The Retail Price of Inequality

David M. Frankel

Tel Aviv University

Eric D. Gould

Hebrew University of Jerusalem

November 19, 1999

[1] We thank Avner Ahituv, Saul Lach, John Pencavel, Yoram Weiss, and seminar audiences at Boston University, Hebrew University, Stanford University, and Tel Aviv University for helpful comments. We also thank Dan Feenberg for crucial data and Yiftah Gordoni and Ran Greenwald for research assistance. Frankel acknowledges financial support from the Israel Foundations Trustees. Gould thanks the Falk Institute for financial support. Frankel: Department of Economics, Stanford University, Stanford, CA 94305 (email: dfrankel@leland.stanford.edu; tel: 650-723-9303). Gould: Department of Economics, Hebrew University, Mount Scopus, Jerusalem 91905, Israel (email: mseric@mscc.huji.ac.il; tel: 972-2-588-3247).

Abstract

This paper studies the relation between a city's income distribution and its retail price level using panel data. We find that an increase in the presence of lower middle income households, relative to poor or upper income households, is associated with lower prices. Our findings suggest that greater income inequality raises the prices that poor households face, thus making it harder for them to invest in human capital.

J.E.L. Number: H71.

Keywords: retail prices, income inequality, poverty, income distribution.

The Retail Price of Inequality

1 Introduction

The relation between retail prices and the local income distribution has long been controversial in the United States. Racial riots in the 1960s were sparked, in part, by complaints from the poor that they paid more for goods and services than their wealthier neighbors (see Alkaly and Klevorick [1]). Empirical studies have not supported this claim; poverty and prices are usually found to be more or less uncorrelated. We revisit this issue by considering the effects of the entire income distribution, not just the poverty rate as in prior studies. We also improve on previous research by using panel data with instrumental variables, which permits us to correct potential problems of endogeneity and unobserved heterogeneity. In addition, we confirm the robustness of our results using a second source of price data.

Our results indicate that higher prices are a result not of greater poverty *per se*, but rather of greater *inequality*. We divide the population into three groups: poor (below the poverty line), lower-middle income (between one and two times the poverty line), and middle-upper income (above twice the poverty line).[1] If one percent of the population moves from the lower-middle income group to the poor group, prices rise by about 0.8% under OLS (2.8% using instrumental variables). If the same one percent moves to the middle-upper income group, prices rise by about 0.7% under OLS (2.0% using instrumental variables).[2] A higher concentration of people in either the top or bottom income group leads to higher retail prices.

[1] We verify the robustness of these results using a different schema that is based on income quantiles. Results from the other schema also confirm the validity of treating the middle-upper category as a single group for purposes of explaining retail prices.

[2] These estimates are for fixed-effects regressions on a panel dataset; similar estimates are obtained in cross section.

These findings suggest that the recent trend towards greater income inequality may have hurt the poor by raising their cost of living. In 1980, the average poor central-city resident lived in a city in which 18.1% of the population were poor and 21.3% were between one and two times poverty.[3] By 1990, the first figure had risen to 20.3% while the second had fallen to 19.9%. In addition to lowering current consumption, higher prices leave the poor with fewer resources to invest in human capital, thus potentially exacerbating the "poverty trap."[4] This suggests a new way in which current inequality may lead to greater inequality in the future. The growth in the return to education is likely to have strengthened this effect in recent years.

We present results from both OLS and IV estimation. The OLS results answer the question of whether "the poor pay more": they do, but only to the extent that they are isolated from lower-middle income households. The IV results suggest that changes in the income distribution *cause* the changes in prices that we observe. This is important for two reasons. First, it rules out reverse-causality as an explanation: the findings are not due to the locational choices of the near-poor being more sensitive to cost-of-living differences. Second, some urban policies have the potential to affect the distribution of income (e.g., welfare policy, income taxation, choices of public good bundles). Our IV results suggest that such policies may have large effects on retail prices.

We test several explanations for our findings. One is based on search behavior. There is empirical evidence that search intensity is greatest among buyers with mod-

[3] 18.1% is the weighted average central city poverty rate in 1980, where the weights are the numbers of poor persons in each central city. Other percentages are defined analogously. A "central city" is defined by municipal boundaries; e.g., the central city of Boston does not include the adjacent municipalities of Cambridge or Brookline. There were 367 central cities in 1980 and 1990. Data are from Census Summary Tape File 3C [22].

[4] This assumes that the poor are credit constrained. For a theoretical treatment, see, e.g., Galor and Zeira [10].

erate incomes (see Carlson and Gieseke [6] and Goldman and Johansson [11]). This may be because search costs are U-shaped in income; in particular, the poor lack cars while the rich have a high cost of time.[5] Poor families also tend to have limited storage space and small budgets (see Kunreuther [14]). This leaves them captive to their local grocery store, while the non-poor are freer to search for bargains outside their immediate area.

Two proxies for the costs and benefits of search are used in the analysis. One reason poor families tend to have higher search costs is that they are less likely to own cars. Hence, to capture the role of search costs, we use the percent of households who do not own cars. To capture the *benefits* of search, we use average household size: search should pay off more for bigger households, which buy larger quantities. Hence, this variable should be negatively correlated with prices.

We also consider several competing hypotheses. If crime raises retail costs and poverty causes crime, the poverty rate could simply be picking up the effect of crime on prices. To test this, we include the property crime rate in the regression. Another hypothesis is that an increase in the middle-high income group leads real estate prices to increase. We test this using an index of housing costs. Finally, higher prices in poorer cities may be due to a prevalence of small convenience stores or "mom-and-pop" grocery stores, which have higher average costs. To test this, we include the percent of retail stores that have ten or more employees; this variable should be negatively related to the price level.

We find only weak evidence for any of these hypotheses: although the variables usually have the predicted signs, they are rarely significant in a statistical sense. Store size and the proxy for search costs are each significant (with the expected sign) in 1 out

[5]Goldman and Johansson [11, p. 178] also suggest that the poor may be less able to "identify, assess, and exploit marketplace opportunities." This can be thought of as a higher cognitive cost of search. For survey evidence indicating that the rich have a higher cost of time, see Marmorstein, Grewal, and Fishe [17].

of 4 specifications, while the other variables are never significant. Furthermore, even after controlling for these additional variables, the income distribution coefficients are still statistically significant, which suggests that either we are not testing these hypotheses with accurate measures or that some other explanation is driving the results. We leave it to future research to examine this further.

