Journal of Economic Perspectives—Volume 23, Number 2—Spring 2009—Pages 77–97

The Role of Prices in Measuring the Poor's Living Standards

Christian Broda, Ephraim Leibtag, and
David E. Weinstein

Almost 50 years after President Lyndon Johnson's famous 1964 State of the Union speech that introduced the "War on Poverty," two facts stand out in the current debate about poverty. First, since David Caplovitz (1963) wrote his path-breaking book, *The Poor Pay More,* numerous researchers have confirmed that the poor indeed pay more than households of higher income for the goods and services they purchase. Second, official poverty rates as measured by the U.S. Census have remained essentially flat since the late 1960s, raising questions about the success of the policies implemented to reduce poverty. In this paper we revisit these two facts by paying close attention to the price data underlying these findings. By examining scanner data on thousands of household purchases we find that the *poor pay less*—not more—for the goods they purchase. In addition, by extending the advances on price measurement in the recent decade back to the 1970s, we find that current poverty rates are *less than half* of the official numbers. This finding underscores the importance of correctly measuring the evolution of prices to determine the appropriate poverty thresholds over time. Both findings are contrary to the conventional wisdom established in the last few decades.

We start by addressing the question of whether the poor pay more for the goods they buy. Since Caplovitz's (1963) book, a number of papers have seemed to confirm his main thesis that the poor pay more. Surveys of food stores often

■ *Christian Broda is Professor of Economics and Business, Booth School of Business, University of Chicago, Chicago, Illinois. Ephraim Leibtag is an Economist with the Economic Research Service, U.S. Department of Agriculture, Washington, D.C. David E. Weinstein is Carl S. Shoup Professor of Japanese Economy, Columbia University, New York City, New York. Broda is a Faculty Research Fellow and Weinstein is a Research Associate, both with the National Bureau of Economic Research, Cambridge, Massachusetts. Their e-mail addresses are ⟨cbroda@chicagobooth.edu⟩, ⟨eleibtag@ers.usda.gov⟩, and ⟨dew35@columbia.edu⟩.*

conclude that low-income households shop in stores with higher item prices. For example, Chung and Myers (1999) and Cotterill and Franklin (1995) have found that poor neighborhoods have fewer discount stores than the suburbs and therefore poor people are likely to shop more frequently in higher-priced small convenience stores. However, no information about household expenditures by store type was used in these studies, and so it was impossible to quantify with any precision the magnitudes of these effects. Other papers have relied on survey data that only provide the unit costs paid for broad food categories by different income groups, rather than the actual prices paid for specific items (for example, Kaufman, MacDonald, Lutz, and Smallwood, 1997).[1] As pointed out by Attanasio and Frayne (2005), the use of unit costs makes it difficult to distinguish whether the poor pay more for identical goods or choose lower-quality varieties of these products.

In this paper, we circumvent the problems of previous studies by using a dataset that contains actual purchases of around 40,000 households collected by Nielsen. By focusing on the actual prices paid by households, we show that poor households systematically pay *less* than richer households for identical goods. The poor pay less in part because they shop in cheaper stores and in part because they pay less for the same goods even in the same store. This latter effect probably arises because poorer households are more likely than richer households to buy goods on sale, even in the same store.[2] We also confirm that the poor shop more in convenience stores—where prices are 11 percent higher than in traditional grocery stores—but show that this effect is dominated by their higher share of expenditure in supercenters where prices are 10 percent *lower* than in grocery stores.

The second fact that has emerged in the literature on poverty is related to how the Consumer Price Index (CPI) affects the evolution of poverty rates over time. For many years, the CPI used to be based on comparing the price of a fixed basket of goods at different points in time. As was widely recognized (for an excellent summary, see Lebow and Rudd, 2003), this approach tended to overestimate the true rise in the cost of living for two main reasons: 1) the fixed basket of goods did not allow for substitution from goods that became relatively more expensive to goods that became relatively less expensive; and 2) the fixed basket of goods does not properly adjust for new goods or the quality improvements in existing goods (for discussion in this journal, see the symposia in the Winter 1998 and Winter 2003 issues). In the last 10 years or so, the CPI has been dramatically improved. It is no longer calculated on the basis of a fixed basket of goods, but rather with an approach that allows for some substitution between goods over time. Both goods of improved quality and new goods are rotated more quickly into the mix of goods

[1] Similarly Leibtag and Kaufman (2003) use Nielsen HomeScan data to look at the differences in average prices across product groups (for example, milk and cheese) and also found poorer households had lower unit value expenditures.

[2] This could also be consistent with findings in Aguiar and Hurst (2007) that suggest that prices people pay are related to the value of time and the amount of time that people decide to invest in shopping. If poor people were to shop more, they would pay less.

whose prices are being compared, and in some cases hedonic adjustments are used for quality changes as well, although many analysts believe that the CPI does not yet fully capture the gains of improved quality and new goods.

While these changes have improved how well the Consumer Price Index captures the rise in the true cost of living going forward, they are not projected backward. Looking back over several decades, the official poverty and real wage measures are based on the earlier, unimproved price indexes. The effect of correctly measuring the evolution of price levels has important implications for our perceptions of how poverty rates have evolved over time. While official measures of poverty rates (and real wage growth, for that matter) appear to have moved very little over the last 40 years, much of this arises from the fact that measurements of the change in price levels have been consistently *overstating* the cost of living. Properly measured, real wage growth has been robust—even for the lowest 10 percent of workers—and poverty rates based on the official poverty line but using an improved price index would measure less than half of the official numbers. In this sense, the United States has been making gradual progress in the war on poverty, but without acknowledging it.

Prices Actually Paid by the Poor

Past research suffered from the lack of data detailed enough to match prices paid with the households that actually made the purchases. As a result, prior work focused on inferring the linkages between prices paid and household characteristics—for example, by using approaches based on neighborhood effects and unit costs. Here, we first describe household-level data that allows us to look at the linkages between prices paid and household characteristics. We then use that data to reconsider the commonly held notion that the poor pay higher prices than those with high incomes and that this behavior is driven by the larger share in expenditure of the poor in high-priced convenience stores.

Data on the Prices That Households Actually Pay

In this paper, we work with one of the richest datasets of consumer behavior available. We use 2005 Nielsen Homescan household-level data that contain information on every purchase of a food item with a barcode made by approximately 40,000 demographically representative households across the United States. This unique dataset allows the comparison of the prices paid by goods with the same barcode—for example, a 32 oz. bottle of Orange Gatorade—across households with different incomes. The data is collected from handheld scanners that Nielsen provides to each household. After each shopping trip, households scan the barcodes of every purchased item with a barcode. If the good is purchased from a chain for which Nielsen already has store data, the weekly average price is included directly from the store scanner data.[3] Otherwise, the household is asked to hand enter the price inclusive of any coupons or discounts that applied to the purchase. Once the data of each household is collected, Nielsen produces a unique dataset that provides the price and quantity for a large set of food and grocery products together with household and store characteristics.

