



Comparing Two Sources of Retail Meat Price Data

William Hahn, Janet Perry, and Leland Southard



Visit Our Website To Learn More!

Data on Meat Price Spreads

www.ers.usda.gov/data/meatpricespreads/

Data on Retail Scanner Prices for Meat

www.ers.usda.gov/data/meatscanner/

National Agricultural Library Cataloging Record:

Hahn, William

Comparing two sources of retail meat price data.

(Economic research report (United States. Dept. of Agriculture. Economic Research Service); no. 88)

1. Meat—Prices—United States—Statistics, 2. United States—Statistical services. I. Perry, Janet E.

II. Southard, Leland. III. United States. Dept. of Agriculture. Economic Research Service. IV. Title.

HD9415

Cover photo credit: Shutterstock.

Recommended citation format for this publication:

Hahn, William, Janet Perry, and Leland Southard. *Comparing Two Sources of Retail Meat Price Data*, ERR-88, U.S. Dept. of Agri., Econ. Res. Serv. November 2009.

The U.S. Department of Agriculture (USDA) prohibits discrimination in all its programs and activities on the basis of race, color, national origin, age, disability, and, where applicable, sex, marital status, familial status, parental status, religion, sexual orientation, genetic information, political beliefs, reprisal, or because all or a part of an individual's income is derived from any public assistance program. (Not all prohibited bases apply to all programs.) Persons with disabilities who require alternative means for communication of program information (Braille, large print, audiotape, etc.) should contact USDA's TARGET Center at (202) 720-2600 (voice and TDD).

To file a complaint of discrimination write to USDA, Director, Office of Civil Rights, 1400 Independence Avenue, S.W., Washington, D.C. 20250-9410 or call (800) 795-3272 (voice) or (202) 720-6382 (TDD). USDA is an equal opportunity provider and employer.



United States
Department
of Agriculture

Economic
Research
Report
Number 88

November 2009



A Report from the Economic Research Service

www.ers.usda.gov

Comparing Two Sources of Retail Meat Price Data

William Hahn

whahn@ers.usda.gov

Janet Perry

jperry@ers.usda.gov

Leland Southard*

Abstract

The livestock industry uses information on meat prices at different stages in the marketing system to make production decisions. When grocery stores began using electronic scanners to capture prices paid for meat, it was assumed that the livestock industry could capitalize on having these point-of-sale data available as a measure of the value of its products. This report compares scanner price data with publicly available data collected by the U.S. Department of Labor's Bureau of Labor Statistics (BLS). Of the two data types, scanner data provide more information about retail meat markets, including a wider variety of meat-cut prices, multiple measures of an average price, the volume of sales, and the relative importance of discounted prices. The scanner data sample, however, is not statistically drawn, and complicated processing requirements delay its release, which makes scanner data less useful than BLS data for analyzing current market conditions.

Keywords: price spreads, meat, meat pricing, scanner data, retail prices, retail meat prices, farm-to-retail

Acknowledgments

The authors would like to acknowledge Bill Cook and Mark Bowman of the U.S. Department of Labor's Bureau of Labor Statistics for reviewing early drafts of this report. We would also like to thank those in USDA's Economic Research Service (ERS) who provided technical reviews: Ephraim Leibtag, Steven Zahniser, Barry Krissoff, and Daniel Pick, as well as two anonymous outside reviewers. The authors thank John Weber and Curtia Taylor of ERS for editorial and design assistance.

*Leland Southard, now retired from ERS, was the project manager for the development of the retail meat scanner database, and his contributions to the development and documentation of the database were essential to the success of the program.

Contents

Summary iii

Introduction 1

Scanner Data to the Rescue (Or Not). 3

Comparing BLS and Scanner Prices and Price Spreads 7

Statistical Modeling With Scanner and BLS Data. 12

BLS Prices, Sales-Weighted Scanner Prices, and
the “Ideal” Measure of Retail Prices 16

Conclusion 23

References 25

Data Sources 27

Appendix – Dynamic-Adjustment, State-Space
Model: Equations, Specifications, and Tests 28

Summary

USDA's Economic Research Service has a long history of calculating meat prices at different stages in the marketing system. ERS uses data collected by the U.S. Department of Labor's Bureau of Labor Statistics (BLS) as the basis of its measure of retail meat prices. The livestock industry uses this information to make production decisions.

What Is the Issue?

Recent legislation required USDA to investigate the use of an alternative source of data on retail meat prices. The purpose of the legislation was to address livestock industry concerns regarding the quality of retail meat price data used by ERS. This report compares average retail prices calculated using data from BLS with data from grocery stores using point-of-purchase scanners to record prices. It analyzes the value of both data sets in forecasting near-term market conditions.

What Did the Study Find?

Both BLS data and scanner data have relative strengths. BLS data have several advantages over scanner data:

- BLS uses statistical sampling to select retail outlets to survey. By contrast, scanner data are volunteered by stores and may exclude certain retailers that BLS makes efforts to include.
- Because BLS uses sampling to select outlets, statistical theory implies that price averages derived from BLS data ought to be unbiased. To the extent that retailers whose price history is not captured in the scanner data set have different price structures than retailers who volunteer their data, scanner averages may or may not be biased.
- BLS data are generally available 12-20 days after the end of the month they are gathered. Because of processing issues, 7-8 weeks are required before scanner data become available.

Scanner data, in turn, have several advantages over BLS data:

- Scanner data include more meat cuts. The latest iteration of scanner data showed 188 cuts, including domestic and imported lamb. By contrast, the BLS database lists 32 meat and poultry cuts, some of which have been discontinued.
- Scanner data provide some quantity measures. BLS data provide none. Scanner data provide an index that compares a month's sales of a particular meat cut with average monthly sales of the same cut for a base year. Scanner data also show the share of meat cuts sold at a discounted price.
- Scanner data provide a wider range of price-related statistics than the BLS data. For example, scanner data are weighted by actual sales and can be used to calculate standard deviations, not just average prices. Standard deviations are a measure of the variation in prices.

In their current form, scanner data provide a different, but not necessarily better, view of retail meat markets than BLS-based data. Given the present production lag, scanner data for meat do not appear to contribute value to the analysis.

How Was the Study Conducted?

ERS routinely uses BLS data to calculate retail composites for Choice beef, pork, broilers, whole chickens, and whole frozen turkeys. This study calculated the same composites using scanner data and compared the two data sets. This study also used an econometric analysis comparing scanner and BLS prices to determine why scanner prices are more volatile than BLS prices. Both scanner data and BLS data were then examined for their value in analyzing current market conditions.

Introduction

Economics is the study of the interaction of supply and demand, with prices transmitting information to producers and consumers about where they should allocate their resources. Gathering information on prices can be costly. As a public service, the Federal Government has programs that report prices. USDA's Agricultural Marketing Service (AMS) collects and publishes livestock and wholesales meat price data. USDA's Economic Research Service (ERS) uses AMS data in its analysis of livestock markets. To supplement its analysis, ERS relies on data collected by the U.S. Department of Labor's Bureau of Labor Statistics (BLS) to calculate retail meat prices and farm-to-wholesale-to-retail price spreads.

Recent legislation required USDA to investigate the use of an alternative source of retail meat price data.¹ The purpose of this provision was to alleviate producers' concerns regarding the quality of the retail meat price and price spread data used by ERS.² This report compares the calculations of average retail prices using data from BLS and data from grocery stores using point-of-sale scanners to record prices. The report offers a statistical analysis of the relative value of the two data sets in forecasting near-term market conditions.

Why Do Grocery Store Prices Matter?

Meat producers, consumers, and policymakers make decisions based on price information from the livestock, meat, and poultry industries. AMS collects and disseminates a wide range of information on livestock and meat prices, with some prices reported twice daily. Other Federal Government agencies also collect and publish a wide range of meat and poultry quantity estimates tracing meat production and trade, providing a continuous time series for these statistics. Data for the final stages of marketing—specifically, on meat sold through foodservice and retail stores—are much less complete.

Consumers are spending an increasing share of their food dollars on away-from-home food. Still, a survey analyzed by Davis and Lin (2005) revealed that 78 percent of all pork and 65 percent of all beef were purchased in grocery stores. This finding suggests that grocery store prices are important measures of the value of meat to consumers.

To measure activity in the retail sector, ERS uses meat- and poultry-cut prices gathered by BLS (to calculate its cost-of-living indices) to calculate retail values for pork, beef, and poultry (see box, "Bureau of Labor Statistics Retail Price Data"). Hahn (2004) discusses ERS procedures for calculating retail, wholesale, and farm prices for beef and pork and their associated price spreads. The price spread is the difference between the value of the animal and its meats at different levels in the marketing system. Hahn also provides evidence that BLS-based retail prices can be useful leading indicators of farm and wholesale price changes for beef and pork. High (low) retail prices in one month are often followed by higher (lower) wholesale and farm prices in the following months.

¹ The Livestock Mandatory Reporting Act of 1999 (LMRA) required USDA to acquire an alternative measure of retail meat prices independent of the Bureau of Labor Statistics data. Part of the LMRA required ERS to continue the use of its pre-existing price-spread procedures so that the two data sources could be compared. Prior to the Act, the National Cattlemen's Beef Association had purchased and analyzed data from a grocery store data syndicate. ERS developed a pilot project to expand coverage to include additional retail cuts of beef, pork, poultry, and lamb and then began purchasing this type of data as well.

² The main purpose of the legislation was to revamp USDA's *Market News* program, which aims to aid the efficient marketing of agricultural commodities by providing the public with price and sales information drawn from transactions around the country. Fundamental changes in livestock industries had called into question the effectiveness of *Market News* reporting for livestock and led to a major redesign of the program through the *LMRA*. A recent ERS report (Perry et al., 2005) reviews developments leading up to the *LMRA* and assesses its impact on cattle markets after its implementation in 2001. GAO (2005) indicates that further work is required for the prices to reflect market transactions. The text of the legislation can be seen at: www.ers.usda.gov/data/meatscanner/livestockmandatoryreportingact.pdf

Bureau of Labor Statistics Retail Price Data

The U.S. Department of Labor's Bureau of Labor Statistics (BLS) collects data on a wide range of prices paid by U.S. consumers for various products from various outlets. BLS uses the data to calculate various versions of the Consumer Price Index (CPI). The CPI is a measure of inflation focused on the buying habits of urban consumers. On its website, BLS states that "the all urban consumers group represents about 87 percent of the total U.S. population." The items and outlets sampled by BLS are based on periodic surveys of consumer purchases. These surveys occur roughly every 10 years. BLS uses a rotating sample of randomly selected outlets. The price data that BLS collects are then used to calculate national averages that are reported each month. BLS publishes prices only when it has enough observations for a statistically reliable estimate. It contacts outlets to collect price information every business day, either via phone or through onsite visits. It surveys tens of thousands of outlets, although it does not collect all possible prices from each outlet. BLS reports monthly average retail price data on its website, <http://www.bls.gov/data/>. Data for a given month are released between the 12th and the 20th of the following month.