The paper is organized as follows. The next section discusses the relevant literature. Section 3 discusses the data and Section 4 presents the basic OLS results. IV estimation is presented in Sections 5 and 6. Explanations are tested in Section 7 and we conclude in Section 8.

2 Literature Review

The first to study the relation between income and retail prices was Caplovitz [5] in 1963. He found that the poor paid considerably more than the non-poor for major durables such as televisions and washing machines. However, Caplovitz's reliance on retrospective data was seen as problematic. Subsequent studies of food prices based on store surveys tended instead to find that prices are uncorrelated with or slightly increasing in neighborhood income [1, 8, 12, 15, 16]. Most recently, MacDonald and Nelson [15] find that a 1% increase in median income is associated with a (statistically insignificant) .013 to .019 percent food price increase.

All of the cited studies limit themselves to cross sectional OLS regressions; thus, their results may be contaminated by unobserved heterogeneity or endogeneity. In contrast, we confirm our results using both cross-sectional and panel data; thus reducing the likelihood that unobserved heterogeneity across cities is driving the results. Unlike prior studies, we also use instrumental variables to address problems of endogeneity. In addition, we base our price index on a broader consumption basket than either consumer durables (as in the case of Caplovitz) or food (as in the other studies). We also examine the effects of the income distribution on different components of the overall price index.

Our study also differs from previous work by looking at price differences across cities, while prior work examined differences across neighborhoods. We believe that this approach is better suited to measuring the causal effects of income distribution on prices for two reasons. The first reason is that our results are insensitive to the problem of cross-neighborhood shopping, which appears to be widespread (see, e.g., Mitchell [18]). Since a person's neighborhood of residence is not perfectly correlated with where she shops, studies of neighborhoods in a single city are less likely to detect a relationship between the income distribution and retail prices. The second advantage of conducting the analysis on the city level is that our results are less prone to the endogeneity of neighborhood choice within a city. Residents of a city sort themselves according to neighborhood characteristics (including perhaps the cost of living), and therefore it may be that prices determine the income distribution of a neighborhood and not the other way around. Sorting across cities is likely to be less of a problem since intercity mobility is more limited.[6] (This does not entirely satisfy our concerns about endogeneity, so we also use instrumental variables.)

3 The Data

Data were collected from 187 cities in 1979-80 and 1989-90.[7] Summary statistics are shown in Tables 1 and 2. For each city and biannual period, a retail price index was constructed using data from the American Chamber of Commerce Research

---

[6]Reasons include job market frictions, the costs of creating new social networks, and the relative difficulty of finding a house or apartment in a new city.

[7]Hence, for each city we have two biannual price indices. The 187 cities are those for which sufficient data existed to construct both indices.

Associates (ACCRA) [2].[8] To confirm our findings, we also examine a panel of 15 metropolitan areas using price data from the Bureau of Labor Statistics. This second dataset is explained in Section 4.

The ACCRA index is an expenditure-weighted average of price indices for four categories: groceries, transportation, health, and miscellaneous goods and services.[9] ACCRA also presents price indices for housing and utilities, which were excluded since we wish to focus on private sector retail markets.[10] ACCRA attempts to control for quality, either by careful specification or by restricting to given brand names. The following is a partial list of components of the ACCRA price index. The lists are complete in the case of transportation and health. For complete descriptions of the other two indices, see [2].

Groceries: One pound of 100% pork, Jimmy Dean brand sausage; a 6.125 oz. can of Starkist chunk light tuna; a half-gallon carton of whole milk; a 5 pound package

[8]ACCRA data are available only for selected cities and quarters. Each city's index in each quarter is reported by ACCRA as an index relative to the average of cities that reported that quarter. This means that the base of the index depends on the set of reporting cities, which changes from quarter to quarter. To eliminate this dependence, we recomputed each quarterly index so that prices are measured relative to average price level of the 49 cities that reported data in all 16 quarters in our 4 year sample (1979, 1980, 1989, and 1990). These recomputed indices were then used to compute biannual averages. One remaining problem is that the number of quarters for which prices are available varies by city and biannual period. To eliminate the resulting heteroskedasticity, we weighted by the number of quarters used to compute each biannual index. We also weighted by city size to make the results more representative of the urban population.

[9]The results we report use the expenditure weights that ACCRA reports for each year. These weights differ between 1979-80 and 1989-90 because of changes in expenditure patterns. We also computed price indices using, for each biannual period, the average of the weights in the years 1979, 1980, 1989, and 1990. The regression results (available on request) were virtually identical. We also present regressions for the individual indices, which are independent of the weighting scheme.

[10]The transportation and health indices also include some components that are not sold in private sector retail markets. We address this below by analyzing each of the four components separately.

of sugar; one dozen grade A eggs.

Transportation: One-way commuting fare to central business district, up to 10 miles; average price to spin-balance one front wheel; one gallon of regular unleaded gasoline, cash price at self-service pump, including all taxes.

Health: Average cost per day of semiprivate hospital room; fee for routine doctor's office visit; fee for routine adult teeth cleaning and adult oral examination; price of 100-tablet bottle of Bayer 325-mg. aspirin.

Miscellaneous: MacDonald's Quarter-Pounder with cheese; 12" Pizza Hut thin crust cheese pizza; dry cleaning of man's two piece suit; ticket to first-run movie at indoor cinema, evening, no discount; Miller Lite or Budweiser beer, 6-pack, 12-oz. containers, excluding deposit.

Prices do not include sales taxes except for the gasoline component of the transportation price index.

Data on the income distribution come from the decennial U.S. Census.[11] We used two income schemas. The first is based on the ratio of a person's pretax family income to the poverty line for a family of the given size and composition.[12] We use the percents of persons in three groups: below the poverty line, between one and two times the poverty line, and over twice the poverty line. (These were the natural delineations because no breakdown of the third group was available.) The second classification is based on after-tax family income. It consists of the share of families

[11]We used Census places as our geographic unit, since they correspond to the places used by ACCRA. A Census place is defined by metropolitan political boundaries, like a central city. In fact, a central city is just a large Census place that is located in a Census metropolitan area.

[12]The data come from Summary Tape File 3C, Census of Population and Housing [22]. Pretax income includes market income as well as cash transfer payments (AFDC, unemployment compensation, Social Security, etc.). The poverty line is adjusted for inflation over time but does not vary by region.

that fall into each national after-tax income quintile.[13] (For example, if 25% of a city's population falls into the poorest national quintile, then the city is poorer than the nation as a whole.) This schema has the advantage of taking federal and state income taxes into account. Moreover, a finer breakdown of the middle-upper segment is available, so we can test the validity of treating it as a single group. The drawback of this schema is that no family size adjustment is made, unlike the poverty classification.