We focus here on food products, more specifically approximately 290,000 barcodes (approximately 40 percent of the universe of barcode goods analyzed in Broda and Weinstein, forthcoming). The data provides a rich set of information about household characteristics including household size, age of household head or heads, race, marital status, and zip code. For each of the purchases we know not only the price and quantity, but certain information about the store: its zip code; the store's format (grocery store, mass merchandiser, drug store, supercenter, club store, or convenience store); and the store name if it is part of a store chain that is tracked by Nielsen. A mass merchandiser is a retail outlet that primarily sells nonfood items but does have some limited nonperishable foods items available, while a supercenter is an expanded mass merchandiser that also sells a full selection of grocery items. A warehouse club store sells both food and nonfood items in bulk quantities. To get some sense of how the store breakdown is constructed in the Nielsen classification system, Safeway is classified as a grocery store, Rite Aid as a drug store, Wal-Mart as a supercenter, Target as a mass merchandiser, Costco as a club store, and Seven Eleven as a convenience store.

Do Poor Households Buy Goods at Smaller Stores?

One argument in favor of the notion that poorer households pay more for the same goods is that discount retailers tend to locate in suburban areas that may be hard for poor households to reach. Thus, it is commonly argued that poor households are forced to purchase their goods in smaller, higher-priced stores while wealthier households purchase their goods in discount stores that offer lower prices.

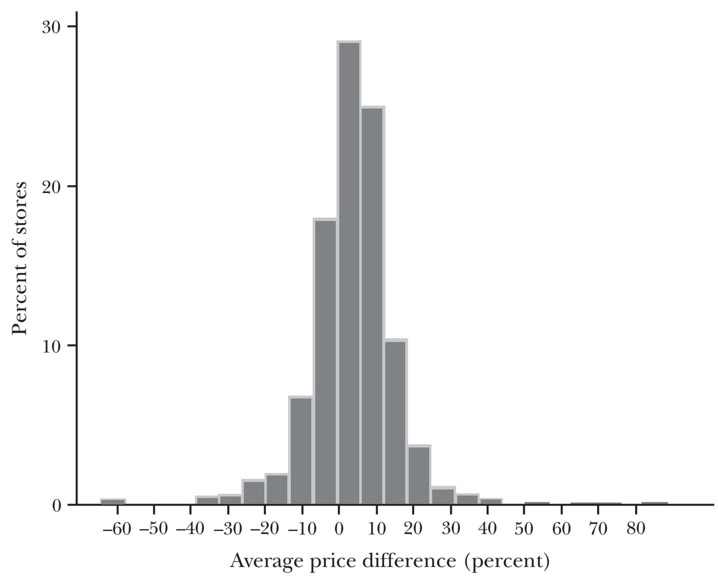
We begin our exploration of this subject by investigating how prices differ across stores. The data contains store names for 452 different retailers. To examine how prices differ across retailers, we first computed the log price paid by every consumer in every store and subtracted the average log price of that good in all stores. We next averaged these price differentials by store to obtain a sense of how average prices varied across stores for the same good.

We plot the distribution of these differentials in Figure 1, which shows that consumers obtain systematically different prices in different stores. The cheapest 10^{th} percentile of stores has prices that are at least 23 percent less than the upper 10^{th} percentile. This pattern suggests that having access to cheaper stores can have a significant impact on real income.

Another interesting feature of the histogram is that the distribution is not

[3] For a more detailed discussion of the implications of this procedure on the Homescan data, see Einav, Leibtag, and Nevo (2008), available at ⟨http://www.ers.usda.gov/Publications/ERR69/⟩.

Figure 1
Distribution of Estimated Average Store Price Differences



centered around zero. Instead the median store in the sample has a price differential of 4.4 percent (4.4 percent more expensive than the average store). This pattern arises from the fact that there are many low-volume stores with higher prices compared with a relatively small number of large deep discounters—a feature of the data that we will explore later.

While it may be the case that supercenters do not locate in the poorest neighborhoods, that does not necessarily mean that poorer households do not have access to these stores. We looked at households by income level and examined what share of their food expenditures took place in different kinds of stores: grocery, drug, mass merchandise, supercenter, club, and convenience stores. The results indicate that while there are some differences in the share of expenditures in different store types, the differences are small. All income categories, from our lowest category of $5,000–$7,999 to our second-highest category, up to $100,000, spend from 52–57 percent of their food purchases at grocery stores. The highest income category, above $100,000, spends 59 percent of its food dollars at grocery stores. While the lowest three income categories of households—from $5,000 to $11,999—spend from 2 to 3 percent of their total expenditures in convenience stores, the highest-income category households spend 0.7 percent of their expenditures there. While prior work is correct to argue that those with lower incomes tend to purchase a greater share of their goods in convenience stores, the difference appears to be quite modest.

Perhaps most surprising is the fact that low- and middle-income households are more likely to purchase food at supercenters, where prices are lower. While households from $5,000 to $49,999 spend 20–22 percent of their food dollars at supercenters, those with incomes above $70,000 spend 13–17 percent of their food dollars in supercenters. However, those with high incomes are much more likely to spend their food dollars at club stores. These stores, while not typically offering discounts that are as deep as supercenters (a point we will quantify later), also offer significant discounts. Households with income levels in the two lowest categories, from $5,000 to $9,999, spend only 6–9 percent of their food dollars in club stores, while those in the highest income categories, with more than $70,000 in income, spend 14–17 percent of their food dollars in club stores. Interestingly, the usage of club stores and supercenters combined together is essentially the same across all income categories.

Overall, we find little difference in the type of stores in which poor and rich households shop for food. Thus, the data contradict the notion that lower-income households are forced to shop at high-priced convenience stores because they lack access to other types of stores.[4]