BLS's use of statistical sampling in selecting the outlets it surveys helps ensure the quality of the data it collects. However, BLS procedures have a well-known weakness: they are a better measure of the prices that consumers *observe* rather than the prices that consumers actually *pay*. Consumers generally buy more of an item when its price is lower. Economists have long recognized that consumer reaction to price changes complicates the BLS estimation of cost-of-living indices. Shapiro and Wilcox (1996) estimate that failing to account for consumer responses results in an average annual increase of 0.7 percent to BLS's Consumer Price Index (CPI). This may make the CPI overstate the actual inflation rate. Measuring the prices consumers actually pay would require a weighted average price—one calculated by weighting each price by the volume of sales for that price.

Thus, use of the BLS data may lead to ERS overstating its retail meat price estimates and the associated retail price spread. ERS-supported research in 1970 (Degner) calculated that switching from a simple-average price to a weighted-average price would reduce estimated retail beef prices by 7.5 percent. In 1999, the General Accounting Office (GAO, now the Government Accountability Office) estimated that ERS's pork retail value overstated the consumer price of pork in December 1998 by about 6 percent, or 14 cents per pound. Grocery store scanners provide data on prices consumers pay rather than on prices they observe. A *sales-weighted* average may provide a more accurate measure of what consumers actually pay for their meat cuts. Using point-of-sale prices and quantities sold at the various prices may result in estimates of average prices that more closely match market activity.

Scanner Data to the Rescue (Or Not)

Scanner data are now available for food products and offer new opportunities for empirical research. Stores use these data to track price and inventory movement and to control product movement along their supply chains. Scanner data allow the gathering of information at the point of sale, which could potentially lead to improved market information.

ERS purchased grocery store scanner data to compare the value of the scanner data with that of the current retail price data from BLS. The scanner data come from retail grocery stores with annual sales of \$2 million or more. The scanner data on meat sales come from national and regional chains that account for about 20 percent of the value of all U.S. supermarket sales. It is not known how much of total U.S. meat sales are covered by these stores, however.

The stores in the scanner data set share common characteristics. All are supermarkets that process their receipts using electronic scanners, sell products through the traditional supermarket retail meat case, and voluntarily provide their scanner data to commercial data firms. Processing of the data is conducted by a private firm that buys the data from a syndicate, which purchases and collects the data for all products from the stores within the syndicate. Confidentiality requirements prohibit ERS from directly accessing the raw scanner data.

Confidentiality requirements also limit ERS's knowledge of the geographic coverage of the data. Some of the chains that provide scanner data may have stores in rural areas or small towns that the BLS data do not cover. Scanner data do not include data on sales from butcher shops, warehouse clubs, mail order firms, selected big-box food distributors, and other retailers that choose not to provide data for third-party use. BLS samples a wider variety of outlets than those included in the scanner data, so BLS prices provide more comprehensive outlet coverage.

BLS uses statistical sampling to determine the outlets that are surveyed; consequently, one could use statistical theory to evaluate price averages. Conversely, the scanner data come from stores that volunteer their data. To the extent that retailers whose price history is not in the data set have different price structures for their products than the remaining stores, the average scanner price could be biased. If averages are biased, one cannot determine whether an average price constructed from the scanner data is higher or lower than the "true" average price that consumers pay.

Scanner data provide a wider range of price-related statistics than the BLS data set. The scanner data set contains both a simple average price (by stores) and a sales-weighted average price for all the meat cuts. And, because this average price is weighted by actual sales, the data set also includes the standard deviation associated with each average, a measure of the variation in prices. The advantages of the scanner data in this case, however, are offset by the lack of statistical sampling, which reduces users' confidence in the data.

The scanner data also provide measures of meat sales quantity: the share of the selected meat cut sold at a "featured," or discounted, price and an index

that measures total sales for the cut. The index compares a month's sales for the selected meat cut with average monthly sales for the base year 2001. These indices show how a meat cut's sales vary seasonally or year-to-year, but they do not allow one to compare relative volumes of meat cuts (e.g., pounds of ground beef versus pounds of bone-in chicken breast).

The use of quantity indices helps ensure the confidentiality of the stores supplying the data. Management of participating stores had expressed concerns that revealing the total pounds of cuts sold by the syndicate might allow competing stores to estimate meat sales by firm. In hindsight, one could have protected total sales-volume information by reporting the percentage of the tonnage of meat-case sales accounted for by each cut and an index for total tonnage. This type of breakout on quantities would protect stores' business interests while providing analysts and the public with more valuable information.³

Improved scanner data will not answer all the questions regarding the quantities of meat purchased because grocery stores are not the only sources of beef and pork for consumers. Restaurants and other foodservice establishments sell large amounts of meat. Further, the scanner database does not include all the meat sold through grocery stores. The scanner database includes only random-weight meat products, including whole beef and pork muscle cuts, ground beef, ham, lamb, veal, turkey, certain frozen items (e.g., patties, whole birds), and random-weight packages of bacon and sausage. The database does not include the following:

- Fixed-weight items sold in standard package sizes, such as luncheon meat,⁴ most branded sausage products, packaged bacon, or precooked items
- The meat in other processed foods, such as soups and frozen dinners
- Deli meats, including rotisserie beef, pork, and chicken

The BLS data cover a much narrower range of meat cuts than scanner data. BLS's database lists only 34 meat and poultry cuts, some of which have been discontinued. Others may be reported irregularly. By contrast, the latest iteration of scanner data shows prices for 188 cuts, including domestic and imported lamb.

Scanner Data Issues

Developing a database using scanner data for meat cuts requires complex processing that delays publishing until 7-8 weeks after the end of the month the data are gathered. AMS releases some of the data needed to calculate wholesale prices daily; all of it is available weekly. BLS data are available 12-20 days after the end of the month they are gathered. For example, January's scanner data are available in late March, and January's BLS data are available in February. To forecast January's prices at the end of December, one would use December's AMS wholesale data, November's BLS data, and October's scanner data.

Two features of scanner data account for the length of the processing time: insufficient standardization of Universal Product Codes (UPCs)⁵ and a lack

³ A tonnage index that addresses this issue is planned for the next iteration of scanner data.

⁴ Market analysis can benefit from the ability to trace the volume of each species' meat that goes into processing, but in many cases, the firms may not divulge product recipes.

⁵ Meat may be cut and packaged in the back of a store or at a central meat-cutting facility owned by the foodstore chain itself. Sources report that centralized cutting now dominates retail meat production. However, this centralized cutting may be specialized to the store's requirements.

of data on featured prices. Grocery store scanners rely on UPCs to record product types and prices based on information supplied by the store's database. Efforts are under way to standardize meat-cut descriptions and UPCs using the Uniform Retail Meat Identity Standards (URMIS) (see box, "UPC and UMRIS Codes").

The volume of scanner data generated by an individual store is huge, and the data that stores share with the syndicate are only a subset (and summary) of the available data. Each store in the syndicate shares a weekly database providing the following information for each purchase:

- UPC
- Description of the item
- Number of packages sold with that UPC
- Total dollar amount of sales of that item
- Regular price

The downloaded data do not include the weight of the packages or their actual price per pound. If the database had just one of these additional items, one could calculate the other item from the sales volume, which is expressed in dollars. Why are package weights and the actual price excluded from the database? Syndicated data were developed largely for the benefit of packaged-goods manufacturers who are not selling random-weight products like meat. In this instance, the UPC is enough to identify the weight or size of the scanned package. The average price of an item is the dollar value of sales divided by the total number of packages. Comparing the calculated average price to the regular price enables one to determine whether the item was on sale and what customers paid, on average, for the item.

The syndicate's data are further processed by a contractor hired by ERS. The contractor has access to another database, not generated by store scanners, that lists all newspaper-advertised special prices. This database enables advertised prices to be matched to individual stores. A series of programmed "exception-checking" flags automatically marks those cases where features may occur. Based on certain criteria, the advertised price replaces the regular price. Exception-flagging may compare the calculated average weight using the regular price with its past average. When items are on sale, the average package weight, based on regular prices, will decline.

The use of exception-checking could lead analysts to understate the average price. Many stores represented in this data set use shopper loyalty cards. Stores estimate that over 90 percent of shoppers use loyalty cards. Some of the meat that is assigned the sale price based on exception-checking may actually have been sold at the regular price because the consumer did not have or did not use a loyalty card. If and when this happens, the feature price may be lower than the store's true weighted-average price for the cut, and the estimated sales volume (expressed in pounds) may be too high.

The processing time for scanner data may be shortened by modifying the programs that stores use to summarize and share their data with the syndicate. For example, stores could split the fields for price, dollar volume, and

packages sold into regular and discount fields. However, because the primary users of the data are packaged-goods manufacturers, the syndicate has little incentive to modify its procedures to account for random-weight products.

The delay in processing time may make scanner data less valuable. Gropp and Kadareja (2006) found that stale public information might contribute to large spikes in price volatility. They conclude that timelier, higher quality public information results in a closer alignment of market activity because players are acting on the same information rather than incorporating various beliefs about the validity of the data. The potential for access to higher quality data may lead one to consider using scanner data for developing a weighted-average price, but the delay in processing the data could negate any added value.

UPC and URMIS Codes

The scanner data's descriptions for each grocery store are carefully matched to the corresponding Uniform Retail Meat Identity Standards (URMIS). First established in 1973 by the Industry-wide Cooperative Meat Identification Standards Committee (ICMISC), the URMIS system was developed to provide an identification system for retail meat cuts and a standardized nomenclature for every retail red meat item (beef, veal, lamb, and pork). The URMIS system was revised in 2003 and is available at <http://www.beefretail.org/uDocs/urmis/start.pdf>. Before the advent of URMIS, a specific retail cut may have had several different names depending on the store or region where it was sold. For example, Kansas City strip, New York strip, and beef loin steak are all the same cut of meat. URMIS was established to eliminate consumer confusion caused by the proliferation of names used to describe retail meat cuts. In October 1984, a guideline for assigning retail identification numbers to URMIS descriptions was approved by GS1 US (formally, the Universal Code Council), thus creating UPCs for random-weight meat products. UPCs have been developed for all URMIS codes and poultry products. Because the guidelines remain voluntary, chains, and even stores within a chain, may not consistently use URMIS codes for their random-weight meats. For this study, both URMIS codes and product descriptions from the packages were used to determine the product type so that average prices could be developed for consistent groups of meat cuts. For more information about the UPC system, see <http://www.gs1us.org/>. For specific UPC numbers and descriptions for meat and poultry, see <http://www.meat-track.com/index.php/>.

Comparing BLS and Scanner Prices and Price Spreads

This analysis presents basic findings about retail meat prices using scanner data, it compares the scanner-based price series to the BLS-based price series, and it provides statistical analysis of the market information provided by both series.⁶ This analysis also includes wholesale prices and the wholesale-to-retail price spreads implied by the different measures of retail prices. The wholesale-to-retail price spread reflects the difference between the retail price of a product and its wholesale price. ERS calculates other price spreads; however, this study examines only the wholesale-to-retail price spread, referred to here as the “price spread,” or “spread.”

