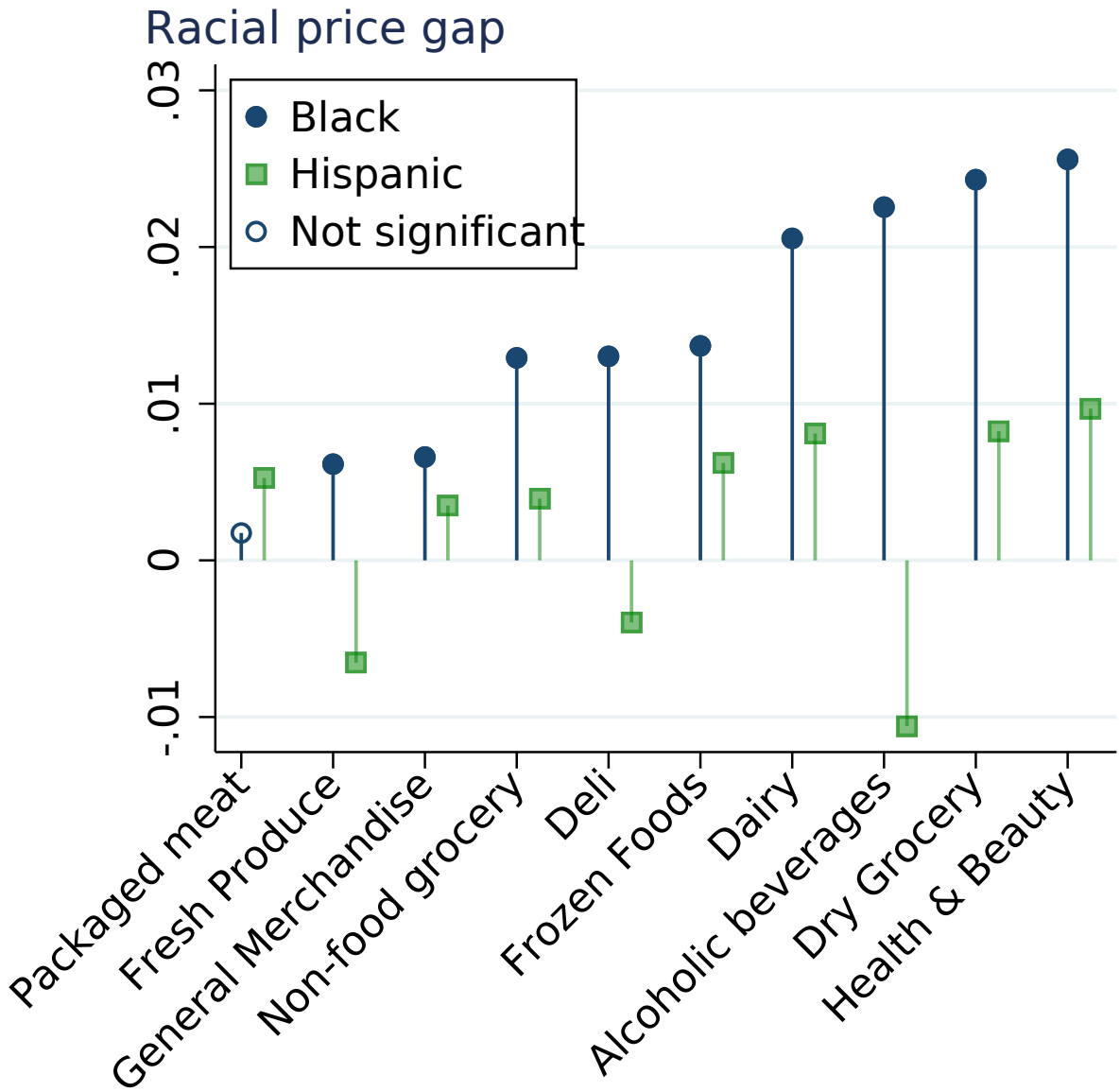
Notes: The dependent variable is the average log price paid at the household-year level, after residualizing out product-year fixed effects. Table reports the coefficients on indicators for each self-reported race/ethnicity category. Income controls are indicators for binned income amounts; demographics are are indicators for educational attainment, household size, married, age of head, and number of children; location controls are fixed effects as indicated . Robust standard errors, clustered on household, in parentheses.