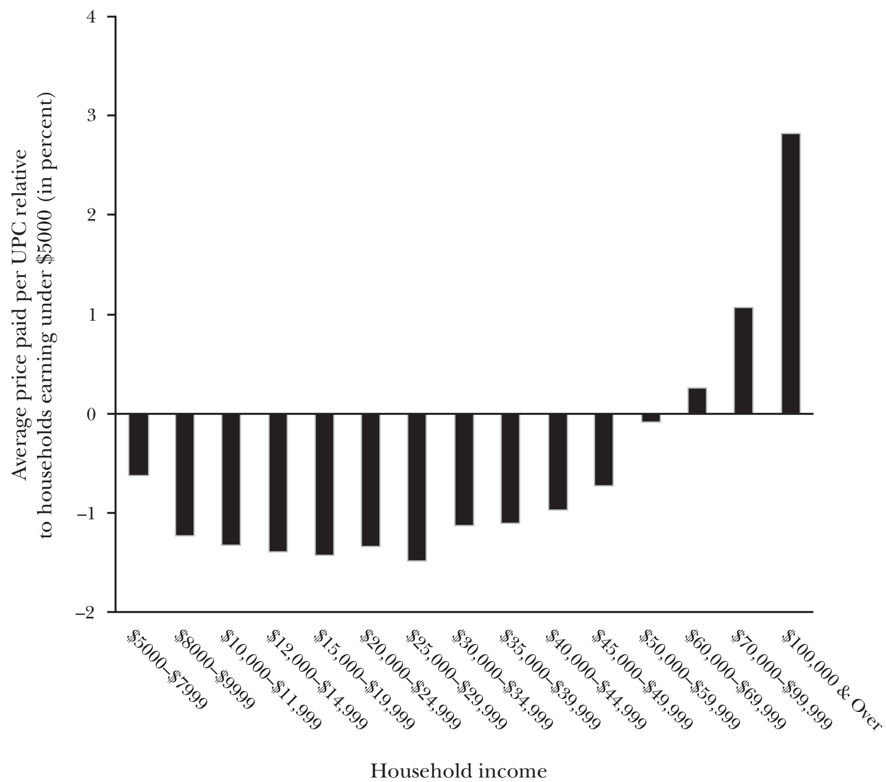
The Poor Pay Less

Given these shopping patterns, do those with low incomes pay higher prices for food? We used the log of the price paid per barcode (Universal Product Code or UPC) as our dependent variable in a series of regressions, in which each regression includes barcode fixed effects and a series of dummy variables that describe the main characteristics for each household (like the household's income level). We report the point estimates for these regressions in Figure 2. The results indicate that while some of the very poorest households—those earning less than $8,000 per year—may pay between 0.5 percent and 1.3 percent more for their groceries than households earning slightly more, households earning between $8,000 and $30,000 tend to pay the least for groceries, whereas higher-income households pay significantly more. In particular, households earning over $100,000 per year pay between 2–3 percent more than poorer households. In short, the conventional wisdom that the poor pay higher prices is not present in a dataset that precisely tracks purchases of individual goods by different households.

To gain a better understanding of why households pay the prices they do, we run a series of additional regressions. In each regression, the dependent variable is the log of the price of a good. The explanatory variables always include a fixed effect for the good—based on a unique barcode—and then various other variables concerning household and store characteristics. Because the income data breaks households up into income classes, we do not have a continuous measure of household income. While we could run every regression with 16 income categories, this results in a large number of reported coefficients. In the interests of parsimony,

[4] There is some concern that Homescan data underrepresent households in the lowest part of the income distribution and these households may be the ones that face the most limited access to supermarkets and supercenters. Additional research is needed to estimate how this underrepresentation, if it exists, may affect our results.

Figure 2
Relationship between Income and Prices Paid



Note: Bars represent coefficients on household income dummies in a regression of log price per unit on UPC fixed effects and income dummies.

we recoded the income values of the households to the average value for each income bracket. For example a household earning between $12,000 and $15,000 per year was assigned an income equal to $13,500. Once we do this, we can create a variable equal to the log of the income in that category. Our results are qualitatively unaffected by using income categories, but we were forced to drop the highest-income category of households, because the midpoint of "$100,000 and up" is not well-defined.

Table 1 reports the results from this regression. As one can see in column 1, a 10 percent increase in income is associated with roughly a 0.1 percent increase in the prices paid per item. This result implies that households earning $80,000 per year pay about 2.9 percent more than households earning only $10,000 per year. In other words, the regression results with the continuous variable present a picture quite similar to what we saw in Figure 2.

The results in column 1 do not control for various demographic characteristics. For example, larger households are going to be poorer on a per-person basis

Table 1
Impact of Income on Prices Paid

Dependent variable	(1) *ln(price)*	(2) *ln(price)*	(3) *ln(price)*	(4) *ln(price)*	(5) *ln(price)*	(6) *ln(price)*	(7) *ln(price)*
Log Household Income	0.0108***	0.0131***	0.0119***	0.0121***	−0.0901***	0.0089***	0.0088***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0045)	(0.0001)	(0.0001)
Log Income in Household's Zipcode			0.0289***	0.0134***	−0.0746***		0.0068***
			(0.0003)	(.0007)	(.0046)		(.0008)
Log Income in Store's Zipcode				0.0176***			0.0021***
				(0.0007)			(0.0008)
ln (Household Income) • ln(Avg. Zipcode Income)					0.0098***		
					(0.0004)		
UPC Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Household characteristic controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Store Controls	No	No	No	No	No	Yes	Yes
N	1.73E+07	1.45E+07	1.45E+07	1.11E+07	1.45E+07	1.09E+07	8.64E+06
(within) R^2	0.0007	0.0121	0.0126	0.0148	0.0127	0.0784	0.0638

Notes: Due to limitations set by the data, the log of household income and the interaction terms were calculated based on averages of ranges. For example, a household whose annual income was $12,300 would be in the range $12,000–$15,000 and would have been assigned the value of $13,500. "Household characteristic controls" include size, age, race, marital status, and city of residence. No city controls were used in the regressions with store names. Standard errors are in parentheses.
***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

than smaller households with the same income. In column 2, we add controls for a slew of demographic characteristics to the regression. In these regressions, we add dummies for household size, race, marital status, and age of the adult members. Adding these controls increases the strength of the relationship between price paid and income. The elasticity of price paid to income rises from 0.011 to 0.013. The belief that poorer households pay more is apparently wrong. On average, households with higher incomes pay slightly more.

A second hypothesis often discussed in the literature is that poor households have less access to low-priced goods because stores face relatively less competition in poor neighborhoods than in wealthier ones. To examine this hypothesis, we added in the log of the per capita income of the zip code in which the household lives. The results are reported in column 3 of Table 1. Once again, the conventional wisdom seems backwards. Even after controlling for household income, a person living in a wealthier neighborhood typically pays more for food—perhaps because stores are nicer or charge higher prices—than a household living in a poorer neighborhood. We find that households that live in zip codes that are two standard deviations poorer than the average zip code pay 1.4 percent less for the same items than households that live in zip codes that are two standard deviations above average, after controlling for income level and demographic characteristics. Clearly, the mix of stores in wealthier neighborhoods does not translate into lower prices for shoppers.

Why Do Those with Higher Incomes Pay More?

One can imagine a number of hypotheses why people living in higher-income neighborhoods pay more for food. One possible explanation is that a good with a certain barcode purchased in a particular store may not truly be identical with same good in a different store because the shopping experiences in the different stores are not the same (Betancourt and Gautschi, 1993, offer a more detailed discussion of the services provided by retailers). Alternatively, poor people may invest more time comparing the prices of goods or learn more from their neighbors about which stores offer deeper discounts. Finally, perhaps rents are higher in higher-income neighborhoods, and so people shopping in those neighborhoods face higher prices in general.