Scanner data include two types of average prices. The *store-weighted average price* averages the price from each store. BLS uses a similar process to calculate its averages. The *sales-weighted average price* uses volume sold at particular prices to weight the average. As noted earlier, this may provide a more accurate measure of the average prices consumers pay for the various meat cuts. ERS publishes retail prices and price spreads for five different meats: a Choice beef composite, a pork composite, a broiler composite, whole chickens, and whole frozen turkeys. This analysis used scanner data to develop both sets of price averages (store-weighted and sales-weighted) to calculate retail prices for the above products, which were then compared with BLS-based retail prices for January 2001-August 2005.⁷

There are two reasons to expect that a sales-weighted scanner average would be lower than the store-weighted average. First, stores that featured a product at a lower price would be expected to have more sales and a larger weight in the average than stores that did not feature the product. Second, high-sales-volume stores tend to have lower prices, and high-volume stores will have a greater weight in the sales-weighted average than they would have in the store-weighted average. One might also expect that sales-weighted prices would be lower than BLS prices. It is unclear as to whether BLS prices should be higher or lower than store-weighted scanner prices.

Table 1 shows the average difference between BLS prices and store-weighted scanner prices, and between store-weighted scanner prices and sales-weighted prices during the sample period January 2001-August 2005. The store-weighted scanner prices for beef were generally lower than the BLS-based prices. For the other four meats, the store-weighted scanner prices were higher. For all five meats, the sales-weighted scanner prices averaged lower than store-weighted scanner prices. Scanner data reveal that when stores lower prices to feature a particular cut, consumers react by increasing their purchases, which lowers the average price that consumers pay for meat. The sales-weighted average prices were less than 90 percent of the store-weighted average prices, and lower prices were particularly noticeable for whole chickens at scanner-data stores. Using the sales-weight prices implies lower average price spreads. Figures 1-5 show the three average retail prices and the wholesale prices for the five meats over the study period.

Store-weighted scanner prices are expected to be higher than sales-weighted scanner prices because weighting prices by sales volume should favor the

⁶ This analysis is based on scanner data for January 2001-August 2005. The *LMRA* lapsed briefly in 2004, but President George W. Bush signed legislation (P.L. 108-444) extending the program through September 30, 2005. The *LMRA* was renewed in late 2006, extending the law until September 30, 2011 (P.L. 109-296). Because of these lapses and other budgetary issues, data after 2005 became available after ERS finished the analysis in this report.

⁷ The database is documented at www.ers.usda.gov/data/meatscanner/. The sales-weighted price corresponds to the price labeled “weighted average, feature-weighted” in the online database. The store-weighted price here corresponds to the “mean feature-weighted price” in the database.

Table 1

Comparison of retail prices, January 2001-August 2005

Product	Average price differences	
	BLS-based price minus store-weighted scanner price	Store-weighted scanner price minus sales-weighted scanner price
	<i>Cents per pound</i>	
Choice beef composite	10.5	40.8
Pork composite	-26.3	48.9
Whole chickens	-41.3	53.4
Whole, frozen turkeys	-11.1	16.2
Broiler composite	-10.6	44.4

Source: USDA, Economic Research Service using electronic scanner data for meat. For summary statistics, see www.retail-lmic.info/CD/default.asp?ERSTab=2. BLS-based price composites are available at www.ers.usda.gov/data/meatpricespreads/.

lower prices. The only exception is for turkey⁸ (see fig. 4), but only for a few months in the sample. When the sales-weighted average was above the store-weighted average, there were a large number of stores with low prices and sales and a few stores with high prices and sales.⁹ High volume and high price is unusual as high-volume outlets generally have lower prices (Leibtag, 2005).

For the Choice beef composite the BLS-based price and the store-weighted scanner price were remarkably similar for the first 49 months, January 2000-January 2004 (see fig. 1). In February 2004, however, the two prices diverged, with the store-weighted scanner price staying below the BLS-based price for the remainder of the sample period.

The divergence may stem from the December 2003 reports of the first confirmed U.S. case of bovine spongiform encephalopathy (BSE, also known as mad cow disease) in a dairy cow. An immediate reaction was that all U.S. beef export markets were closed. Some consumers in the United States chose to stop eating beef—at least for a time. All retail prices for meat (not just beef) for December 2003 were lower than prices for November, and the price decline in January 2004 was even larger. Beef prices were reduced to encourage domestic consumption of beef already on the market because this product could not be exported. In February 2004, the decline in the scanner average prices was much larger than the decline in the BLS-based prices. Starting in February 2004, the store-weighted scanner prices dropped below the BLS-based prices and maintained that position. Based on analysis of ACNielsen Homescan data, Kuchler and Tegene (2006) found that deviations from established purchase patterns following the BSE announcements varied across beef products but were limited to no more than 2 weeks in all cases. Pritchett et al. (2007) analyzed the scanner data's price and quantity data and found a similarly small and transient effect of BSE on consumer demand for meat.

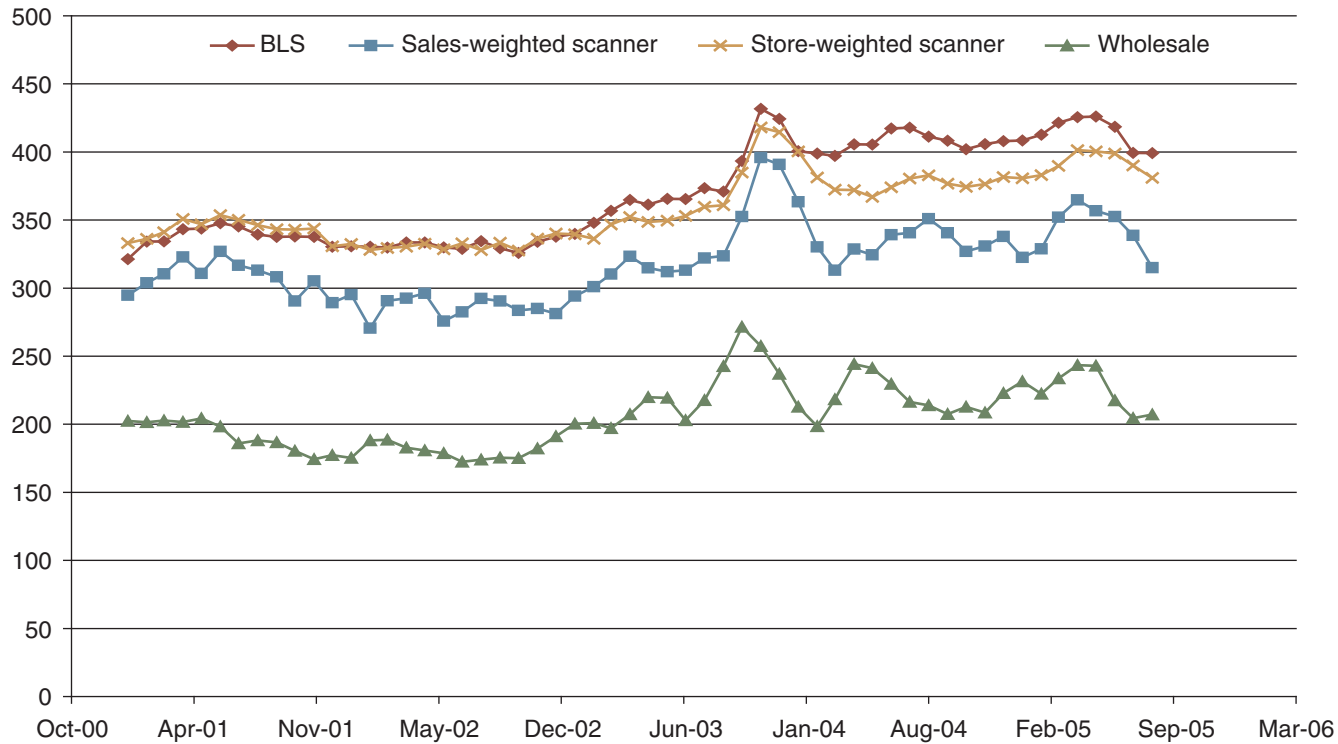
Could the apparent break in Choice beef prices be evidence that stores in the scanner data's sample reacted differently to BSE than the stores in BLS's

⁸ A special effort is made to include turkeys that stores give away as seasonal incentives in exchange for the purchase of a specified dollar amount of other groceries. These turkeys are given a price of \$0 in the scanner data, whereas BLS assigns the regular price to those turkeys.

⁹ For example, suppose that one were averaging prices from three stores—two small stores and one large store. The large store could have twice the sales volume of the two small stores combined. The large store's volume would give it twice the weight of the other two combined in the sales-weighted case, but all three stores would have the same weight in the store-weighted case. If the large store has lower prices than the two small stores, the sales-weighted price will be lower than the store-weighted price. If the large store has higher prices, the store-weighted price will be lower than the sales-weighted price.

Figure 1
Choice beef composite prices

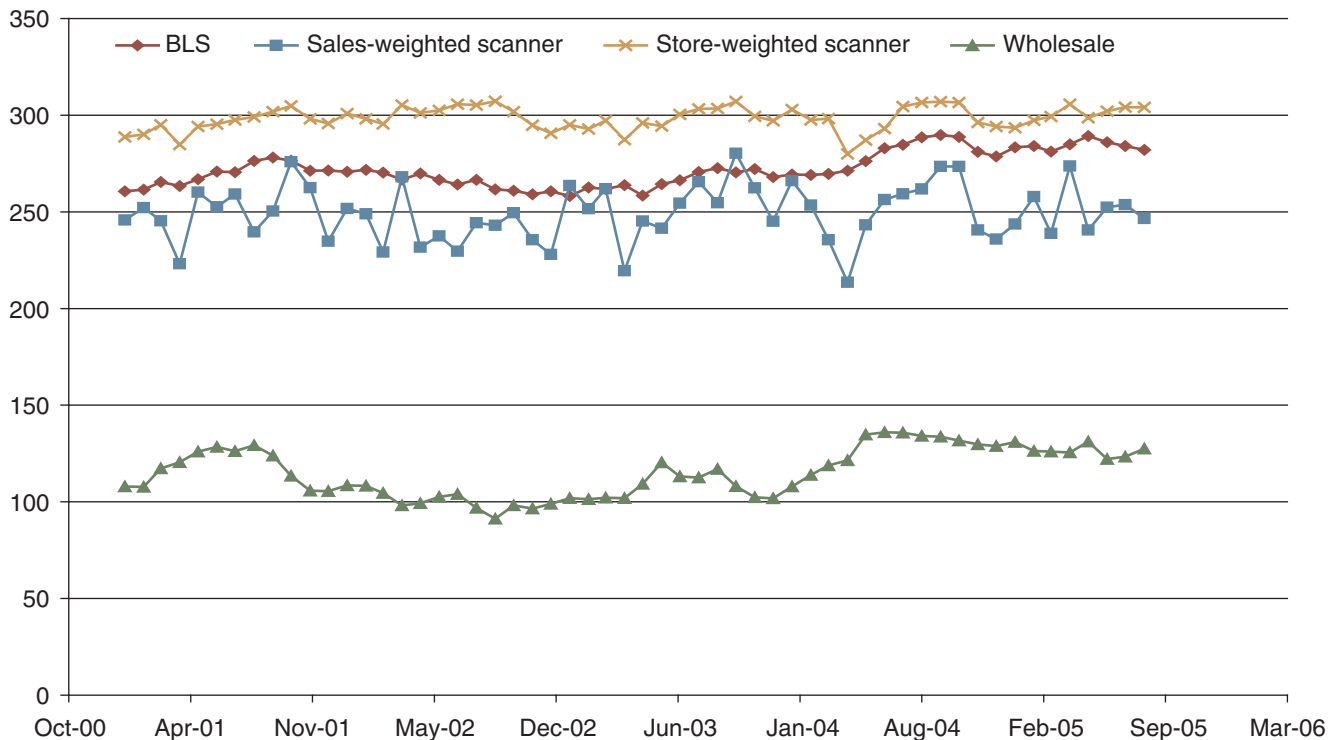
Cents per pound



Source: USDA, Economic Research Service calculations.