Table A.3: Robustness to alternative approaches

Specification	Base	Exclude private label	Exclude warehouse purchase	Take out channel-product fixed effect	Take out package size-product fixed effect	Study pre-coupon price
	(1)	(2)	(3)	(4)	(5)	(6)
Black	0.020	0.020	0.019	0.016	0.013	0.016
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	0.008	0.008	0.010	0.007	0.007	0.005
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
# Households	175,440	175,440	175,434	175,440	175,440	175,440

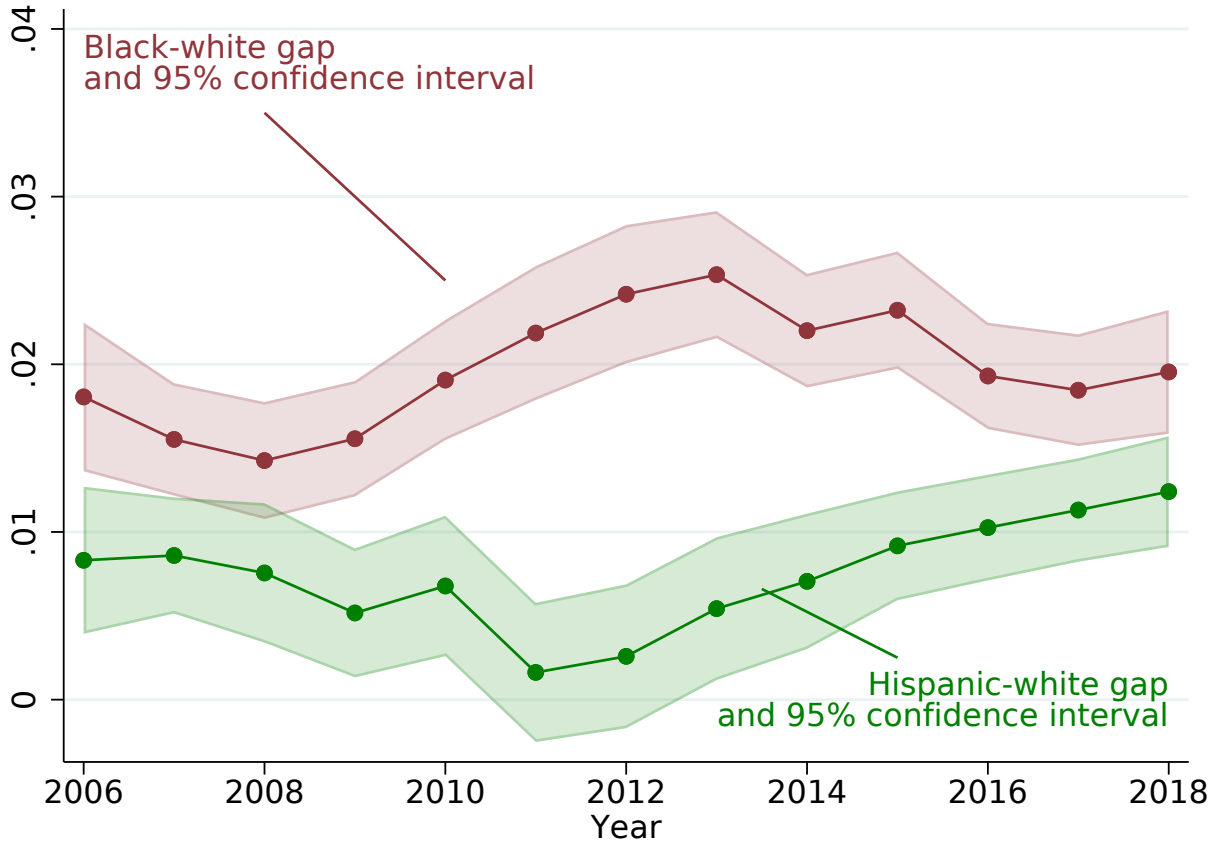
Notes: Columns (2)-(3) show that the racial price gpa is not sensitive to private label product inclusion and is not driven by warehouse clubs. Column (4) shows that the racial price gaps remain within retailer channel because we residualize prices net of Columns (5) and (6) show that package size and coupons remain important when we handle them by changing the dependent variable rather than including direct measures of size and coupon discounts. The dependent variable in column (1)-(3) is the average log price paid at the household-year level, after residualizing out product-year fixed effects. In column (4) we instead residualize out product-year-retail channel fixed effects, in column (5) we residualize out product-year-by-package size fixed effects, and in column (6) we work with the average log price gross of coupons. In column (2) we exclude private label products, and in column (3) we exclude purchases in warehouse clubs. Table reports the coefficients on indicators for non-Hispanic black and Hispanic (mutually exclusive; the omitted category is non-Hispanic white, and we also include an "all other race/ethnicity" category). Robust standard errors, clustered on household, in parentheses.