These hypotheses are not mutually exclusive, and we turn next to trying to quantify their importance. In column 4, we first include both the log of the per capita income of the household's address and the log of per capita income of the zip code in which the store is located. If one thinks that information sharing is important, one should expect that the coefficient on the income of the zip code where the household lives to be critical. On the other hand, if rents are the key driving force, then one should expect that the coefficient on the per capita income of the store's zip code would be the critical factor. As one can see from the results in column 4, both effects are important. Although there is a statistically significant and positive relationship between the average per capita income in the store zip code and the price paid, the relationship is not economically significant. A store in a neighborhood with twice the per capita income of another neighborhood only has prices that are one percent higher. Apparently, very little of the apparent higher cost of shopping in wealthier neighborhoods is due to the same goods being sold at higher prices.

It is also sometimes argued that poor people might purchase goods for less if they had better access to discount stores. To examine this hypothesis, we interacted the income of the household with that of the log per capita income of the zip code in which the household lives. If poorer people are better able to find bargains when they live in wealthier neighborhoods, one should expect to find a negative coefficient on the interaction term. On the other hand, if bargains are less available in wealthier neighborhoods, one should expect to see a positive sign. Column 5 shows a positive coefficient on the interaction term—that is, low-income people living in poor neighborhoods tend to pay less for the same items as low-income people in high-income neighborhoods.[5]

One possible reason that poor people pay less is that they shop in stores that provide fewer or lower-quality amenities. Stores in poorer neighborhoods might be

[5] The negative coefficients on the household and zip code variables do not indicate that prices are higher for poorer people, because the elasticity of price paid with respect to income is also a function of the per capita income in the store's zip code. The elasticity of price with respect to income is rising for purchases in stores located in zip codes that have per capita incomes within two standard deviations of the average.

less clean, less well-staffed, and offer less customer service than stores in wealthier neighborhoods. In other words, there may be a positive correlation between household income and the quality of the store at which the consumer shops that spuriously leads us to believe that poorer households face lower prices, when in fact poorer households simply receive fewer services with their goods.

Since the quality of the services provided in a store is likely to be highly correlated with the store in which the good was purchased, we decided to control for this possible bias by including a vector of retail chain dummies in our regression. To the extent that the quality of the services provided in one branch of a chain of stores is similar to that provided in another branch, this would eliminate the impact of unobserved quality differences across chains on the prices paid by consumers. Column 6 of Table 1 shows that adding chain controls does result in a small drop in the elasticity of price paid with respect to income, but the effect is small. Comparing the coefficient of .013 on the log of household income from column 2 with the coefficient of .009 in column 6 suggests that about one-third of the higher price paid by richer households for the same good is attributable to them shopping at nicer stores while the rest is due to their shopping behavior within stores. The fact that poorer households pay less for the same goods *even in the same retail chain* indicates that the price differential between poor and wealthy households is likely to represent a shopping effect.

But how much does the access to specific types of stores matter? We can answer this question by replacing our store-chain dummies with a series of store-type dummies (Table 2). This expands the specifications presented earlier in Table 1. The first interesting result is the dramatically lower prices paid for the same good by store type. Supercenters and mass merchandisers charge 10 to 12 percent less for the same goods than do grocery stores. Similarly, convenience stores charge 11 percent more than grocery stores for the same good. Taken together, this implies that a consumer purchasing a particular barcode good in a convenience store pays about 20 percent more than if the same good was purchased at a supercenter or mass merchandiser. This pattern suggests that consumers are willing to pay substantial premiums for the convenience of buying a certain good at a nearby location.

These price estimates can be combined with the expenditure by store types to examine the hypothesis that poorer households face higher prices because they purchase more of their groceries in convenience stores. If we multiply the average price differential by store type estimated in column 1 of Table 2 with the share of expenditures by income group in that store type and sum across all store types, we can compute the effect of different shopping patterns on prices paid. We found that although the lowest-income households do tend to spend a slightly higher share of their food dollar in convenience stores than higher-income households, they purchase so much more food in supercenters that it more than offsets the higher prices they pay in convenience stores. Higher-income households tend to shop in a more expensive mix of stores, and this tends to raise the prices they pay for the same goods by around 0.8 percent.

Table 2
Impact of Income, Store Type, and Race on Prices Paid

Dependent variable	*ln(price)*	*ln(price)*	*ln(price)*	*ln(price)*
	(1)	(2)	(3)	(4)
Log Household Income	0.0115***	0.0108***	0.0117***	−0.0635***
	(.0001)	(.0001)	(.0001)	(.0044)
Log Income in Household's Zipcode		0.0176***	0.0104***	−0.0579***
		(.0003)	(.0007)	(.0045)
Log Income in Store's Zipcode			0.0059***	
			(.0007)	
ln(Household Income) • ln(Avg. Zipcode Income)				0.0071***
				(.0004)
Race: Black	0.0022***	0.0014***	0.0016***	0.0015***
	(.0002)	(.0002)	(.0003)	(.0002)
Race: Asian	−0.0213***	−0.0222***	−0.0199***	−0.0224***
	(.0005)	(.0005)	(.0006)	(.0005)
Race: Hispanic	−0.0002	−0.0001	−0.0017***	0.0000
	(.0003)	(.0003)	(.0004)	(.0003)
Race: Other	−0.0003	−0.0005	0.0004	−0.0006
	(.0004)	(.0004)	(.0004)	(.0004)
Store type dummy (Drug Store)	−0.0605***	−0.0609***	−0.0389***	−0.0608***
	(.0005)	(.0005)	(.0006)	(.0005)
Store type dummy (Mass Merchandiser)	−0.1174***	−0.1172***	−0.1007***	−0.1172***
	(.0004)	(.0004)	(.0004)	(.0004)
Store type dummy (Super Center)	−0.0961***	−0.0948***	−0.0963***	−0.0948***
	(.0002)	(.0003)	(.0003)	(.0002)
Store type dummy (Club Store)	0.0784***	0.0782***		0.0781***
	(.0009)	(.0009)		(.0009)
Store type dummy (Convenience Store)	0.1062***	0.1065***		0.1065***
	(.001)	(.001)		(.001)
Store type dummy (Other)	−0.1509***	−0.1507***		−0.1506***
	(.0003)	(.0003)		(.0003)
UPC fixed effects	Yes	Yes	Yes	Yes
Controls for household characteristics	Yes	Yes	Yes	Yes
Fresh produce included	No	No	No	No
N	1.45E+07	1.45E+07	1.11E+07	1.45E+07
R^2	0.0408	0.041	0.0304	0.041

Notes: Due to limitations set by the data, the log of household income and the interaction terms were calculated based on averages of ranges. For example, a household whose annual income was $12,300 would be in the range $12,000–$15,000 and would have been assigned the value of $13,500. Race dummies are relative to "White." Store type dummies are relative to "Grocery." Household characteristic controls include size, age, race, marital status, and city of residence. Standard errors are in parentheses. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

The data also reveals some evidence of racial differences in the prices paid for goods. African American households pay about 0.2 percent more for the same goods than white households, and most mysteriously, Asian households pay 2 percent less than do white households. Our data does not allow us to investigate the underlying causes of these patterns, which may have roots in some form of discrimination or in cultural differences in shopping behavior. However, there does not seem to be any evidence that minorities are charged systematically higher prices.