Figure 2
Pork composite prices

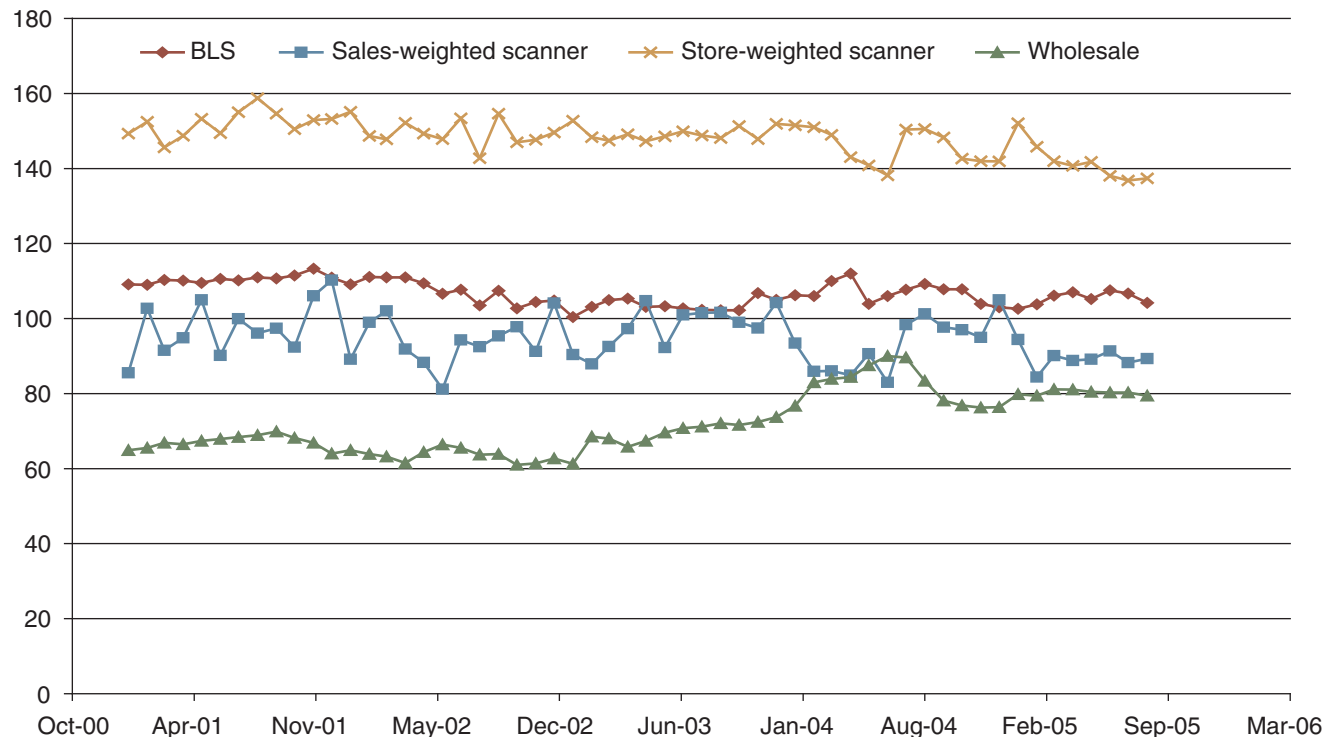
Cents per pound



Source: USDA, Economic Research Service calculations.

Figure 3
Whole chicken prices

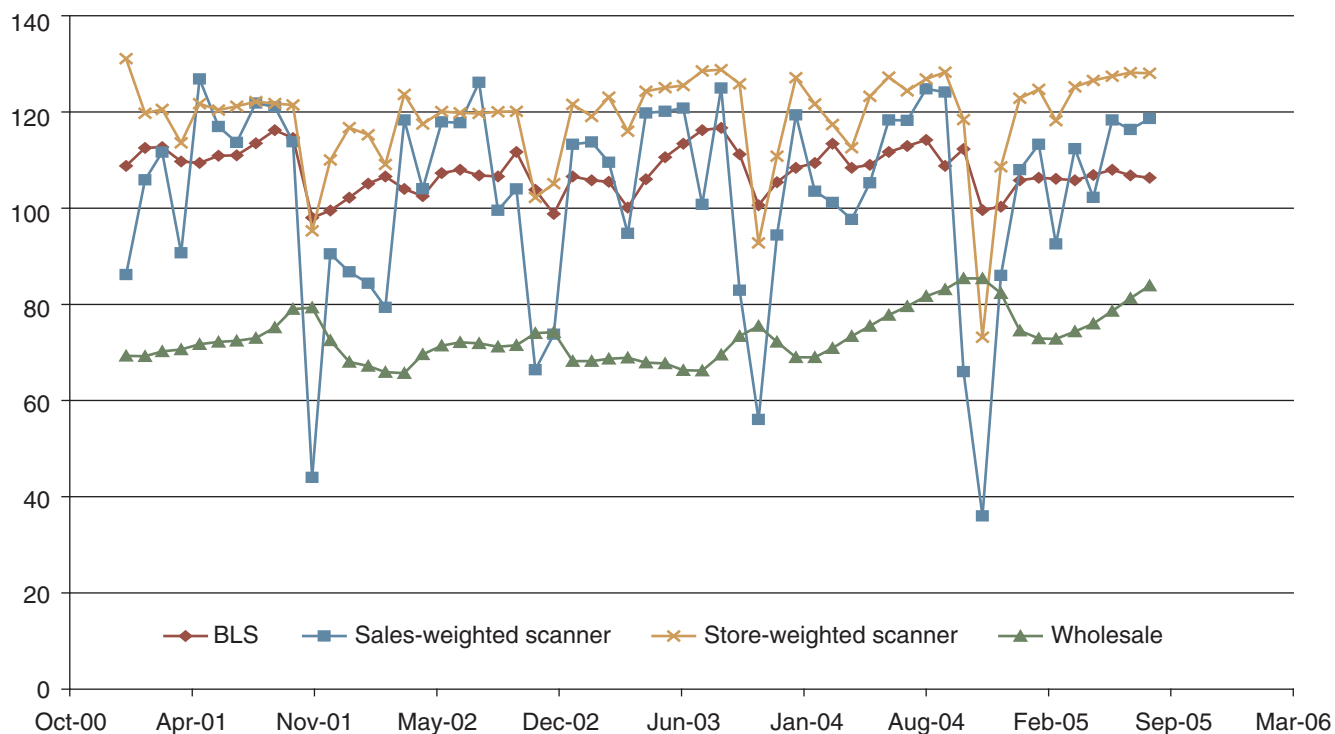
Cents per pound



Source: USDA, Economic Research Service calculations.

Figure 4
Whole turkey prices

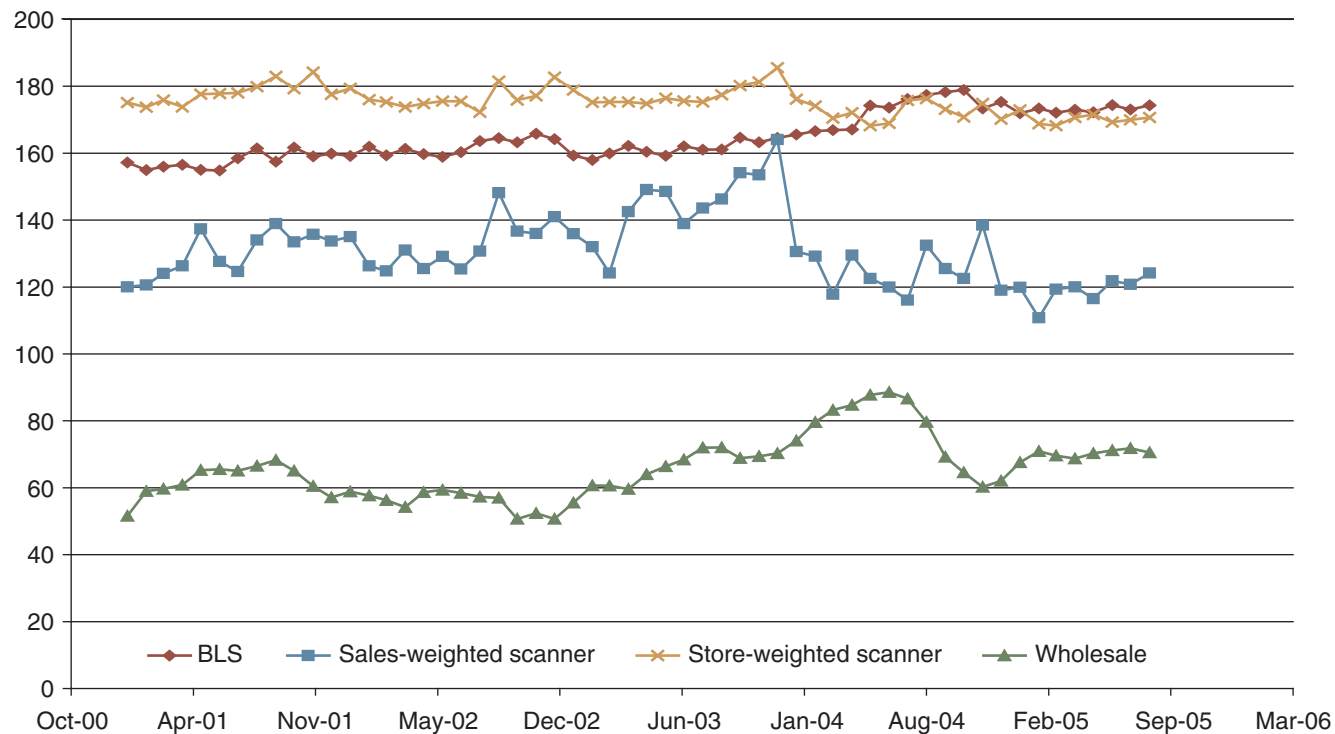
Cents per pound



Source: USDA, Economic Research Service calculations.