Prices can also depend on sales taxes. We control for this using the combined city, state, and county sales tax rate from the National Sales Tax Directory [26]. A small number of our cities (10 in 1982 and 11 in 1992) also have city or county income taxes. We tried entering the average combined city and county income tax rate.[14] Its coefficient was always insignificant, so we removed it from the specifications.

All regressions using this dataset also control for a fixed set of other demographic variables.[15] These are the percent of persons aged 65 and over, percent Hispanic, percent female, percent of households that are not family households, and the log of population density. We began with a larger set of demographic controls but removed

[13]The Census Bureau gives distribution information for 17 and 25 family income categories in 1980 and 1990, respectively. For each city-year, we converted this to an after tax income distribution by reducing the category boundaries by the combined state and federal tax rate for each bracket for each year. We then aggregated the categories, interpolating where necessary, to estimate the proportion of families in each of the national after tax income quintiles. Tax rates were computed using the internet TAXSIM program of the National Bureau of Economic Research [9]; we thank Dan Feenberg for working to make this program available and helping us use it. In computing tax rates, we assumed a married couple filing jointly with two children and no other deductions.

[14]We used the rates in 1982 and 1992 for observations in 1979-80 and 1989-90, respectively. Rates were computed as total tax receipts (from [21]), divided by aggregate personal income in the city (from [22]).

[15]We also perform regressions on a smaller dataset of 15 metropolitan areas; these regressions do not include any control variables.

a subset whose effects were always insignificant.[16] All regressions are also weighted by city size to make the results more representative of the urban population. A city's weight is the average of its 1980 and 1990 populations.

4 OLS Results

We first estimate the relationship between the income distribution and the price level using simple OLS regressions. In order to reduce problems of unobserved heterogeneity across cities, we exploit the panel structure of the dataset by regressing the 1979/80 to 1989/90 differences of the dependent variable on the differences in the independent variables. (Later in this section we present pooled, cross-sectional results, which are broadly similar.)

Results for the overall retail index are presented in Table 3.[17] In the first column, the only income measure included is the percent of households below the poverty line.[18] The coefficient estimate is not statistically significant. This is consistent with the existing literature, reviewed above, which has found a small or insignificant relation between poverty and prices.

In order to capture more features of the income distribution, column two of Table 3 includes the percent of households over two times the poverty line. The suppressed variable is the size of the intermediate group: the percentage of households between the poverty line and twice the poverty line. Both income coefficients are now positive

[16]These include the percent aged 17 and under, percent black, percent of households that are married couple households, percent unemployed, male and female labor force participation rates, and the percent of persons living in urban areas.

[17]Each specification in Table 3 controls for the sales tax and the demographic covariates described in section 3. For parsimony, we present only the coefficient estimates for the income variables.

[18]All income variables are numbers between 0 and 1, measuring the proportion of persons or families in a given group.

and statistically significant. Relative to the intermediate group, a one percentage point increase in the poor group or the higher income group is associated with an increase of 0.76% and 0.69%, respectively, in retail prices. The R-squared increases from 0.33 to 0.42 by adding the second income variable.

Column three in Table 3 uses the income quintiles. The suppressed group is the set of households falling in the second national after-tax quintile. (The "first quintile" refers to the poorest group, with the others following in order.) This is the group that most closely matches the suppressed group using the poverty schema. The coefficient estimates for the other four quintiles in the third column are all positive, which is consistent with the effects found in column two using the poverty classification. The effects are not significant for the third and fourth quintile groups, and are barely significant for the lowest quintile group.

An F-test for the equality of the coefficients of the top three income groups yields a p-value of 0.26. Therefore, we cannot reject that these categories have the same effects on prices. In the last column of Table 3, the top three income quintiles are aggregated together. With this specification, the coefficient estimates for the poorest group and the upper group are now significant. Relative to the second quintile, a one percentage point increase in the poorest quintile (respectively, the upper three quintiles) is associated with a price increase of 0.52% (0.57%). The aggregation of the top three quintiles does not reduce the adjusted R-squared, which supports this more parsimonious specification.[19] For this reason, all subsequent regressions using this income classification will combine the upper three quintiles.

Up to now, our dependent variable has been the (logged) overall retail price index. This index is an expenditure-weighted average of four individual price indices: groceries, miscellaneous, health, and transportation. We now examine each individual

[19] Combining the quintiles also reduces problems of multicollinearity. The standard errors for quintiles 3, 4, and 5 in column 3 are .35, .23, and .18, respectively, while the standard error for the combined category in column 4 is just .14.

index in turn.

Table 4 presents OLS estimates for each price component using the poverty classification; Table 5 uses the quantile categories. The results are similar to what was found for the overall index: an increase in either the low or middle-high income group is usually associated with higher prices. However, the transportation index does not follow the pattern of results for the other indices. Most notably, the estimates for the quantile categories are statistically insignificant and negative. For the poverty categories, the estimates are positive but still far from significant. Hence, transportation goods appear to be different. This difference could be explained by two factors. One is that public transportation is one of three components of the transportation index, and these prices are set by local governments rather than private bodies. Another reason is that gasoline (the second of three components of the index) is subject to high state excise taxes. In fact, gasoline is the only component of the overall retail price index that includes taxes. For these reasons, we might not expect transportation to follow a similar pattern as the other indices.

The preceding results come from ten-year difference regressions. Table 6 presents cross sectional regressions. The two periods (1979-80 and 1989-90) are pooled, with the addition of a year dummy to capture changes in the average price. Results for both income schemas are consistent with those found in the difference regressions in the preceding tables. This strengthens the evidence for a positive relation between inequality and prices.

Alternative Dataset: 15 Metropolitan Areas

To test whether our results depend on the source of the price data, we also studied the 15 metropolitan areas for which the U.S. Bureau of Labor Statistics collected price data during the 1980s. The price index consists of all items except shelter. The 15 metropolitan areas are listed in Table 7. (Metropolitan areas tend to be considerably larger than cities, the geographic unit studied in our main dataset.)