Quality and Variety by Income Group

One of the other ways that lower-income households might save money is by buying lower-quality varieties of the same goods (although this would not have clear welfare implications, as we discuss below). Our data show this pattern in two different ways. Fresh produce is the one exception to the basic rule in the data that the quality of a good is identical for all goods sold with the same barcode. The quality of, say, a Chiquita banana, may vary depending on how long the banana has been sitting on the shelf—and stores may charge different prices for the same brand of banana depending on how fresh it is.

To examine how important this effect is in our data, we split the sample into items that are sold at varying weights ("random weight" goods) and those that are sold in fixed weight common units. Random weight goods are much more likely to have quality, and therefore price, variation that is determined by the freshness of the goods. Thus, while one can of Campbell's Alphabet Soup is indistinguishable from another, the same may not be true for bananas. If lower-income households tend to purchase lower-quality versions of the same goods, one should expect to see the poor pay even lower relative prices for random weight goods.

Table 3 presents results in which we split the sample into random weight and nonrandom weight goods. Goods that are not sold by random weight constitute the vast majority of our sample and thus the coefficient estimates presented in the first two columns, which are based on the sample of nonrandom weight goods, are quite similar but a little smaller than those in Table 1. Dropping the random weight goods suggests that every log unit (that is, 69 percent) increase in household income is associated with a 0.8 percent increase in the price paid for identical goods. This set of regressions provides some of the strongest evidence we have that lower income households pay less, not more, for the same goods.

The last two columns of Table 3 present the results for random weight goods. The results of these regressions suggest that the prices paid for fresh produce are approximately five times more sensitive to income than other prices. Comparing the results from column 1 and column 3 suggests that 20 percent (one fifth) of this effect can be attributed to the better ability (perhaps through effort) of the poor to find bargains, while the remaining 80 percent is due to the poor finding lower-quality versions of the same produce.

The magnitude of this estimate suggests that the gains to households from shopping are likely to be quite modest, but the quality variation in household consumption is substantial. To examine this for a broader set of goods, we turn to a dataset that has been analyzed elsewhere. Broda and Weinstein (2008) work with a sample of 3000 Nielsen "Homescan" households that were surveyed in the fourth quarter of 2003. Although their sample of households is significantly smaller, the sample covers a much wider range of product categories than food.

To investigate the importance of quality variation in household purchases, we

Table 3
Impact of Income on Prices Paid: With and Without Produce

	Without random weight items		*With random weight items*	
Dependent Variable	*ln(price)*	*ln(price)*	*ln(price)*	*ln(price)*
	(1)	(2)	(3)	(4)
Log Household Income	0.0079***	−0.0296***	0.0408***	−0.0241
	(.0001)	(.0045)	(.0021)	(.0677)
Log Income in Household's Zipcode	0.0069***	−0.0294***	0.0112	−0.052
	(.0007)	(.0046)	(.0107)	(.0688)
Log Income in Store's Zipcode	0.0027***		−0.0081	
	(.0007)		(.0108)	
ln(Household Income) • ln(Avg. Zipcode Income)		0.0036***		0.0061
		(.0004)		(.0064)
UPC Fixed Effects	Yes	Yes	Yes	Yes
Controls for Household Characteristics	Yes	Yes	Yes	Yes
Store Controls	Yes	Yes	Yes	Yes
N	8.36E+06	1.06E+07	2.75E+05	3.36E+05
R^2	0.0764	0.0929	0.0474	0.0859

Notes: Due to limitations set by the data, the log of household income and the interaction terms were calculated based on averages of ranges. For example, a household whose annual income was $12,300 would be in the range $12,000–$15,000 and would have been assigned the value of $13,500. Household characteristic controls include size, age, race, marital status, and city of residence. Standard errors are in parentheses.
***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

regressed the log average price paid by a household for all barcodes in a product group (like milk) on the log of the household's average expenditure per adult. If we believe that higher-income households are likely to buy more expensive organic milk while lower-income households are likely to buy cheaper varieties, then one should expect to see even stronger positive relationships between income and prices paid than we saw at the barcode level.

On some level, one might not expect to see much of an effect at the product group level because the definition of these goods is quite narrow. In many national and international price comparison studies, categories of goods like "butter," "eggs," and "sugar" are classified as homogeneous goods. To some extent, one can view the examination of how the prices of these goods varies with income as an examination of one of the basic assumptions of international and national price comparison studies. If there are not large income-based differences in the average prices paid by households for these goods, then this implies that one can compare egg prices across locations and assume that these prices are not identifying quality differences. Since we know from our earlier analysis that the prices of identical goods hardly move with income, if the average prices of product groups are highly sensitive to income, then we conclude that the quality mix within such groups is changing in important and systematic ways. We looked at 128 product groups—

including vitamins, fragrances, soup, coffee, pain remedies, baking supplies, milk, and many more—focusing on the coefficient for how the average price paid for products in the group varied according to income. In our regressions, the median of a one log unit increase in income is an 8.6 percent increase in the amount paid in each product group. This quality effect is about 10 times larger than the shopping effect that we identified earlier. In other words, while the rich only pay slightly more for precisely the same type of eggs or milk or cheese (to name three of our product categories) than poor households, they spend much more for eggs or milk or cheese in general. This result almost surely arises from higher-income households purchasing higher-quality eggs than poorer households.

Broda and Weinstein (2008) develop a summary measure to understand the importance of the interaction between product heterogeneity, quality, and income on measured prices. To start, they decompose differences in the prices paid by different income groups into two components: differences in prices of the specific goods that they both consume, and the average price difference for goods within a product category. The first difference can be called the "shopping effect," because it captures differences in shopping behavior when purchasing the same goods—like looking for sales or purchasing in supercenters—that might cause households in a particular expenditure class to pay more or less for the same goods. The second difference is called the "quality effect" because it reflects the tendency of households in an income category to spend more or less for a particular category of goods.