Figure 5
Broiler composite prices

Cents per pound



Source: USDA, Economic Research Service calculations.

sample? Consider the price data for pork, chicken, and broilers (see figs. 2, 3, and 5). For each of these meats, both scanner-based prices dropped during the first part of 2004, unlike BLS-based prices. The downward trend was most notable for the sales-weighted prices. For broilers, scanner prices remained lower than BLS-based prices after the first quarter of 2004, whereas pork and chicken prices returned to their previous levels.

For pork and chicken, the store-weighted scanner price was consistently above the BLS-based price, whereas the sales-weighted price was consistently below the BLS-based price. The store-weighted price for broilers started out above the BLS price, but after the first quarter of 2004, it began to track the BLS-based price.

Throughout the sample period, both the store-weighted and sales-weighted scanner prices for beef were significantly lower than the BLS-based prices, while scanner prices for the other meats were significantly higher than BLS-based prices. This pattern demonstrates that the stores providing scanner data are somehow different from those in BLS's sample. Locations of the scanner-data stores skew toward medium-sized cities. By contrast, BLS's sample is concentrated in metropolitan areas. Stores vary in their approaches to pricing and sales; scanner-data stores tend to have high regular prices for most meats, at least when compared with stores in BLS's sample. It is possible that the beef-pricing strategies of scanner-data stores changed in 2004, relative to those in BLS's sample.

Statistical Modeling With Scanner and BLS Data

As part of its analysis, ERS developed several questions about the data series being compared:

- Which price series best represents the national average retail price?
- What contributes to the differences in the month-to-month variability of the price series?
- In light of the expense, complex processing requirements, and reporting lag times, do scanner data sufficiently increase the understanding of meat markets?

To address these questions, ERS developed and tested a series of statistical models and applied these models to each of the five meats. ERS used the same basic statistical structures for all five of the meats. This structure attempts to account for the interaction between a meat's wholesale price, its BLS price, and its sales-weighted scanner (SWS) price. The technical details on the statistical structures are outlined in the appendix of this report. All five sets of prices were analyzed separately.

ERS used a combination of “dynamic-adjustment modeling” and “state-space” techniques to analyze the differences in the two data sources for meat prices. Dynamic-adjustment models were used to analyze Choice beef and pork composites (see Hahn, 2004). Dynamic adjustment implies that it takes time for the market to react to changes. One of the implications of dynamic adjustment is that the current month's prices may be affected by the previous month's prices.

State-space is a specific type of statistical estimation that is appropriate in this analysis, particularly for dynamic adjustment models that have missing or unobserved data (Durbin and Koopman, 2001). (The missing and unobserved variables are called state variables.) State-space uses the variables one sees to estimate the unseen variables. The word *estimate* is important here; these estimates are likely to be inaccurate. Part of state space is measuring the accuracy of the estimates. These estimated states are used to predict the next period's state and observed variables. Once one observes the next period's information (whatever it is), one can improve the estimates of the state variables and make the forecast for the following month. Generally speaking, as one gains more information, estimates of the state variables improve and forecasts are more accurate. The ability of state-space techniques to handle missing data is helpful in this analysis. ERS had a limited amount of scanner data at the start of the analysis. SWS data ran from January 2001 to August 2005, or 56 months. BLS and WHL data pre-date SWS data and continued after the end of the scanner data. SWS data are missing prior to January 2001 and (initially) after August 2005. If SWS data are helpful for explaining meat-market conditions, state-space techniques allow the analyst to account for SWS data effects on the BLS and WHL data when SWS is not available. State space also allows one to measure the degree to which observing SWS will improve the accuracy of forecasts and analysis.

Which price series best represents the national average retail price? The rationale behind the legislation requiring the publication of scanner prices was that BLS prices did not adequately reflect meat prices. Without correct pricing signals, producers may make incorrect production decisions, so it was assumed that having prices that more accurately reflect consumers' preferences would be of value. ERS tested whether BLS or SWS prices best measure the retail price. One of the unobserved variables in the state-space analysis is the best measure of the retail price (BMRP). Note that BMRP does not include the term *average*. BLS prices are based on simple-averages. SWS prices are sales-weighted averages. It may well be the case that neither simple averages nor sales-weighted averages are ideal. One extreme possibility is that there is a single firm that the other firms follow—that firm's price could be the BMRP.

Neither BLS nor SWS prices by themselves cover the entire at-home meat market. Stores that are outside either the BLS or the SWS sample may have an effect on the market. The average price from these excluded outlets is called some other price (SOP). SOP is another unobserved variable. SOP could also include the influences of foodservice and export markets. The state-space model makes BMRP a weighted average of the BLS, SWS, and SOP. Part of the estimation process is finding the weights for BLS, SWS, and SOP prices that give the most accurate forecasts for WHL, BLS, and SWS prices. One can test to determine if any of these weights is 1; that would make its price the BMRP. If BLS is the BMRP, that would mean that retail meat pricing is determined by sales in urban markets. If SWS is the BMRP, then grocery stores with scanners drive retail pricing. If SOP is BMRP, then neither BLS nor SWS are driving meat market pricing. However, data on BLS and SWS prices will still be valuable for analysts in this case as these two prices can provide information with regards to SOP/BMRP.

What contributes to the differences in the month-to-month variability of the price series? Sales-weighted scanner prices show more variation month to month than BLS prices. The statistical model can estimate and test two sources of this additional volatility. The first source of a price-volatility differential is variation in adjustment speeds.

Market conditions change daily. Economists have observed that short-term reactions to market conditions usually differ from long-term reactions. For example, cutting beef production by 5 percent for a month has a different effect on prices than cutting beef production by 5 percent for a year. Economists have developed the dynamic-adjustment model to examine this type of phenomenon.

The dynamic-adjustment model for this analysis estimates the long-term, or "full-adjustment," effects of changes in economic conditions on the five prices in the model: WHL, BLS, SWS, (the three that are observed sometimes) SOP, and BMRP. The full-adjustment effects are often called target values. The current month's prices are a function of the current month's targets and the previous month's prices. If the current month's price is somewhere between the previous month's price and the target, the effect is partial adjustment. The current month's price may equal its target (full adjustment) or it may overadjust. ERS has used dynamic-adjustment models to measure adjustment patterns in the Choice beef and pork composites (see

Hahn, 2004). This research revealed that price adjustment was dynamic, and mostly “partial.” When there is partial adjustment, the target prices vary more month-to-month than do the actual prices. As prices adjust more quickly, volatility increases. Hahn’s research implies that wholesale Choice beef and pork values adjust more quickly than (BLS-based) retail prices. Consequently, wholesale values are more volatile than BLS-retail values.

The higher volatility of scanner data may mean that the scanner prices are rapidly adjusting to changes in the market. For technical reasons, ERS tested faster adjustment speed by SWS in two phases. In the first phase, the adjustment speeds for SWS and BLS were checked to determine if they were the same. If the speeds were not the same, ERS used model simulations to determine which of the two retail prices adjusted faster.

The second source of a differential in price volatility is referred to as a transient effect. The partial-adjustment model uses the current month’s prices to forecast the next month’s prices. A transient effect is that part of the current month’s price change that does not help forecast the next month’s price. Transient effects are purely random. The average or expected transient effect for a month will be 0; the “size” of a transient effect is measured by its standard deviation. The standard deviation is a statistical measure that is never negative. Larger transient-effect standard errors produce more volatile prices. Model estimates were tested to determine if the standard-deviations of the BLS and SWS prices’ transient effect were the same. (ERS also tested model estimates to determine if prices have transient effects at all.) If the difference between the transient effects was statistically significant, the estimates were compared to determine which of the two was larger.

There are two potential sources for transient effects; either or both could be present in the data. The first is that transient effects could be a fundamental part of the pricing processes. The second is that the transient effects are actually the result of statistical problems. Both data sources are based on samples. BLS does not sample all prices in all stores every day. The scanner data do not include all supermarkets with scanners. Statistical theory holds that the averages one derives from a sample will differ from the average based on the entire population and that larger, randomly selected samples tend to increase the accuracy of sample averages as estimates of the true average. Statistical theory also requires that the samples be independently drawn. Many individual stores are owned by large, multi-outlet firms, and pricing within such firms is likely to be coordinated. According to the U.S. Census Bureau (2005), the top four grocery firms in 2002 controlled 31 percent of U.S. grocery sales. The number of independent prices (associated with separate firms) in both the scanner and BLS-based data sets was much smaller than the number of individual stores, raising the potential for sampling error. Finally, it also may be the case that the best possible measure of a retail price would use neither a simple average (BLS) nor a sales-weighted average (SWS).

Is it possible to determine the causes of transient effects? The raw data for WHL, BLS, and SWS prices come from different sources using different data-collection methods. If the transient effects were purely the result of measurement errors, one might expect these errors to be independent of one another—their correlations should be zero. True transient effects seem more

likely to be correlated. ERS tested the correlations of the transient effects to determine if they differ from zero. Correlated transient effects would seem to rule out pure measurement errors.

Separating the transient changes from the nontransient changes is difficult, as transient effects are another set of unobserved variables. The statistical model requires state variables for WHL and BLS prices to account for these transient effects. The WHL and BLS state variables are not actually the transient effects; they are the prices minus the transient effects. The SWS state variable is also basically the price minus its transient effect. These observed-price states are those parts of the prices that are useful for predicting future price movements. If a price has transient effects, then its state can only be estimated. Even if the price is observed, the state will be measured with error.

Do scanner data sufficiently increase the understanding of meat markets?

This analysis formalizes the process of forecasting by using state-space techniques to measure the degree to which SWS data may improve analysis and near-term forecasting of meat-market conditions. ERS used the current month's prices to forecast the following month's prices and state variables. If SWS are not available, one has to estimate their value and apply that estimate to the forecast, which affects the accuracy of the forecast. If SWS data are available, their use will improve the estimates of the other state variables, BMRP and SOP, further improving the forecasts. The degree to which SWS data improve forecasts is a measure of their value. Part of the estimation of state-space models is the calculation of the accuracy of the forecasts and state-variable estimates under different scenarios.

BLS Prices, Sales-Weighted Scanner Prices, and the “Ideal” Measure of Retail Prices

Table 2 provides estimates of the averaging parameters. These parameters transform the BLS, SWS, and SOP into an estimate of the BMRP. If a price’s averaging parameter is one, that price is the BMRP. Neither the BLS price nor the SWS price is the BMRP in any of the five meats. In three of the five meats—beef, pork, and whole chickens—the weight for SOP is exactly one. An SOP weight equal to one implies that the BLS and SWS prices are reacting to WHL and SOP; SOP and WHL do not react to BLS and SWS. The primary value of BLS and/or SWS in market analysis in this case is their value in measuring SOP. The weight for Turkey for SOP is close to one; however, the small weight for the BLS price is statistically significant. Broiler is the one meat for which SOP has a weight of zero. For the broiler composite, the BMRP is an average of the SWS and BLS prices. The weight for broilers for SWS is close to one; again, the small weight on the BLS price is statistically significant. For the broiler composite, a combination of the SWS and BLS prices appears to make the BMRP.