The results for this dataset are shown in Table 8. The dependent variable is the log of retail prices.[20] The first two columns are differences-on-differences regressions, like Tables 3, 5, and 4. These coefficient estimates are broadly similar to those for our main dataset and mostly significant. The last two columns are cross sectional regressions, as in Table 6. These estimates have the same signs, but are smaller and less significant than those for our main dataset. This may be due to the absence of control variables, which are precluded by the small sample size.

The data used in the second column of Table 8 are plotted in Figure 1. (Names of metropolitan areas are abbreviated.) Note the pronounced negative relation between the change in the price level and the change in the proportion of lower-middle income persons.

Discussion of OLS Results

Despite differences in data and techniques, our first regression replicates the findings of the existing literature: by itself, the poverty rate is essentially uncorrelated with retail prices (see column 1 of Table 3). However, the results change sharply when a more complete income schema is used. Growth in the poor or middle-upper income group relative to the lower-middle income group is associated with higher prices. This illustrates the shortcomings of using a dichotomous poor/nonpoor income schema. These results appear in both cross-section and panel regressions on our main dataset and can also be detected in panel regressions using a different source of price data.

5 IV Estimation

The OLS results show that prices are not higher in high poverty cities per se, but rather in cities with high inequality: those with relatively few lower-middle income

[20] To maintain consistency with our primary dataset, the 1980 (1990) price index computed here is actually the average of the years 1979-80 (1989-90). Using only 1980 and 1990 data gives essentially the same results.

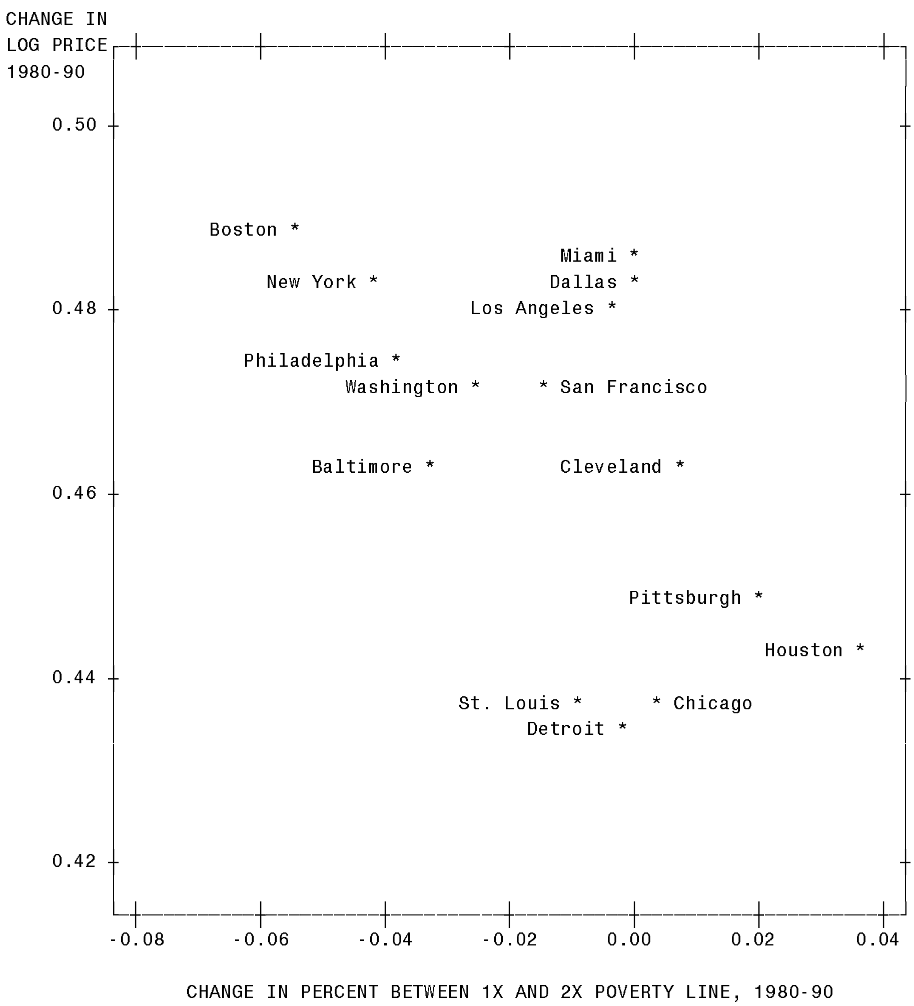

Figure 1: Using BLS Price Data from 15 Metropolitan Areas

households. One potential explanation for this is reverse causality: in deciding where to live, lower middle income households may be especially sensitive to high retail prices as they have tighter budgets than the rich while still having sufficient resources to move (unlike many poor households). If so, they may be more likely than the rich or the poor to choose to live in low-cost cities. In this section we test this explanation via instrumental variables.

Our strategy consists of using information about a city's weather in 1980 to predict the change in income distribution from 1980 to 1990.[21] The poor are likely to be more sensitive to extreme weather conditions since cooling and heating costs are a larger portion of their budgets. This means that a hot summer or a cold, snowy winter in 1980 should lead poor residents to leave disproportionately. Because of moving costs, such migration is likely to have an enduring effect on the income distribution, and thus should affect the change in income distribution from 1980 to 1990. Indeed, controlling for changes in demographics and weather, the partial correlation between initial winter precipitation and the change in the poverty rate equals -0.30 ($p = .0002$). The partial correlation between initial winter precipitation and the change in the proportion above twice the poverty rate is 0.27 ($p = .0006$). Thus, more snow in 1980 lowered the poverty rate over the decade and raised the proportion of those with middle to high incomes.[22]

Extreme weather in 1980 might also affect prices in 1980. This is a potential problem because it means that even if 1980 weather has no effect on 1990 prices, it may

[21] Weather data is from the National Oceanic and Atmospheric Administration [19].

[22] The other variable that is significantly associated with the income distribution changes is the square of the minimum temperature in the initial year. This variable's partial correlation with the change in the poverty rate was .20 ($p = .011$). Its partial correlation with the change in the proportion over twice poverty was -.22 ($p = .007$). Thus, cold winters also lower the poverty rate relative to the proportion of people with middle-high incomes. (The corresponding correlations based on the *unsquared* minimum temperature in the initial year were of the same signs as those reported above, but were not statistically significant.)

still be correlated with the error term in the structural equation as its correlation with 1980 prices may imply a correlation with the 1980-90 change in prices (e.g., if prices are mean-reverting). We address this problem by including, in the structural equation, the *change* in the weather instruments and in other weather variables between 1980 and 1990. The effect of 1980 weather on the change in prices from 1980 to 1990 should disappear once we control for the change in weather from 1980 to 1990. To the extent that this is not so (e.g., because of nonlinear effects of weather on prices), the weather instruments should fail tests of overidentifying restrictions. We use the standard Basmann test to check this. Although the Basmann test is known to reject too often, the instruments pass this test in 11 out of 12 specifications. Thus, there appears to be little cause for concern that the instruments may be correlated with the error term.