Broda and Weinstein (2008) compute these indexes at the highly disaggregated "module" level. For example, in the calculation we described a moment ago, "canned fruit" was one of the 128 product categories, but in the more disaggregated Broda and Weinstein paper, canned apples, canned grapes, and canned plums are all different modules of the product category "canned fruit." Their data is about eight times more disaggregated than the 128 product categories we described earlier. In this highly disaggregated setting, differences in how much is spent for a particular category of goods is especially likely to represent a quality effect, because the goods in these disaggregated categories are so similar.[6]

The decomposition of prices across three income groups divided into terciles shows that, looking at total purchases of goods with barcodes, higher-income households spend 74 percent more than poor households (possibly buying more or different groceries), as shown in Table 4. If one simply computed the average price paid by higher-income households as the average price paid for a module, one would see that these households paid about 32 percent more than poor households for each module. However, the table reveals that 84 percent of this difference in average prices per module is due to the fact that wealthy households buy system-

[6] If you use more aggregate product categories, differences in prices are more likely to reflect different goods than differences in quality. For example, unit prices in alcoholic beverages might reflect comparisons between wine and beer and therefore not reflect quality per se. This is less of a problem if one compares beer prices.

Table 4
Decomposing Price Differences Across Income Categories

	Comparisons relative to the low-income group			
			Decomposition of price difference	
	Overall consumption difference	*Overall price difference*	*"Quality"*	*"Shopping"*
Middle-income group	34.6%	11.9%	87.6%	12.4%
High-income group	74.1%	32.9%	84.4%	15.6%

Source: Broda and Weinstein (2008)

atically more expensive goods within a module, leaving 16 percent of this difference due to the rich paying more for the same goods. So the difference in prices paid for common goods is only about 5 percent ($32\% \times .16 = 5.1\%$).

Overall, those with low incomes pay perhaps 1–3 percent less than those with high incomes for the same goods. If the poor can access the same goods as the rich at a lower price, the use of "common" price indexes between income groups misses an important difference in the relative real incomes of poor and rich households. If we believe that the lower price paid by the poor is a pure income gain, then the ratio of the real income of the 90th percentile household to that of the 10th percentile household is not 4.49 (as estimated in Congressional Budget Office, 2006) but actually 4.26, or 5 percent smaller. Under the assumption that only 50 percent of the lower prices paid by the poor implies higher real income (the other half reflecting extra time shopping or a less pleasant shopping experience), inequality would be 2.5 percent less than suggested by national statistics. Overall, the effect on the level of inequality of differences in price levels across income groups is small.

One implication for research of this strong relationship between the average price paid for a product group and the income of the purchaser is that attempts to compare prices of non-identical goods across locations—even goods that seem to be in similar product categories—are likely to mix quality and product choice differences with true price differences. Higher-income areas are likely to have higher prices for two reasons. First, as we have documented, there is a modest effect of household and neighborhood income on the price paid for goods with identical barcodes. Second, there is a much larger effect arising from the fact that higher-income households tend to purchase much more expensive varieties of the same goods. This second effect is not a real price difference between locations but rather an income-related taste difference between the locations. Since datasets that do not use barcode data are likely to confound these two effects, it is very difficult to know how to interpret price differences across locations from standard datasets.

Poverty and Prices in the Long Run

A common concern for policymakers and economists alike is that low-income U.S. workers have not experienced any real income growth in recent decades. Table 5 presents the real wage series for the lowest 10^{th} percentile of wage earners in the U.S. economy between 1979 and 2005. We normalize the wages to equal one in 1979 to highlight how real wages have evolved over time. A standard methodology to understand real wage growth is to divide wages by the CPI for urban consumers (the CPI-U), which we do in the first row of the table. Labor economists have generated an enormous body of research attempting to understand why this real wage series has not increased. Most of this research has focused on the numerator of real wages—that is, trying to understand why the wages of less-skilled, poor workers have not risen faster. However, relatively little work has been done on the denominator—that is, in thinking about how to construct a price index that captures changes in the standard of living.

Our understanding of the computation of price indexes has evolved substantially in recent decades, and the methodology used by the Bureau of Labor Statistics to compute these indexes has also changed. The Consumer Price Index is constructed using a two-tiered aggregation of prices. At the lower level, data collectors from the Bureau of Labor Statistics collect data on about 80,000 specific items each month. These prices for very specific products are combined into 305 "entry-level items," which are relatively broad categories like "new cars" or "breakfast cereal" or "bedroom furniture" and calculated for each of 38 urban areas. Then at the upper level, these price indexes for the entry-level items are combined into the overall CPI.

A long-standing historical concern with the Consumer Price Index is that it produces upwardly biased estimates of inflation due to substitution bias: the tendency of consumers to consume less of goods whose relative price is rising. An example can illustrate this bias. Assume that the typical consumer buys one bottle of Brand C cola and one bottle of Brand P cola each month at a cost of \$1 per bottle. If the price of Brand C rises to \$2 and the price of Brand P falls to 50 cents, a fixed quantity index like the old CPI would record this set of price changes as inflation; the price index would rise from 1 to 1.25—that is, $((1 + 0.5)/(1 + 1))$—revealing 25 percent inflation. However, if at least some consumers substitute away from the more costly Brand C to the relatively cheaper Brand P, then they need not pay the additional 25 percent in cost. To be sure, those who substitute away do suffer a reduction in utility, since in the absence of the price change they would have preferred another choice, but their reduction in utility is ameliorated to some extent by the availability of a fairly close substitute. Because the CPI used to be calculated on the basis of the overall cost of a fixed basket of goods that did not allow for this type of substitution away from what becomes relatively more expensive, it tended to overstate the actual rise in the cost of living. The official CPI still assumes no substitution at the upper level of aggregation, but the Bureau of Labor Statistics now publishes an

Table 5
Wages in 1979 Dollars of the 10th Percentile Income Distribution

Deflated by:	*1979*	*1990*	*1994*	*2000*	*2005*
CPI-U	1.00	0.89	0.88	0.99	1.00
CPI-U-RS	1.00	0.90	0.91	1.04	1.05
C-CPI-U	1.00	0.92	0.94	1.09	1.12
C-CPI-U-BW	1.00	0.96	1.01	1.22	1.31

alternative price index—the "chained" price index—that does allow for some substitution at the higher level of aggregation as well.[7]

Following the Boskin Commission report (discussed in this journal in Boskin et al., 1998), a major change was introduced into the Consumer Price Index that has mitigated the substitution bias. Since 1999, geometric averaging has been used when the 80,000 prices of specific items are combined into the indexes for entry-level items. For instance, in the example of Brand C and Brand P cola drinks, a geometric average of the new prices would be $(2 \times 0.5)^{1/2} = 1$, resulting in 0 percent inflation. A fixed-weight basket assumes that individuals will purchase the same quantity of goods over time, regardless of how the price changes; in contrast, a geometric average assumes that individuals will make the same expenditure on goods over time. This assumption is equivalent to assuming a unitary price elasticity of demand; for example, if the price of a good fell 10 percent, but the consumer purchased 10 percent more of the good, then expenditure on the good would be unchanged. Using a unitary price elasticity of demand for goods within the same entry-level category is imprecise, but surely an improvement over assuming a price elasticity of zero.