Figure A.1: Racial differences in prices paid, by department



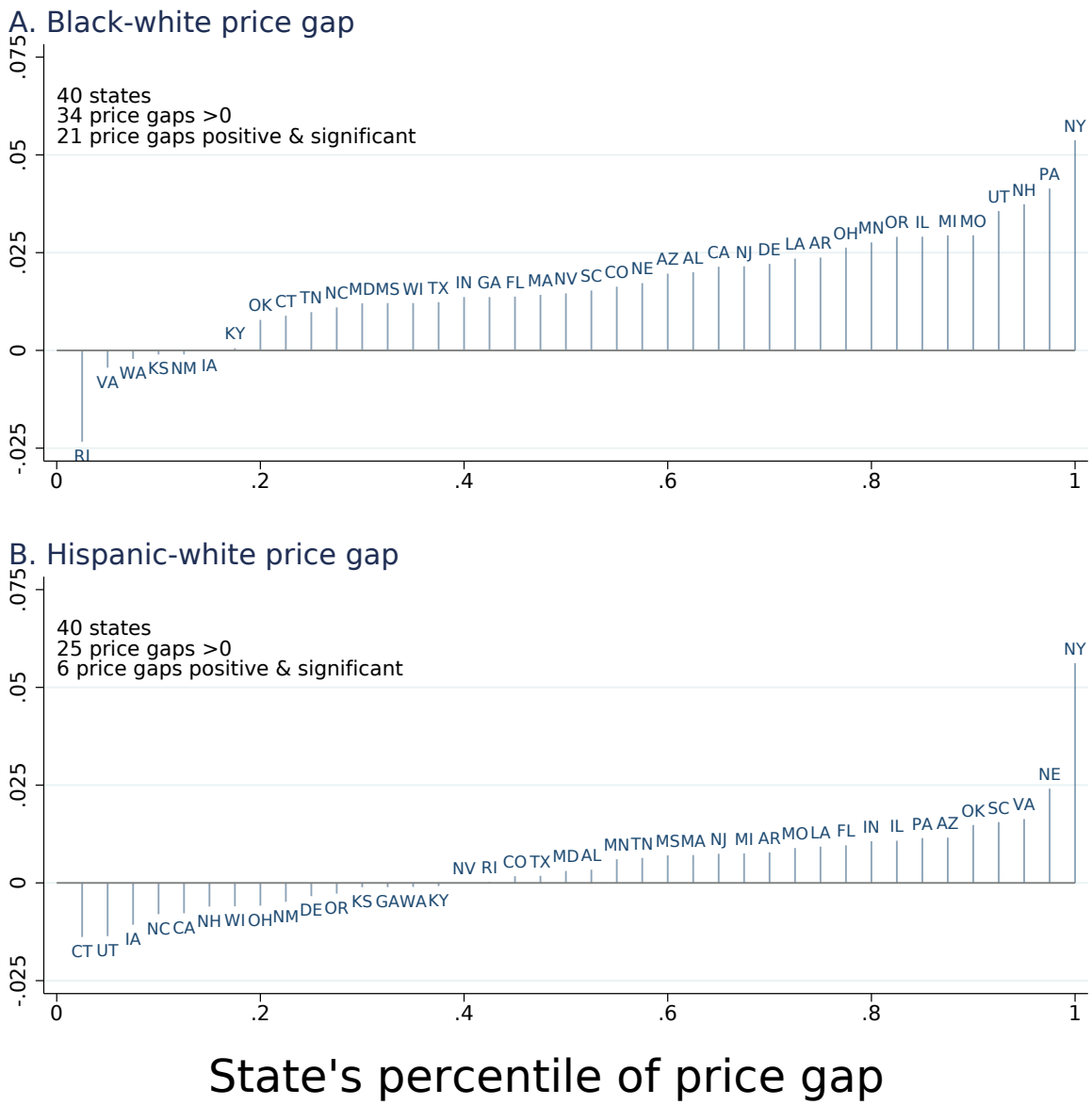
Notes: Figure plots the average racial price gap, by department (broad product category). The racial price gap is defined as the difference in average relative price index by racial category. See Section 3 for details on the construction of the price index. Hollow symbols indicate statistically insignificant differences.