The set of instruments consists of the following five variables from 1979 and 1980: winter precipitation in inches, the annual maximum and minimum of the 12 monthly mean temperatures, and the squares of these maximum and minimum temperatures. In the second stage we control for the 10-year changes in these variables as well as in several other weather variables that affect prices: average temperatures and precipitation for each season.[23]

6 IV Results

Table 9 presents two-stage-least-squares results using the poverty classification. For each price index, the coefficient estimates for the poorest group and the middle-upper income group are positive and significant. The IV estimates for both coefficients are larger in magnitude than the OLS estimates (Table 4). The IV coefficients on the

[23]These additional variables are the change in summer precipitation and summer mean temperature from 1979 to 1989 and from 1980 to 1990, as well as the changes in spring, fall, and winter precipitation and mean temperatures between both pairs of years.

percent poor range from 2.37 to 5.56, compared to OLS estimates of 0.05 to 1.50. For the percent over twice the poverty rate, IV estimates range from 1.63 to 3.29, compared to 0.36 to 1.26 for OLS. This suggests that OLS induces a *downward* bias in the coefficients. Thus, reverse causality does not appear to underlie our findings.

Table 10 presents the same two-stage least squares specifications but instead using the quantile classification. Again, we see the downward bias of OLS. For the lowest quintile, the IV coefficient estimates range from 0.58 to 4.84 in Table 10, while OLS estimates range from -0.64 to 1.18 in Table 5. For the top 60% of the income distribution, IV estimates range from 0.13 to 2.50 while OLS estimates are from -0.18 to 1.04.

7 Explanations

One explanation for our findings is that lower-middle income households have lower search costs than either of the other two groups. This is supported by the finding of Carlson and Gieseke [6] and Goldman and Johansson [11] that search intensity is greatest among buyers with moderate incomes. In this section we attempt to test the search hypothesis. Our test is necessarily indirect since we do not have observations of the search costs of different households. Instead, we examine the effects of adding proxies for the costs and benefits of search in the regression. To proxy for the benefits of search, we use the log of average household size: search should pay off more for bigger households, which buy larger quantities. Hence, this variable should be negatively correlated with prices. Our proxy for the cost of search is the percent of households that do not own a car. Since not having a car raises search costs, an increase in the percentage of households without cars should discourage search and lead to higher prices.

We also consider several competing hypotheses. If higher crime rates lead to higher retail costs and if crime is concentrated in poorer areas, our measure for the poor could be picking up the effect of higher crime rates in poorer areas. Consequently, we include the log of reported property crimes per capita in the regression.[24]

Another hypothesis could be that an increase in the middle-high income group leads real estate prices to increase. This drives up retailers' operating costs, forcing them to raise prices. We test this using ACCRA's housing cost index, which is an average of the cost of renting an unfurnished two bedroom apartment and the monthly mortgage payment on a new 1,800 square foot home.

Finally, prices may move in response to unmeasured changes in the quality of service. Smaller stores provide a service to poor residents by locating closer to their homes and selling smaller quantities. If smaller stores have higher unit costs than larger stores, it could be that the poor pay for this service through higher prices.[25] To test this, we include the percent of retail stores that have ten or more employees using data from the Zip Code Business Patterns [23]. If prices in smaller stores are higher because their costs are higher, this variable should be negatively related to the price level.[26]

The OLS and IV estimates including these alternative variables are presented in Table 11. The first two columns show results for the poverty categories while the last two columns use the quantiles. The dependent variable is the overall retail index (in

[24]This data are from a CD-ROM of the U.S. Census Bureau [24]; the original source is the Federal Bureau of Investigation. We did not have city level data, so we used the county crime rate. In a few cases a city overlapped several counties; here, we used the overall crime rate for the overlapping counties taken together. Property crimes comprise burglaries, robberies, and larcenies. We also tried replacing the overall property crime rate with each of these components in turn; the results were essentially the same.

[25]Kunreuther [14] presents evidence that unit prices are indeed higher in small stores.

[26]Another service-related hypothesis could be that retail firms in richer areas invest more in the upkeep of their stores or hire better workers. We have no way to test this theory. However, this would offer at most a partial explanation for our findings, since it does not explain why growth in the poor group pushes prices up.

log form). The specifications include all of the variables used previously, in addition to the new hypothesis variables. As predicted by the theory, household size is negatively related to the price level. However, this variable is statistically insignificant in all specifications. We find inconsistent results for the percent of households without cars: its effect is sometimes positive and sometimes negative, but only once significant.

Table 11 also shows that property crime has a negligible effect on prices in both an economic and a statistical sense. The proxy for retail service – the percent of stores with ten or more employees – is negatively (but only once significantly) related to prices, which is weakly consistent with our hypothesis that small stores provide a valuable service that people are willing to pay for. Finally, the median rental price, which proxies for the cost of real estate, has effects that are both economically and statistically insignificant.

Even with all the hypothesis variables in the equation, the estimated coefficients of the income groups are still large and significant. This may indicate that other, undiscovered explanations play a role. However, it may simply be that our proxies are simply too indirect for our purposes. Further work is needed to give a more definitive explanation for our findings.

8 Discussion

This paper presents evidence for a previously unstudied effect of greater income inequality: it leads to higher retail prices. Prices increase when lower-middle income households in a community are replaced by either poor or middle-higher income residents. One implication is that greater isolation of the poor from lower-middle income households will raise the prices paid by poor residents. In addition to lowering the current consumption of the poor, this is likely also to make it more difficult for them to invest in human capital. Hence, increased inequality may intensify the "poverty trap" through its effect on retail prices.

Our results are stronger and more reliable than those of prior studies for several reasons. Using different data and methodology, we reproduce the insignificant effects found in prior studies, but show that this is a result of using a dichotomous poor/nonpoor income schema. Significant and economically important effects emerge when we turn to a finer schema that captures the income distribution in greater detail. We also address problems of simultaneity and unobserved heterogeneity by using city level panel data with instrumental variables. Our findings are shown to be robust to OLS (cross-section and panel analysis), instrumental variables, the inclusion of several alternative hypotheses variables, and an alternative source of price data.