Table 5 presents the evolution of real wages of the lowest tenth percentile of U.S. wage earners computed using different price indexes. As discussed a moment ago, the first row presents the conventional measure of real wages at the 10th percentile after deflating the wage series by the CPI for urban consumers (CPI-U), which made no correction for substitution bias prior to 1999. Here we can see the standard result that workers at the lower tail of the income distribution had no real wage growth over the last quarter century.

In the next row of the table, we recomputed the real wage series using what the

[7] The chained CPI uses a Tornqvist index, which is a discrete approximation of a continuous Divisia index. A Divisia index is a weighted sum of the growth rates of the index components. When a Tornqvist index is used as an approximation to the continuous Divisia index, the growth rates are defined as the log difference of successive observations of the components and the weights are equal to the mean of the shares of the components in the corresponding time periods. The Tornqvist index represents an improvement over constant base-year-weighted indexes, because as relative prices change, the Tornqvist index allows both quantities purchased to vary and the weights used in summing the inputs to vary, reflecting the relative price changes. For a more complete discussion of Tonrqvist indices, see Diewert (1976).

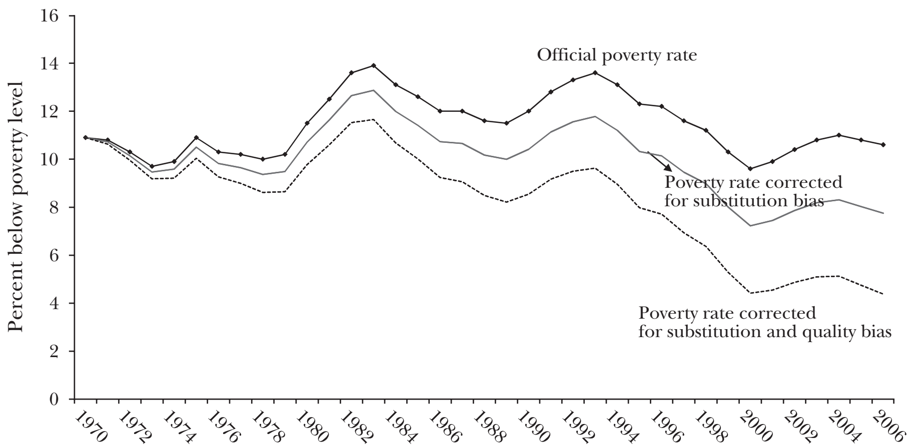
U.S. Bureau of Labor Statistics calls the CPI-U-RS (research series), which is an estimate of what the CPI for urban consumers would have been if the 1999 improvements had been implemented earlier. This calculation suggests that real wages of the lowest paid 10 percent of workers didn't stagnate, but actually rose by a modest 5 percent over this time period. This series, however, still uses the assumption of fixed baskets of goods at the upper level of aggregation between the entry-level price indexes. To be sure, there is probably less substitution at the upper level, but surely some types of substitution are plausible—like purchasing more food from grocery stores if the relative price of restaurant meals rises. The "chained" CPI-U, or C-CPI-U, allows for some substitution at the upper levels of aggregation. Using this less-biased price index, real wages at the 10th percentile would have risen by 12 percent between 1979 and 2005.

Along with substitution bias, the other major category of bias widely recognized in the CPI involves quality improvements in existing goods—including the quality improvement dramatic enough to be named as an entirely new good. With existing goods, it often occurs that quality improvements are not fully considered in the collection of price data, so that what looks in the government statistics like a higher price for a good may actually reflect a higher quality of good, or what looks like an unchanged price may actually reflect improved quality for the money. New goods will eventually be rotated into the basket of goods whose prices are included in the CPI, and any change in the price of those new goods will be tracked, but the benefits to consumers from having the new goods available in the first place are not well-captured.

Broda and Weinstein (2008) use barcode-level data to estimate an aggregate quality bias in the Consumer Price Index. They develop an adjusted CPI, which we will call here the C-CPI-U-BW, which allows some substitution between goods at both lower and higher levels of aggregation, and also adjusts for quality/new goods bias. Using this index, we compute that the real wages at the 10th percentile increased by 30 percent from 1979 to 2005. In other words, the real wages of low earners have not remained stagnant, as suggested by conventional measures, but actually have been rising on average by around 1 percent per year. Roughly half of the difference between the standard measure in the top row, showing no increase in wages for low-income workers, and our estimate in the bottom row is due to the fact that the CPI-U only imperfectly corrects for substitution bias, and the remaining half is due to the fact that many new and improved goods have appeared and the impact of these goods are only imperfectly captured by the CPI.

The effect of these adjusted measures of inflation on poverty rates in the United States is dramatic. Remember, the income thresholds for the poverty line are updated by the CPI for urban consumers each year, but otherwise have undergone only minor modifications since the 1960s. However, the CPI has an upward bias as a measure of the rise in the cost of living, both because of substitution biases (although this bias has been attenuated since 1999) and also because the index does not fully include the benefits to consumers of new and

Figure 3
Percent of Families Below Poverty Level in the United States, 1970–2006



Source: U.S. Census Bureau and authors' calculation.
Note: A constant substitution and quality bias is assumed throughout the entire period. The estimates of the substitution bias are based on the 2000–2006 period and the quality bias on the 1994–2003 period.

better goods. Thus, changes in poverty rates over time will be sensitive to biases in the CPI.

No agreed-upon estimates exist for all of the biases in the Consumer Price Index over recent decades, but some estimates are possible. As a starting point, we focused on movements in poverty between 1990 and 2005. We use the Public Use Microdata Sample (PUMS) and other Census data to compute poverty rates in 1990 for each class of household. First, we update the poverty thresholds using the CPI for urban consumers (CPI-U) between 1991 and 2005; this recreates the official poverty rates (with a slight difference because the sample of households in the PUMS differs from that on which the official poverty rate estimation is based). Next we recompute the poverty thresholds using the chained CPI (C-CPI-U) and the Broda–Weinstein adjusted index (C-CPI-U-BW). Finally, we extrapolate the differences in the last 15 years between the official and the adjusted thresholds back to 1970.

In Figure 3, we recomputed the poverty rate using poverty threshold cutoffs based on different inflation rates. The official poverty rate is based on the poverty line rising with the CPI for urban consumers (CPI-U). This series indicates that 9.6 percent of all families were in poverty in 2000, rising to 10.8 percent by 2005. This number tends to overstate poverty because prices have not been rising as fast as the CPI-U suggests and so using it to compute poverty thresholds raises the poverty thresholds too rapidly. If we use the chained CPI instead—allowing for substitution at both the lower and higher levels of aggregation—the poverty rate in 2005 is 25 percent lower than the official statistic. If we make a further adjustment to the poverty thresholds by taking into account the value of improved quality and new goods with the Broda–Weinstein (2008) estimates, poverty rates in 2005 fall to less than half the rate in the official statistics. In other words, the seeming stability of the poverty rate over recent decades arises from failing to adjust for the fact that the poor can substitute away from expensive goods and have access to new and better goods.