Table 2

Estimates of parameters that transform retail prices into the best measure of the retail price

Product	Price or state		
	BLS-based retail price	Sales-weighted-scanner retail price	Some other retail price
	<i>Averaging parameter</i>		
Choice beef composite	0	0	1.0000
Pork composite	0	0	1.0000
Whole chickens	0	0	1.0000
Whole, frozen turkeys	0.0208	0	0.9792
Broiler composite	0.0353	0.9647	0

Source: USDA, Economic Research Service using state-space model estimates.

Volatility of Sales-Weighted Scanner Prices

Two factors may account for the higher volatility of SWS: faster price adjustment and larger transient effects. Table 3 shows estimated standard deviations for the transient effects for each meat product. Three of the five meats show transient effects for only two of the prices. Two of the meats show transient effects for all three prices. All five meats show transient effects for SWS. These SWS transient effects are considerably larger than those of WHL or BLS. Those transient effects that are not estimated to be zero are statistically significant. Also, the BLS and SWS transient effects are statistically different for all five meats. The SWS transient standard deviation is the largest of the three standard deviations for all meats. In fact, the smallest SWS standard deviation, broiler composite, is larger than the largest standard for any other meat or price, Choice beef’s wholesale price. The large transient effects account for the SWS prices being more volatile than the BLS or WHL prices.

Table 3

Estimated standard errors of the transient effects on meat prices

Product	Wholesale price	BLS-based retail price	Sales-weighted, scanner-based price
<i>Cents per pound</i>			
Choice beef composite	1.12	*	8.86
Pork composite	0.39	0.18	10.45
Whole chickens	*	0.19	3.94
Whole, frozen turkeys	*	0.35	7.13
Broiler composite	0.01	0.34	2.59

*Denotes estimates that hit the lower bound of zero.

Source: USDA, Economic Research Service using state-space model estimates.

Table 4 shows correlation estimates. Correlations come in pairs and cannot be calculated if one or both members of the pair have a standard deviation of zero. Consequently, three of the prices have only one correlation while two prices have a full set. The hypotheses that the meats' transient effects are all uncorrelated can be rejected. As mentioned earlier, measurement errors are a source of transient effects. Pure measurement errors are likely to make the transient effects uncorrelated. The significant correlations are evidence of there being no pure measurement errors. Also, evidence of the fact that the BLS price for Choice beef and the WHL prices for chickens and turkeys have zero transient effects shows that these variables are not measured with error.

Transient effects are a source of SWS volatility. Differences in adjustment speeds of prices are another potential source. All five meats were tested using model simulations to determine if BLS and SWS adjust at the same rate. This hypothesis was accepted for pork and rejected for the other meats. To determine how BLS and SWS prices differ in their adjustment speeds, the various models' estimates were simulated. The model simulations made the target prices for WHL, BLS, SWS, SOP, and BMRP 1 cent per pound higher than the starting prices. The target prices were fixed for 12 months, and the models were simulated without transient or other random effects. In the real world, the target values vary monthly and there are other unpredictable effects on prices. A month's simulated price is a function of the target and the previous month's simulated price. For ease of comparisons, all five variables were started at zero. The closer the price is to the target of one, the more fully adjusted it is.

Figure 6 illustrates the price adjustments for the Choice beef composite. The wholesale price initially over-reacts to changes in the target, whereas the BLS-based retail price does not react at all. The scanner-based retail price reacts slightly at first. The wholesale price then over-corrects in the second month, whereas the two retail prices adjust slowly. For the Choice beef composite, the simulated scanner price is generally closer to its full-adjustment value than the BLS-based price, indicating quicker adjustment.

Figure 7 illustrates the price adjustments for the pork composite category. In this case, adjustments are the same for the two observed retail prices. The wholesale price is always closer to its full-adjustment value than either of the retail prices.

Table 4

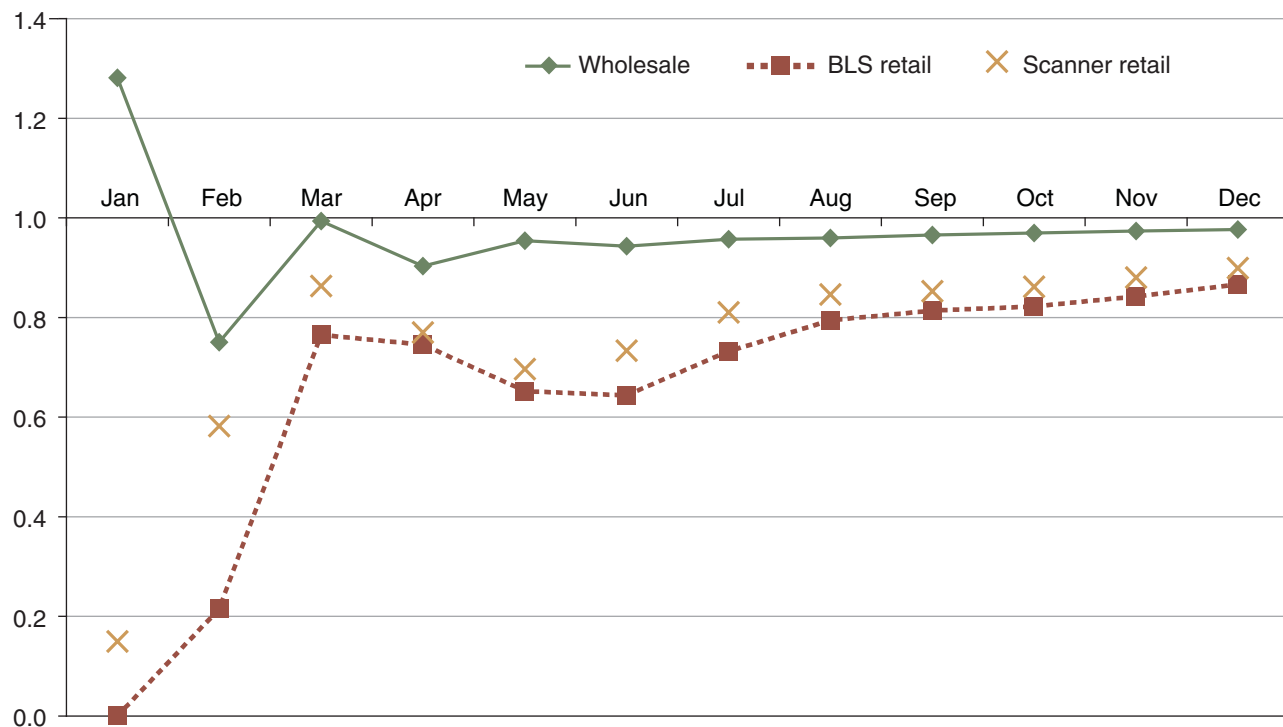
Estimated correlations of transient effects

Product	Wholesale price and BLS-based retail price	Wholesale price and sales-weighted, scanner-based price	BLS-based retail price and sales-weighted, scanner-based price
<i>Correlation estimates</i>			
Choice beef composite	NA	-0.11	NA
Pork composite	0.63	-0.50	0.36
Whole chickens	NA	NA	-0.23
Whole, frozen turkeys	NA	NA	0.20
Broiler composite	-0.64	1.00	-0.59

NA=not available, cannot be calculated as one or both of the standard errors of the price is 0.
 Source: USDA, Economic Research Service using state-space model estimates.

Figure 6
Twelve months of price adjustment for the Choice beef composite

Cents per pound

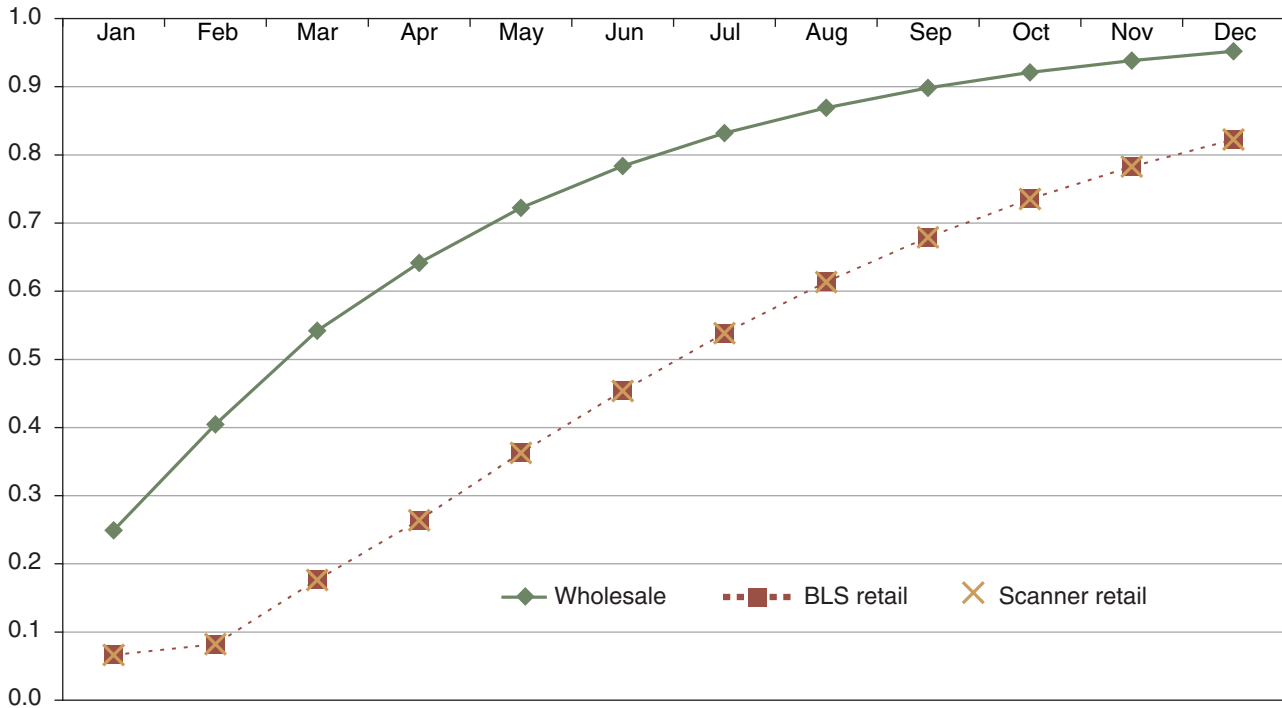


Source: USDA, Economic Research Service using state-space model estimates.

Figure 8 shows the price adjustment for whole chickens. For this product, the BLS-based retail price adjusts more rapidly than the scanner-based retail price, whereas the wholesale price adjusts more rapidly than either retail price.