We also considered several competing hypothesis that can explain these results. However, empirical attempts to test these hypotheses were inconclusive, if not negative. Further work is needed to explain the mechanism behind our findings. But one clear conclusion from our study is that changes in the income distribution can have a large causal effect on retail prices, which suggests that cities should consider these effects in their assessment of policies that may affect the local distribution of income.

References

[1] Alcaly, Roger E., and Alvin K. Klevorick. 1971. "Food Prices in Relation to Income Levels in New York City." *Journal of Business* 44:380-397.

[2] American Chamber of Commerce Researchers Association. Various years. *Inter-City Cost of Living Index.*

[3] Altonji, Joseph G., and David Card. 1991. "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives." In *Immigration, Trade, and the Labor Market*, edited by John M. Abowd and Richard Freeman. Chicago: University and of Chicago Press and National Bureau of Economic Research.

[4] Bound, John, David A. Jaeger, and Regina M. Baker. 1995. "Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable Is Weak." *Journal of the American Statistical Association* 90:443-50.

[5] Caplovitz, David. 1963. *The Poor Pay More*. New York: The Free Press of Glenco.

[6] Carlson, John, and Robert Gieseke. 1983. "Price Search in a Product Market." *Journal of Consumer Research* 9:357-365.

[7] Dixon, Donald F., and Daniel J. McLaughlin, Jr. 1970. "Low Income Consumers and the Issue of Exploitation: a Study of Chain Supermarkets." *Social Science Quarterly* 51:320-328.

[8] Donaldson, Loraine, and Raymond S. Strangways. 1973. "Can Ghetto Groceries Price Competitively and Make a Profit?" *Journal of Business* 36:61-65.

[9] Feenberg, Daniel, and Elizabeth Coutts. Winter 1993. "An Introduction to the TAXSIM Model." *Journal of Policy Analysis and Management* 12:189-194.

[10] Galor, Oded, and Joseph Zeira. 1993. "Income Distribution and Macroeconomics." *Review of Economic Studies* 60:35-52.

[11] Goldman, Arieh, and J.K. Johansson. 1978. "Determinants of Search for Lower Prices: An Empirical Assessment of the Economics of Information Theory." *Journal of Consumer Research* 5:176-186.

[12] Goodman, Charles S. 1968. "Do the Poor Pay More?" *Journal of Marketing* 32:18-24.

[13] Kumar, V., and Robert Leone. 1988. "Measuring the Effect of Retail Store Promotions on Brand and Store Substitution." *Journal of Marketing Research* 25:178-85.

[14] Kunreuther, Howard. 1973. "Why the Poor Pay More for Food: Theoretical and Empirical Evidence." *Journal of Business* 46: 368-383.

[15] MacDonald, James M., and Paul E. Nelson, Jr. 1991. "Do the Poor Still Pay More? Food Price Variations in Large Metropolitan Areas." *Journal of Urban Economics* 30:344-359.

[16] Marcus, Burton H. 1969. "Similarity of Ghetto and Nonghetto Food Costs." *Journal of Marketing Research* 6:365-368.

[17] Marmorstein, Howard, Dhruv Grewal, and Raymond P.H. Fishe. 1992. "The Value of Time Spent in Price-Comparison Shopping: Survey and Empirical Evidence." *Journal of Consumer Research* 19:52-61.

[18] Mitchell, Alison. June 6, 1992. "Where Markets are Never Super; Some Urban Neighborhoods Fight Absence of Major Chains." *The New York Times*, late edition, section 1, p. 25.

[19] National Oceanic and Atmospheric Administration. 1998. *United States Historical Climatology Network Serial Temperature and Precipitation Data.* Asheville, North Carolina.

[20] Staiger, Douglas, and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65: 557-86.

[21] U.S. Department of Commerce, Bureau of the Census. 1982, 1992. *Census of Governments.*

[22] U.S. Department of Commerce, Bureau of the Census. 1980, 1990. *Census of Population and Housing, Summary Tape File 3C.*

[23] U.S. Department of Commerce, Bureau of the Census. Various years. *Zip Code Business Patterns.*

[24] U.S. Department of Commerce, Bureau of the Census. 1996. *USA Counties 1996*. Washington, D.C. (CD ROM)

[25] U.S. Department of Labor, Bureau of Labor Statistics. Various years. *Consumer Price Index*. Washington, D.C.

[26] Vertex Systems Inc. *National Sales Tax Rate Directory*, various years. Berwyn, Pennsylvania.

[27] Walters, Rockney. 1991. "Assessing the Impact of Retail Price Promotions on Product Substitution, Complementary Purchase, and Interstore Sales Displacement." *Journal of Marketing* 55:17-28.

Variable	Mean	Std. Dev	Minimum	Maximum
Retail Price Index	100.0	4.58	90.1	116.7
Groceries Prices	99.8	4.15	90.5	112.9
Health Prices	99.1	13.26	73.3	157.8
Transportation Prices	99.4	6.41	85.9	126.1
Miscellaneous Prices	100.6	4.76	89.6	118.4
Population	183539	325698	10034	3485398
Percent Under Poverty	0.1867	0.0702	0.0334	0.5800
Pct Btwn 1X and 2X Poverty	0.2011	0.0351	0.0620	0.3061
Pct Over 2X Poverty	0.6122	0.0918	0.2027	0.9046
Pct in 1st Income Quintile	0.2512	0.0774	0.0417	0.6812
Pct in 2nd Income Quintile	0.2195	0.0345	0.0834	0.3219
Pct in 3rd Income Quintile	0.2012	0.0309	0.0829	0.2799
Pct in 4th Income Quintile	0.1747	0.0336	0.0448	0.2770
Pct in 5th Income Quintile	0.1534	0.0709	0.0163	0.5246
Population Density	1046.5	607.5	202.7	3891.4
Pct Stores with >9 Employees	0.3404	0.0389	0.2037	0.4689
Total Sales Tax Rate	0.0603	0.0136	0	0.0900
Property Crime Rate	0.0158	0.00832	0.000484	0.0448
Pct Households with No Cars	0.1348	0.0717	0.0206	0.4527
Mean Household Size	2.5803	0.2243	2.0078	3.6619
Pct Aged 65+	0.1329	0.0386	0.0357	0.2810
Pct Hispanic	0.0818	0.1458	0.00201	0.7685
Pct Female	0.5242	0.0148	0.4807	0.5538
Pct Non-Family Households	0.3678	0.0685	0.1897	0.5552
Annual Maximum Temp., Yr. t	89.06	6.52	76.45	108.40
Annual Minimum Temp., Yr. t	17.60	11.66	-9.39	49.46
Winter Precipitation, Yr. t	2.6758	2.0713	0.00500	9.3356

Table 1: Summary statistics for 187 cities in sample, 1990 values. Population density equals number of persons per square mile. Total sales tax is sum of state, city, and county rates. Property crime rate equals property crimes per capita. Temperatures are in farenheit; precipitation is in inches. The following are shown in absolute form, but logs are used in regressions: all price indices, population density, the property crime rate, and mean household size.