Figure A.2: Racial differences by year



Notes: Figure plots the average racial price gap, by year. The racial price gap is defined as the difference in average relative price index by racial category. See Section 3 for details on the construction of the price index.

Figure A.3: Racial differences by state



Notes: Figure plots the average racial price gap, by state. The racial price gap is defined as the difference in average relative price index by racial category. See Section 3 for details on the construction of the price index. Hollow symbols indicate statistically insignificant differences.

B Zip code racial composition and characteristics

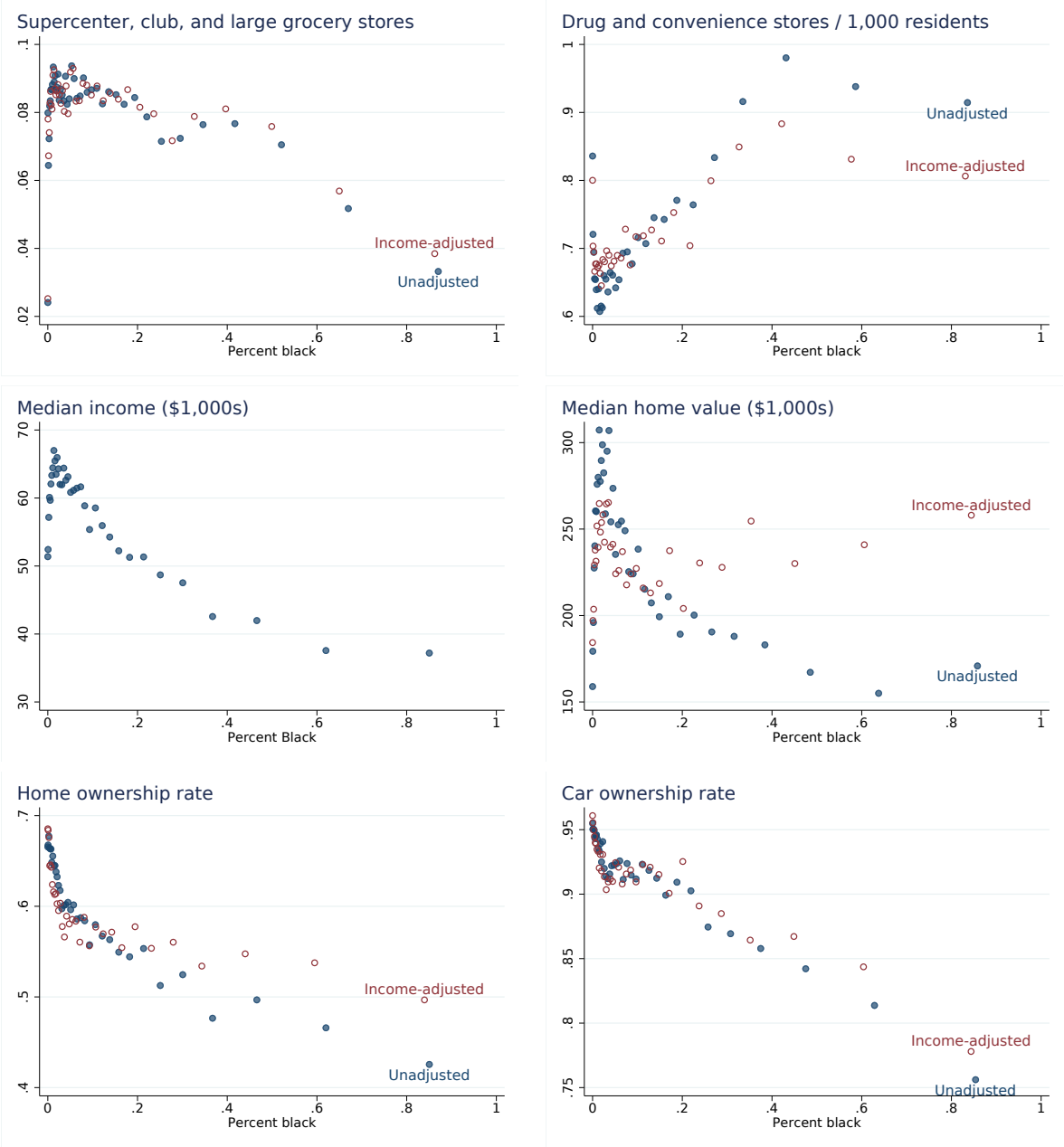
We study zip code racial composition and characteristics using data from the County Business Patterns database and the American Community Survey (ACS). From the County Business Patterns dataset, we construct measures of the number of large grocery, supercenter and club stores per capita as well as the number of gas and convenience stores per capita for each zip code area in 2016. These measures are defined using the criteria of Allcott et al. (2019). From the ACS, we collect the 2011 5-year estimates of demographic characteristics including: the median income, median home value, home ownership, as well as car ownership, defined as the fraction of occupied homes with at least one car.