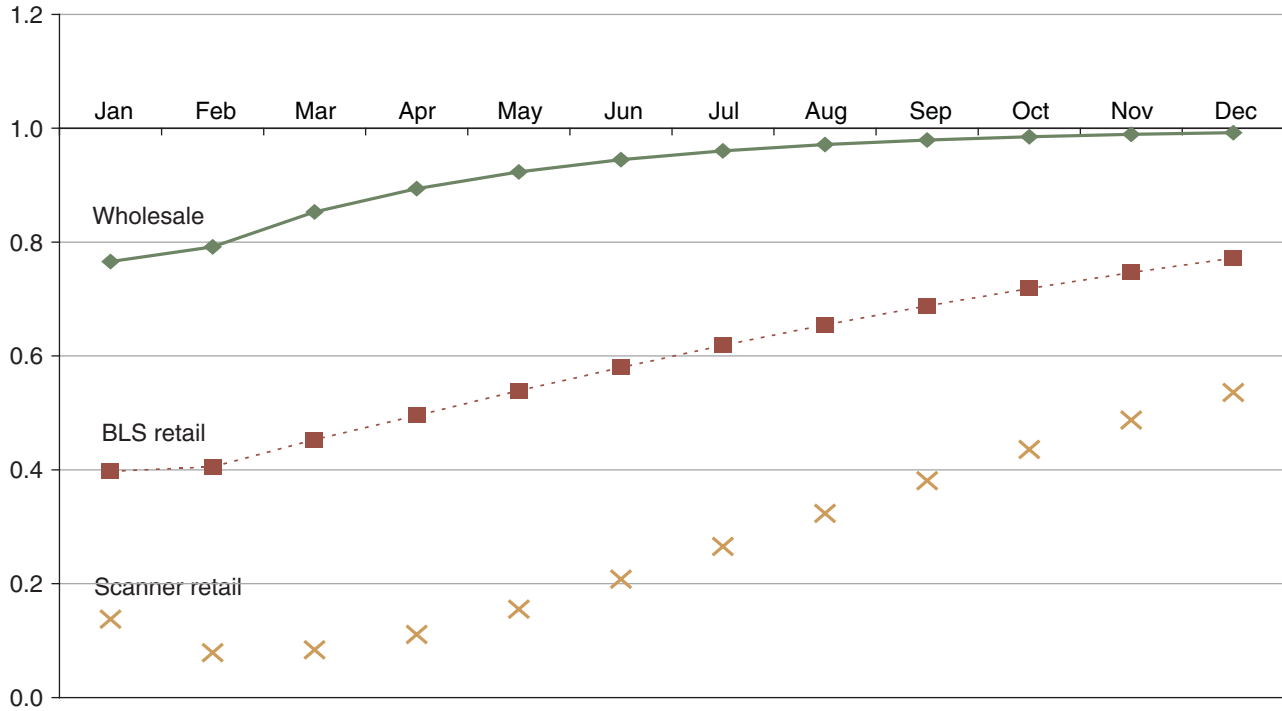
Frozen turkey is the only meat of the five that has a strong cyclical adjustment pattern. As shown in figure 9, it appears that the BLS-based retail price adjusts faster than the scanner-based retail price, which, in turn, adjusts faster than the wholesale price.

Figure 7
Twelve months of price adjustment for the pork composite
 Cents per pound



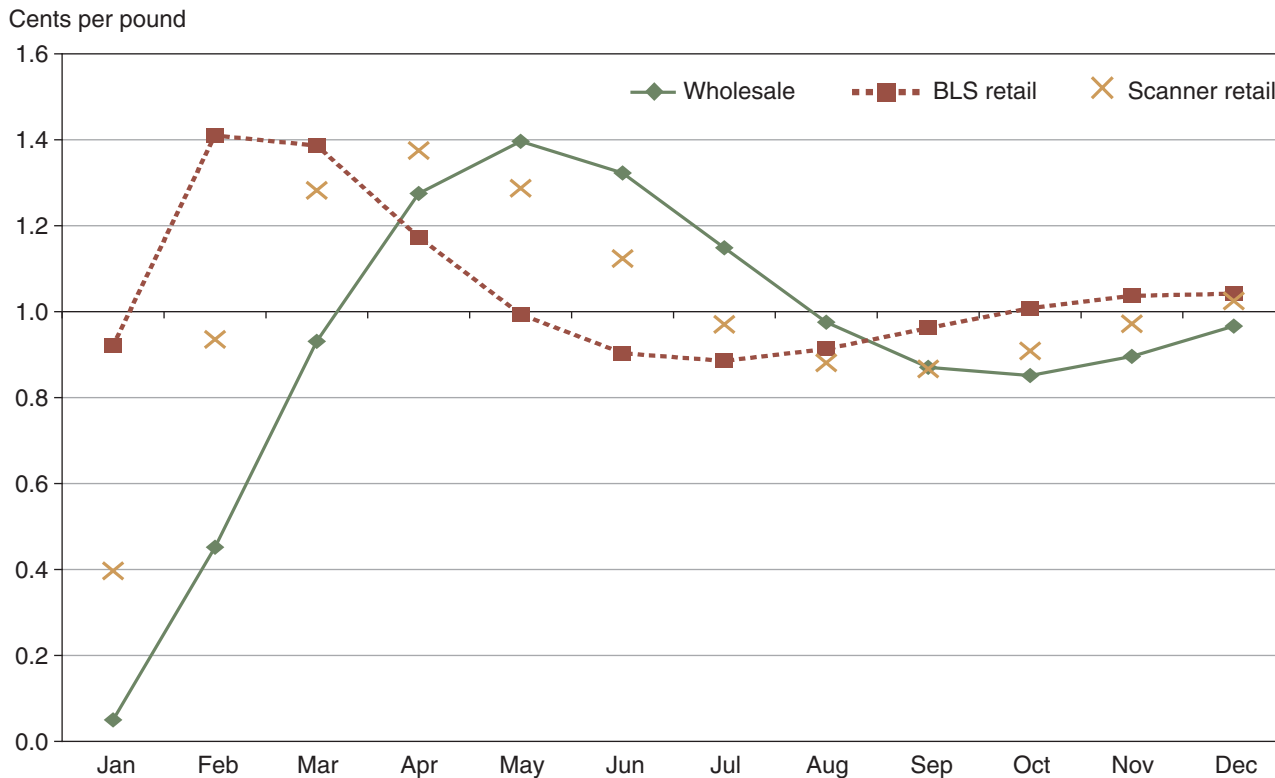
Source: USDA, Economic Research Service using state-space model estimates.

Figure 8
Twelve months of price adjustment for whole chickens
 Cents per pound



Source: USDA, Economic Research Service using state-space model estimates.

Figure 9
Twelve months of price adjustment for whole, frozen turkeys



Source: USDA, Economic Research Service using state-space model estimates.

For the broiler composite, the BLS-based retail price initially shows almost complete adjustment, while the scanner-based retail price and the wholesale price barely move (fig. 10). However, feedback from the wholesale price and the scanner-based price pulls the BLS-based price away from full adjustment. The scanner-based price and the wholesale price show similar adjustments. The BLS-based retail price is closer to its target level than either the scanner-based retail price or the wholesale price.

The Choice beef composite is the only product for which the adjustment of the scanner-based retail price is statistically different from and faster than the adjustment of the BLS-based retail price. The difference between the adjustments of BLS and SWS prices was only large in the second month of adjustment. It appears that most of the volatility of SWS prices is due to their larger transient effects.

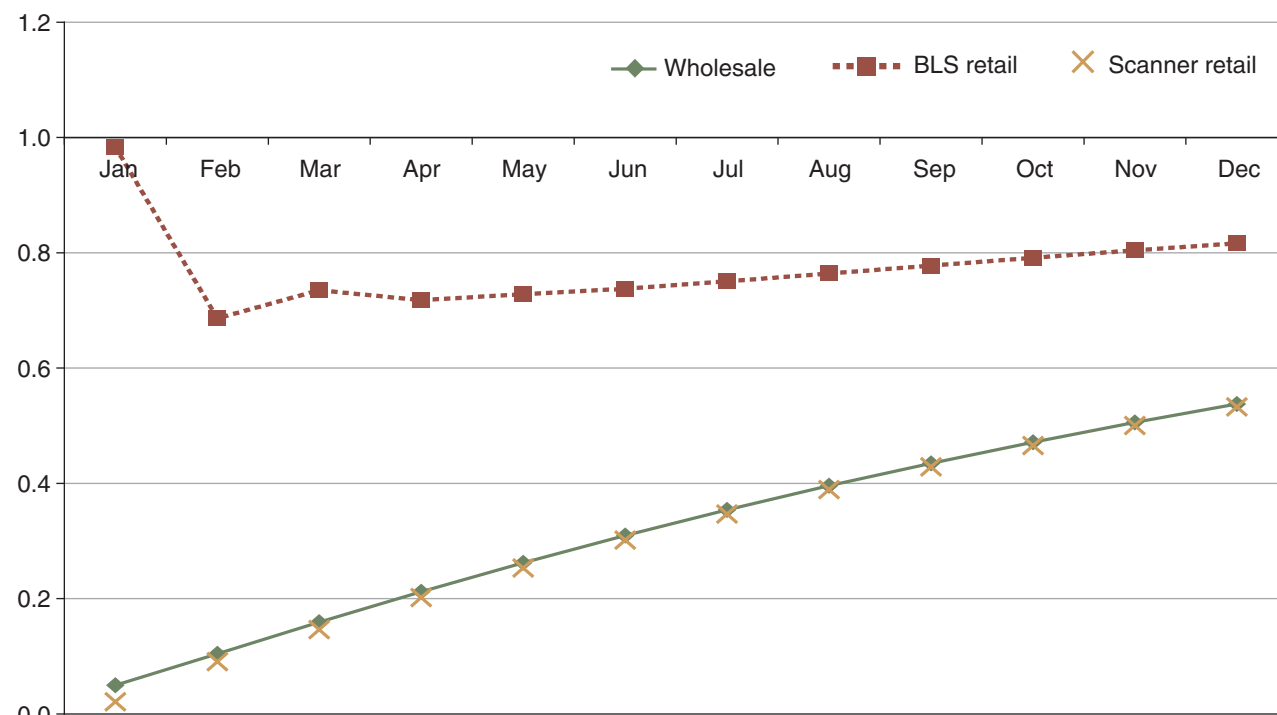
Using Scanner Data To Analyze Current Market Conditions

One result from the state-space models was a state value for the BMRP. Both BMRP and SOP have to be estimated; however, one cannot estimate the transient parts of either. The state variables for BMRP and SOP are those prices that are useful for forecasting future conditions. Table 5 shows the steady-state value of the standard error of the BMRP state’s measurement given four different patterns of data availability. The lower the standard error, the more accurate the measurement.

Figure 10

Twelve months of price adjustment for the broiler composite

Cents per pound



Source: USDA, Economic Research Service using state-space model estimates.