Over the past 15 years, the difference in methodologies for computing the rates has produced a poverty rate that rises by 1.7 percentage points per decade faster than the rate based on the actual bundle of goods consumed. If the U.S. government statistical agencies had used a quality-adjusted chained index instead of the CPI-U to adjust the poverty line, the actual poverty rate would have fallen by about 60 percent from 1970 to 2005.

Conclusion

For decades the examination of poverty rates and real wages has focused primarily on the level of wages and incomes, ignoring the crucial role that variation in prices paid and in price levels play in any comparison over time. Focusing on prices of individual goods and price indexes suggests several striking findings about poverty and low wages in the United States that are contrary to some widely held beliefs.

First, by examining the actual prices paid by households we find that poor households pay less for the goods they buy than higher-income households. Moreover, we find that poor households shop more, not less, in discount stores such as supercenters and that, even in stores of the same retail chain, poor households pay a lower price.

Second, when the standard Consumer Price Index is adjusted to take into account common estimates of the substitution and new-goods/quality biases, it turns out that real wages of low-wage U.S. workers at the 10^{th} percentile have risen substantially over the last 30 years. Similarly, there is a notable gap between the way poverty thresholds are adjusted over time and the way researchers believe standard of living should be compared over time. This gap between practice and theory is not a theoretical curiosity. Using an upgraded CPI that controls for existing substitution and new goods/quality biases in a conservative way, we find that poverty rates in 2005 would be half what the official measures suggest.

■ *We wish to thank the National Science Foundation for supporting this work under grant SES-0820536. Broda thanks the IGM initiative at the Booth school of business for support. Morgan Hardy provided excellent research assistance. Views expressed in this paper are solely those of the authors and do not reflect official USDA policy positions on these issues.*

References

Attanasio, Orazio, and ChristineFrayne. 2005. "Do the Poor Pay More?" http://www.ifs.org.uk/publications/4318.

Aguiar, Mark, and Erik Hurst. 2007. "Lifecycle Prices and Production." *American Economic Review,* 97(5): 1533–59.

Betancourt, Roger, and David Gautschi. 1993. "The Outputs of Retail Activities: Concepts, Measurement and Evidence from US Census Data." *The Review of Economics and Statistics,* 75(2): 294–301.

Boskin, Michael J, Ellen R. Dullberger, Robert J. Gordon, Zvi Grilliches, and Dale W. Jorgenson. 1998. "Consumer Prices, the Consumer Price Index, and the Cost of Living." *Journal of Economic Perspectives,* 12(1): 3–26.

Broda, Christian, and David E. Weinstein. 2008 *Prices, Poverty and Inequality: Why Americans Are Better Off Than You Think.* Washington, DC: AEI Press.

Broda, Christian, and David E. Weinstein. Forthcoming. "Product Creation and Destruction: Evidence and Price Implications." *American Economic Review.*

Caplovitz, David. 1963. *The Poor Pay More: Consumer Practices of Low-income Families.* Free Press of Glencoe.

Chung, Chanjin, and Samuel L. Myers, Jr. 1999. "Do the Poor Pay More for Food? An Analysis of Grocery Store Availability and Food Price Disparities." *Journal of Consumer Affairs,* 33(2): 276–96.

Cotterill, Ronald W., and Andrew W. Franklin. 1995 "The Urban Grocery Store Gap". Food Marketing Policy Issue Paper 8, Food Marketing Policy Center, University of Connecticut.

Diewert, Walter Erwin. 1976. "Exact and Superlative Index Numbers." *Journal of Econometrics,* 4(2): 115–45.

Einav, Liran, Ephraim Leibtag, and Aviv Nevo. 2008. *On the Accuracy of Nielsen Homescan Data.* Economic Research Report 69. Economic Research Service, U.S Department of Agriculture. http://www.ers.usda.gov/Publications/ERR69/.

Finke, Michael S., Wen S. Chern, and Jonathan J. Fox. 1997. "Do the Urban Poor Pay More for Food? Issues in Measurement." *Advancing the Consumer Interest,* 9(1): 13–17.

Hayes, Lashawan Richburg. 2000. "Do the Poor Pay More? An Empirical Investigation of Price Dispersion in Food Retailing." Industrial Relations Section Working Paper 446, Department of Economics, Princeton University.

Johnson, Lyndon. 1964. "State of the Union 1964." Speech, January 8. http://www.presidentialspeeches.org/State-of-the-Union-1964-Lyndon-Johnson.php.

Journal of Economics Perspectives. 1998. "Measuring the CPI" symposium. Vol, 12, no. 1, pp. 3–78.

Journal of Economics Perspectives. 2003. "Consumer Price Index" symposium. Vol, 17, no. 1, pp. 3–58.

Kaufman, Phillip R., James M. MacDonald, Steven M. Lutz, and David M. Smallwood. 1997. *Do the Poor Pay More for Food? Item Selection and Price Differences Affect Low-Income Household Food Costs.* Agricultural Economics Report No. AER759, U.S. Department of Agriculture, Economic Research Service.

Lebow, David E., and Jeremy B. Rudd. 2003. "Measurement Error in the Consumer Price Index: Where Do We Stand?" *Journal of Economic Literature,* 41(1): 159–201.

Leibtag, Ephraim S., and Phil R. Kaufman. 2003. "Exploring Food Purchase Behavior of Low-Income Households: How Do They Economize?" U.S. Department of Agriculture, Economic Research Service, AIB 747-07.

Schultze, Charles. 2002. *At What Price: Conceptualizing and Measuring Cost-of-Living and Price Indexes.* National Research Council.

This article has been cited by:

1. Douglas J. Besharov, Marianne Bitler, Steven J. Haider. 2011. An economic view of food deserts in the united states. *Journal of Policy Analysis and Management* **30**:1, 153-176. [CrossRef]
2. Gita Gopinath, , Pierre-Olivier Gourinchas, , Chang-Tai Hsieh, , Nicholas Li. 2011. International Prices, Costs, and Markup Differences. *American Economic Review* **101**:6, 2450-2486. [Abstract] [View PDF article] [PDF with links]
3. Martin Burda, Matthew Harding, Jerry Hausman. 2011. A Poisson mixture model of discrete choice. *Journal of Econometrics* . [CrossRef]
4. Craig Gundersen. 2010. Being Consumed: Economics and Christian Desire. *Economic Record* **86**:275, 648-649. [CrossRef]
5. Andrew Leicester, ZoÃ« Oldfield. 2009. Using Scanner Technology to Collect Expenditure Data*. *Fiscal Studies* **30**:3-4, 309-337. [CrossRef]