Figure B.1 show the association between zip code average characteristics and the share of the population that is black. Each point in the scatter plot is an equally sized bin. We report unadjusted averages as well as averages adjusted for differences in zip code median income (except when we study median income itself). We select the number of points to plot, and we adjust for income, using the procedures and software of Cattaneo et al. (2019, 2021). We present analogous binned scatter plots by Hispanic share in Figure B.2.

We see, first, that large supermarkets are less common in zip codes with large black population shares, and drug and conveniences stores are more common. This association does not primarily reflect area income, despite the strong association between grocery store channel and income documented by Allcott et al. (2019); we see a negative association between large supermarkets and black share even after adjusting for income. Second, we also see area income, home value, home ownership, and car ownership. Zip codes with larger black population face several challenges in obtaining low price groceries: they have less access to low-price retailers, lower income, and less home and car ownership.

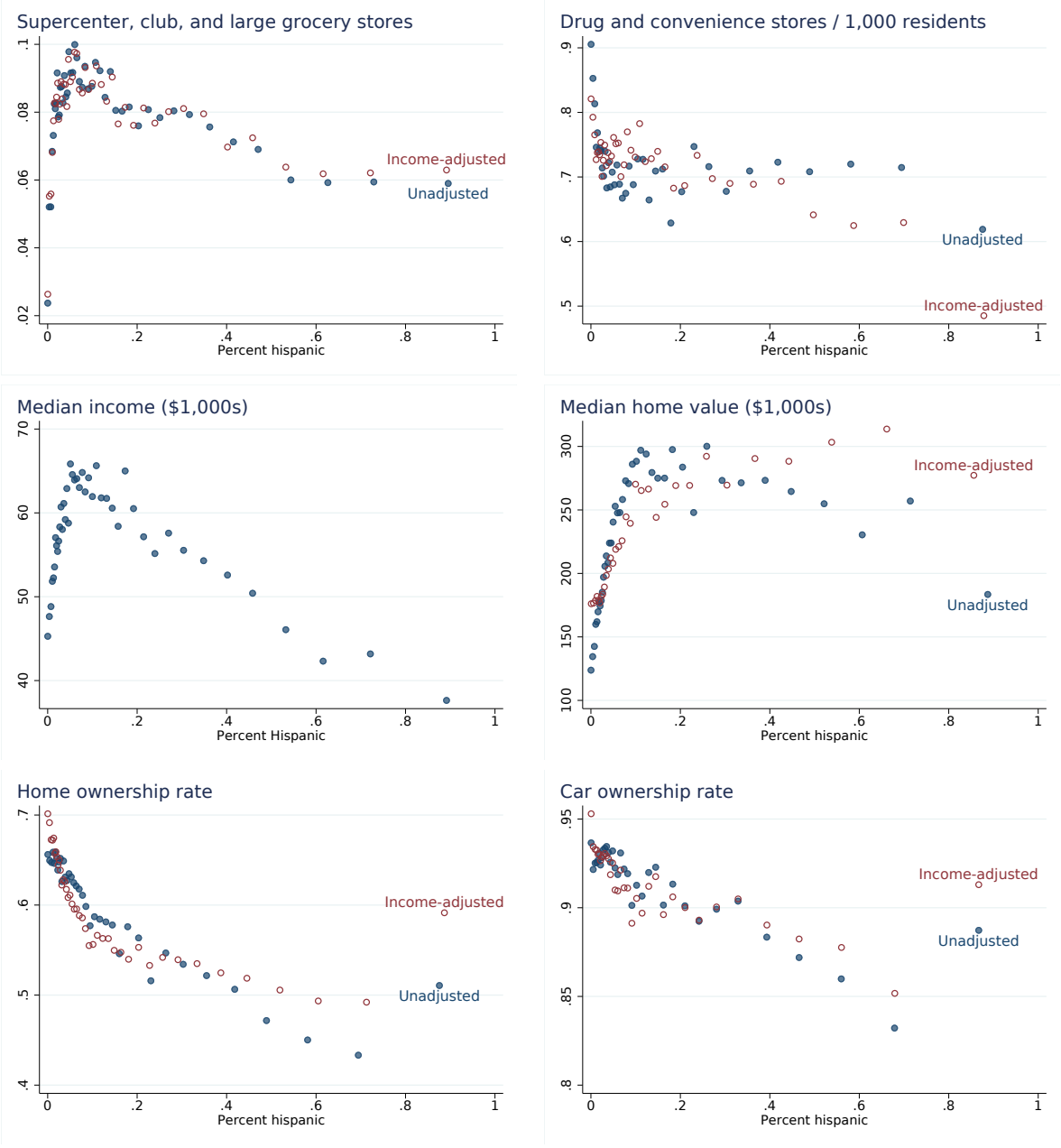
The patterns for Hispanic households are similar, although less extreme, and more sensitive to income adjustments. Neighborhoods with the highest Hispanic share have fewer large supermarkets, lower income, lower home value, lower home ownership, and less car ownership, than neighborhoods with low Hispanic populations.

Figure B.1: Zip code level characteristics by black share