Variable	Mean	Std. Dev	Minimum	Maximum
Δ Retail Price Index	0.2246	3.9768	-8.8408	13.2116
Δ Groceries Prices	-0.5910	4.3201	-12.4158	9.7813
Δ Health Prices	-0.7127	10.0218	-44.0192	26.9837
Δ Transportation Prices	1.2151	7.0627	-15.7497	26.6486
Δ Miscellaneous Prices	0.4298	7.4634	-30.2416	18.1142
Δ Population	14326.7	51587.7	-60577.0	518548
Δ Percent Under Poverty	0.0331	0.0318	-0.0419	0.1932
Δ Pct Btwn 1X and 2X Poverty	-0.00852	0.0260	-0.1066	0.1112
Δ Pct Over 2X Poverty	-0.0246	0.0450	-0.1653	0.1079
Δ Pct in 1st Income Quintile	0.0322	0.0368	-0.0847	0.2065
Δ Pct in 2nd Income Quintile	0.000335	0.0258	-0.0681	0.0913
Δ Pct in 3rd Income Quintile	0.00573	0.0184	-0.0518	0.0589
Δ Pct in 4th Income Quintile	-0.0116	0.0205	-0.0678	0.0492
Δ Pct in 5th Income Quintile	-0.0266	0.0360	-0.2310	0.0591
Δ Population Density	43.5611	134.9	-351.6	529.6
Δ Pct Stores with >9 Employees	0.0377	0.0327	-0.1513	0.1419
Δ Total Sales Tax Rate	0.0142	0.00963	-0.00500	0.0350
Δ Property Crime Rate	-0.00079	0.00532	-0.0183	0.0124
Δ Pct Households with No Cars	-0.00600	0.0190	-0.0683	0.1116
Δ Mean Household Size	-0.1223	0.1123	-0.6044	0.2487
Δ Pct Aged 65+	0.0135	0.0151	-0.0247	0.1205
Δ Pct Hispanic	0.0131	0.0235	-0.00759	0.1184
Δ Pct Female	-0.00079	0.00762	-0.0281	0.0224
Δ Pct Non-Family Households	0.0360	0.0195	-0.0196	0.0923
Δ Annual Maximum Temp., Yr. t	-3.18	3.28	-12.88	4.73
Δ Annual Minimum Temp., Yr. t	-4.38	3.44	-11.93	10.18
Δ Winter Precipitation, Yr. t	0.5283	1.8830	-12.4067	4.8600

Table 2: Summary statistics for 187 cities in sample: absolute changes in all variables between 1980 and 1990. Variable definitions are as in Table 1.

OLS Analysis for the Retail Index				
	Dependent Variable:			
Independent Variables	**Δ Log of Retail Price Index**			
Δ Percent Under Poverty	-0.14 (-1.35)	0.76 (3.94)		
Δ Pct Over 2x Poverty		0.69 (5.36)		
Δ Pct in 1st Income Quintile			0.40 (1.7)	0.52 (2.64)
Δ Pct in 3rd Income Quintile			0.40 (1.16)	
Δ Pct in 4th Income Quintile			0.27 (1.15)	
Δ Pct in 5th Income Quintile			0.60 (3.44)	
Δ Pct in Top 3 Income Quintiles				0.57 (4.11)
Δ Other Control Variables	Yes	Yes	Yes	Yes
Adjusted R^2	0.33	0.42	0.40	0.40
N	187	187	187	187

Table 3: OLS Regression Results: Other control variables are total state and local sales tax, percent of the population over 65 years old, percent Hispanic, percent female, percent of households that are not family households, and log of the population density. All regressions include an intercept. t-statistics in parentheses.

OLS Analysis for the Retail Index using Poverty Categories					
	Dependent Variable				
Independent Variables	**ΔAll**	**ΔGroc**	**ΔMisc**	**ΔHealth**	**ΔTrans**
Δ Pct Under Poverty	0.76	0.46	1.50	1.08	0.05
	(3.94)	(1.84)	(4.12)	(2.03)	(0.11)
Δ Pct Over 2x Poverty	0.69	0.40	1.26	0.53	0.36
	(5.36)	(2.39)	(5.21)	(1.47)	(1.19)
Δ Other Control Variables	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.42	0.39	0.26	0.14	0.3
N	187	187	187	187	187

Table 4: IV Estimation for Retail Index Using Quintile Categories: Other control variables are defined in Table 3. Only income distribution variables are instrumented. t-statistics in parentheses.

OLS Analysis for the Retail Index using Quantile Categories					
	Dependent Variable				
Independent Variables	**ΔAll**	**ΔGroc**	**ΔMisc**	**ΔHealth**	**ΔTrans**
Δ Pct in 1st Income Quintile	0.52	0.39	1.11	1.18	-0.64
	(2.64)	(1.58)	(2.97)	(2.20)	(-1.42)
Δ Pct in Top 3 Income Quintiles	0.57	0.43	1.04	0.77	-0.18
	(4.11)	(2.46)	(3.95)	(2.04)	(-0.57)
Δ Other Control Variables	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.4	0.4	0.22	0.14	0.3
N	187	187	187	187	187

Table 5: IV Estimation for Retail Index Using Quintile Categories: Other control variables are defined in Table 3. Only income distribution variables are instrumented. t-statistics in parentheses.