Notes: Each panel displays a binscatter of a zip code area level statistic on the share of that zip code area's population that is black. In each case (except for the median income panel), both the raw binscatter and the binscatter after controlling for the role of zip code area median income are displayed. In the top two panels, the number of large grocery/supercenter stores per capita and drug/convenience stores per capita gathered from the zip code version of the County Business Patterns data for 2016 are displayed. In the middle two panels, the median zip code area income and median home value amongst owner occupied units for the zip code area gathered from the 2011 American Community Survey (ACS) are displayed. In the bottom two panels, the share of owner-occupied units and the share of households with access to at least one car gathered from the 2011 ACS are displayed.

Figure B.2: Zip code level characteristics by Hispanic share



Notes: Each panel displays a binscatter of a zip code area level statistic on the share of that zip code area's population that is hispanic. In each case (except for the median income panel), both the raw binscatter and the binscatter after controlling for the role of zip code area median income are displayed. In the top two panels, the number of large grocery/supercenter stores per capita and drug/convenience stores per capita gathered from the zip code version of the County Business Patterns data for 2016 are displayed. In the middle two panels, the median zip code area income and median home value amongst owner occupied units for the zip code area gathered from the 2011 American Community Survey (ACS) are displayed. In the bottom two panels, the share of owner-occupied units and the share of households with access to at least one car gathered from the 2011 ACS are displayed.