Pooled, Cross Sectional OLS Analysis		
Independent Variables	**Dependent Variable: Log of Retail Price Index**	
Percent Under Poverty	0.582 (4.11)	
Pct Over 2x Poverty	0.457 (4.81)	
Pct in 1st Income Quintile		0.667 (5.80)
Pct in Top 3 Income Quintiles		0.634 (7.10)
Other Control Variables	Yes	Yes
Adjusted R^2	0.48	0.52
N	374	374

Table 6: Pooled Cross Sectional OLS: Log of retail price index (not differenced) regressed on income distribution variables, a year dummy for 1989-90, and the other control variables used in Table 3. All regressions include an intercept. t-statistics in parentheses.

Metropolitan Area Name	1990 Population
Boston–Lawrence–Salem, MA–NH (CMSA)	4,171,643
Chicago–Gary–Lake County, IL–IN–WI (CMSA)	8,065,633
Cleveland–Akron–Lorain, OH (CMSA)	2,759,823
Dallas–Fort Worth, TX (CMSA)	3,885,415
Detroit–Ann Arbor, MI (CMSA)	4,665,236
Houston–Galveston–Brazoria, TX (CMSA)	3,711,043
Los Angeles–Anaheim–Riverside, CA (CMSA)	14,531,529
Miami–Fort Lauderdale, FL (CMSA)	3,192,582
New York–Northern N.J.–Long Island, NY–NJ–CT (CMSA)	18,087,251
Philadelphia–Wilmington–Trenton, PA–NJ–DE–MD (CMSA)	5,899,345
Pittsburgh–Beaver Valley, PA (CMSA)	2,242,798
San Francisco–Oakland–San Jose, CA (CMSA)	6,253,311
Baltimore, MD (MSA)	2,382,172
St. Louis, MO–IL (MSA)	2,444,099
Washington, DC–MD–VA (MSA)	3,923,574

Table 7: Metropolitan areas for which BLS price data is available.

OLS Analysis for 15 Metropolitan Areas in 1980 and 1990				
Dependent Variable: Log of BLS Price Index				
Independent Variables	**Differences**		**Cross Section**	
Percent Under Poverty	1.406		0.30	
	(1.48)		(0.82)	
Pct Between 1x and 2x Poverty		-0.491		-0.18
		(-2.67)		(-1.67)
Pct Over 2x Poverty	0.880		0.22	
	(2.01)		(1.37)	
Adjusted R^2	0.30	0.31	0.996	0.996
N	15	15	30	30

Table 8: OLS Regression Results for 15 Metropolitan Areas using BLS price data. Price index is all items excluding shelter. Regressions are weighted by mean population in 1980 and 1990. All regressions include an intercept and a time dummy but no other control variables. t-statistics in parentheses.

IV Estimates for Δ Log of Retail Index Using Poverty Categories					
	Dependent Variable				
Independent Variables	**ΔAll**	**ΔGroc**	**ΔMisc**	**ΔTrans**	**ΔHealth**
Δ Pct Under Poverty	2.76	2.37	3.49	3.49	5.56
	(2.75)	(2.06)	(2.13)	(1.73)	(2.09)
Δ Pct Over 2x Poverty	1.96	1.63	2.82	2.82	3.29
	(2.44)	(1.78)	(2.16)	(1.75)	(1.55)
Δ Structural Weather Variables	Yes	Yes	Yes	Yes	Yes
Δ Other Control Variables	Yes	Yes	Yes	Yes	Yes
Basmann Test (P-Value)	0.565	0.795	0.518	0.185	0.116
Adjusted R^2	0.40	0.46	0.30	0.44	0.16
N	184	184	184	184	184

Table 9: IV Estimation for Retail Index Using Quintile Categories: Other control variables are defined in Table 3. Structural weather variables are 1979-89 and 1980-90 changes in (a) spring, summer, fall, and winter average precipitation and temperatures and (b) minimum and maximum annual temperatures and their squares. Only income distribution variables are instrumented. t-statistics in parentheses.

IV Estimates for Δ Log of Retail Index Using Quantile Categories					
	Dependent Variable				
Independent Variables	**ΔAll**	**ΔGroc**	**ΔMisc**	**ΔTrans**	**ΔHealth**
Δ Pct in 1st Income Quintile	1.63	0.58	2.79	0.67	4.84
	(2.24)	(0.71)	(2.06)	(0.50)	(2.22)
Δ Pct in Top 3 Income Quintiles	0.95	0.13	2.20	0.33	2.50
	(1.72)	(0.21)	(2.13)	(0.32)	(1.50)
Δ Structural Weather Variables	Yes	Yes	Yes	Yes	Yes
Δ Other Control Variables	Yes	Yes	Yes	Yes	Yes
Basmann Test (P-Value)	0.087	0.212	0.495	0.004	0.118
Adjusted R^2	0.48	0.55	0.30	0.563	0.17
N	184	184	184	184	184

Table 10: IV Estimation for Retail Index Using Quantile Categories. Other control variables are defined in Table 3. Structural weather variables are defined in Table 9. Only income distribution variables are instrumented. t-statistics in parentheses.

Effects of Explanatory Variables				
	Dependent Variable: ΔAll			
Independent Variables	**OLS**	**IV**	**OLS**	**IV**
Δ Pct Under Poverty	0.68	2.54		
	(3.35)	(2.72)		
Δ Pct Over 2x Poverty	0.56	1.83		
	(3.71)	(2.37)		
Δ Pct in 1st Income Quintile			0.47	1.49
			(2.35)	(2.17)
Δ Pct in Top 3 Income Quintiles			0.45	0.89
			(2.98)	(1.66)
Δ Log Household Size	-0.09	-0.24	-0.12	-0.29
	(-0.60)	(-1.00)	(-0.81)	(-1.60)
Δ Pct Households with No Cars	0.06	-0.13	0.095	-.017
	(1.22)	(-1.53)	(2.07)	(-0.28)
Δ Log Property Crime	0.01	-0.001	0.006	0.004
	(1.06)	(-0.19)	(1.09)	(0.67)
Δ Pct Stores with $>$ 9 employees	-0.09	-0.32	-0.057	-0.17
	(-0.96)	(-2.30)	(-0.61)	(-1.51)
Δ Log of Housing Price Index	0.01	-0.0006	0.010	-0.006
	(1.36)	(-0.04)	(1.17)	(-0.41)
Δ Structural Weather Variables	No	Yes	No	Yes
Δ Other Control Variables	Yes	Yes	Yes	Yes
Basmann Test (P-Value)	-	0.121	-	0.286
Adjusted R^2	0.42	0.44	0.41	0.49
N	187	184	187	184

Table 11: Estimation With Alternative Hypothesis Variables.