Table 5

Standard errors for predicting the value of the best measure of the retail price (BMRP) state

Observed prices	Timing	Product				
		Choice beef composite	Pork composite	Whole chickens	Whole, frozen turkeys	Broiler composite
<i>Cents per pound</i>						
WHL and BLS	All available at end of month	0.819	1.653	1.918	0.907	4.182
	Actual release pattern	3.729	1.783	1.999	1.087	4.246
WHL, BLS, and SWS	All available at end of month	0.809	1.483	1.594	0.884	1.488
	Actual release pattern	3.728	1.639	1.951	1.069	3.895
<i>Percent improvement adding scanner data</i>						
Prices all available at end of month		0.58	5.45	9.21	1.29	47.51
Actual release pattern		0.01	4.21	1.21	0.84	4.31
<i>Percent improvement for making release dates more timely</i>						
WHL and BLS		63.99	3.79	2.07	9.02	0.76
WHL, BLS, and SWS		64.32	5.02	10.06	9.47	44.71

Note: WHL is wholesale price, BLS represents prices from the Bureau of Labor Statistics, and SWS is the series of prices from sales-weighted scanner data.

Source: USDA, Economic Research Service using state-space model estimates.

Does the timing of data releases affect the value of the estimates? As noted earlier, newer data provide the most value, so the delay in the release of BLS and SWS data reduces their value in near-term forecasting—the state-space method allows one to measure how changing the timing of SWS data affects their value. The four patterns of data availability come from combining the two sets of observed prices with the two data release schedules—the actual release pattern, and a second estimate of the contribution of having access to all three types of data at the end of the month. The cases in which all the data are available at the end of the month are also relevant to historical analysis, when all the data for a specific time period are available. Adding scanner data to the mix (that is, having all three types of data available at the end of the month) should improve the accuracy of the models, reducing their standard errors.

The pattern of data release is especially important for analysis of current market conditions and forecasts. Wholesale prices are released daily, and by month's end, the entire month's set of daily wholesale prices is available. By contrast, BLS's data are released 2-3 weeks after the end of the month, and scanner data are released 7-8 weeks after the end of the month. Thus, at month's end, monthly wholesale price data are current, but the latest monthly observation of the BLS-based retail prices is 1 month old, and the latest monthly observation of scanner-based retail prices is 2 months old. Not surprisingly, the results reported in table 5 indicate that having all three prices available at the end of the month would provide more accurate measurements than the current pattern of data release.

Because of this timing issue, the scanner data contribute little to the price analysis for four of the five meat products. The broiler composite is the exception; scanner data greatly improve the estimates of the national average retail price, but only in the analysis of historical data. The 7-8 week lag in the availability of scanner data makes this data “stale” and of little value to the analysis of current market conditions in this context.

Conclusion

Scanner data have a number of advantages over BLS data. Scanner data cover a wider selection of meat cuts and animal species, and they provide information that is not offered by BLS, such as the share of each cut sold at a sale price, several different measures of “average” prices, and quantity indices for cuts and aggregates. With additional coverage and refinements to the scanner data, regional estimates and the relationship of price to quantities purchased could be developed—but at additional costs. Importantly, analysis of scanner data demonstrates that switching from averages based on fixed weights per outlet to sales-weighted averages almost invariably leads to lower estimated average prices and lower estimated price spreads. This analysis did not have access to quantities of meat sold, only to quantity indices. The indices allow one to demonstrate that the price affects the quantity purchased and a lower price stimulates buying. The indices do not allow one to compare the relative volume of one meat cut sold with another. Relative volume measures would greatly improve the quantity data.

BLS-based data have several advantages over scanner data. First, BLS’s data are available to the public free of charge as a byproduct of other BLS activities, whereas access to scanner data must be purchased. Second, BLS selects stores and products using statistically representative sampling techniques, whereas scanner data are based on a voluntarily assembled group of participating stores. To conduct this analysis, ERS assumed that the coverage of the scanner data is adequate to provide statistical reliability, but there is no way to assess the quality of the data. Scanner data exclude certain retailers that BLS makes efforts to include, which may affect average prices.

BLS-based data are released more quickly than scanner data—about 2 weeks after the end of the month. The lag of 7-8 weeks for scanner data stems from data collection and processing issues. Scanner data are a byproduct of data collected on packaged goods from data collection systems designed for packages, not random-weight products. Scanner-data collection procedures could be modified to speed the processing of random-weight data; these modifications might not improve the analysis of package-goods data.

It is expected that speeding up the release of scanner data may enhance their value. More current data may improve the analysis of current retail price conditions. ERS models allowed for the estimation of the additional information that scanner data provide when added to pre-existing data on wholesale and BLS-based retail prices. For most of the meats studied, the addition of scanner data contributed little to the analysis of market prices. Given the present production lag, scanner data for meat do not appear to contribute value to the analysis.

All the meat price data used in this study are byproducts of other activities. ERS calculates wholesale composites using data that AMS publishes on wholesale prices in its *Market News* program. ERS estimates of retail meat prices are based on the prices that BLS collects to measure inflation. Scanner data for random-weight meat packages are a byproduct of data collected for packaged goods companies. Processing of scanner data for random-weight products will likely improve only if demand for the data increases. While

demand has recently grown, it is unclear if the change has affected the cost of the data. In their current form, scanner data provide a different, but not necessarily better, view of retail meat markets than BLS-based data.

References

- Baumol, William J. *Economic Dynamics*, Macmillan Publishing Company, New York, 1970.
- Davis, Christopher G., and Biing-Hwan Lin. *Factors Affecting U.S. Pork Consumption*, Outlook Report No. LDP-M-130-01, U.S. Department of Agriculture, Economic Research Service, May 2005, www.ers.usda.gov/publications/ldp/may05/ldpm13001/ldpm13001.pdf.
- Davis, Christopher G., and Biing-Hwan Lin. *Factors Affecting U.S. Beef Consumption*, Outlook Report No. LDP-M- 135-02, U.S. Department of Agriculture, Economic Research Service, October 2005, www.ers.usda.gov/publications/ldp/oct05/ldpm13502/ldpm13502.pdf.
- Degner, Robert L. *Impact of Price Specials on Estimates of Retail Meat Prices*, Master's thesis, Texas A&M University, Department of Agricultural Economics and Rural Sociology, December 1970.
- Durbin, J., and S.J. Koopman. *Time Series Analysis by State Space Methods*, Oxford University Press, Oxford, UK, 2001.
- Gropp, Reint, and Arjan Kadareja. *Stale Information, Shocks and Volatility*, European Central Bank, Working Paper Series No. 686, October 2006, <http://www.ecb.eu/pub/pdf/scpwps/ecbwp686.pdf>
- Hahn, William F. *Beef and Pork Values and Price Spreads Explained*, Outlook Report No. LDP-M-118-01, U.S. Department of Agriculture, Economic Research Service, May 2004, www.ers.usda.gov/publications/ldp/apr04/ldpm11801/
- Kuchler, Fred, and Abeyayehu Tegene. *Did BSE Announcements Reduce Beef Purchases?* Economic Research Report No. 34, U.S. Department of Agriculture, Economic Research Service, December 2006, www.ers.usda.gov/publications/err34/
- Leibtag, Ephraim S. "Where You Shop Matters: Store Formats Drive Variation in Retail Food Prices," *Amber Waves*, U.S. Department of Agriculture, Economic Research Service, November 2005, www.ers.usda.gov/amberwaves/november05/pdf/featurewhereyoushopnovember2005.pdf.
- Mathews, K.H., Jr., M. Vandever, and R.A. Gustafson. *An Economic Chronology of Bovine Spongiform Encephalopathy in North America*, Outlook Report No. LDP-M-143-01, U.S. Department of Agriculture, Economic Research Service, June 2006, www.ers.usda.gov/publications/ldp/2006/06jun/ldpm14301/
- Perry, Janet, James MacDonald, Ken Nelson, William Hahn, Carlos Arnade, and Gerald Plato. *Did the Mandatory Requirement Aid the Market?—Impact of the Livestock Mandatory Reporting Act*, Outlook Report No. LDP-M-135-01, U.S. Department of Agriculture, Economic Research

Service, September 2005, www.ers.usda.gov/publications/ldp/sep05/ldpm13501/ldpm13501.pdf

Pritchett, James, Kamina Johnson, Dawn Thilmany, and William Hahn. "Consumer Response to Recent BSE Events," *Journal of Food Distribution Research*, 38:2:57-68, July 2007.

Shapiro, Matthew D., and David W. Wilcox. "Mismeasurement in the Consumer Price Index: An Evaluation," *Macroeconomics Annual*, 11:93-142, 1996.

U.S. Department of Commerce, U.S. Census Bureau. *Sector 44: Retail Trade: Subject Series—Establishment & Firm Size: Concentration by Largest Firms for the United States: 2002*, November 25, 2005, http://factfinder.census.gov/servlet/IBQTable?_bm=y&-ds_name=EC0244SSSZ6.

U.S. General Accounting Office (GAO). *Pork Industry: USDA's Reported Prices Have Not Reflected Actual Sales*, GA0/RCED-00-26, December 1999, <http://www.gao.gov/new.items/rc00026.pdf>

U.S. Government Accountability Office (GAO). *Livestock Market Reporting: USDA Has Taken Some Steps To Ensure Quality, But Additional Efforts Are Needed*, GAO-06-202, December 2005, <http://www.gao.gov/new.items/d06202.pdf>

Data Sources

ERS calculated the retail values of the Choice beef, pork, and broiler composites using BLS average prices for meat cuts. Data are available on the ERS website at www.ers.usda.gov/data/meatpricespreads/

The beef and pork tables are available with the most recent data for the Choice beef and pork composites. The retail-prices table contains the recent broiler composite. “Older” data may be accessed in the historical data worksheet.

Whole chicken and whole frozen turkey prices were taken directly from BLS data. In the BLS database, these items are identified as APU0000706111 and APU0000706311, respectively. Data are available on the BLS website at <http://data.bls.gov/cgi-bin/srgate>.

Appendix: Dynamic-Adjustment, State-Space Model: Equations, Specifications, and Tests

State-space models are generally specified using two sets of equations: one for the unobserved state variables and one for the observed variables. The three observed variables in this study are the wholesale price (WLS), the BLS-based retail price (BLS), and the sales-weighted scanner-based retail price (SWS). There are three state variables associated with these three prices. Two more state variables are meant to measure the effects of the two unseen retail prices, the best measure of the retail price (BMRP), and the third retail price (some other price, or SOP) that is averaged with the BLS-based retail price and scanner-based price to make the BMRP. Two more states are added to improve the models' fit and provide more price-spread effects.

The three observed prices, as noted above, are WHL, BLS, and SWS. The two unobserved price states are BMRP and SOP. The last two states are the wholesale price driver (WPD) and the retail price driver (RPD).

The observation equations, written in scalar form equations (1a-c), are:

$$(1a) \quad y_{t,WHL} = s_{t,WHL} + s_{t,WPD} + e_{t,WHL}$$

$$(1b) \quad y_{t,BLS} = s_{t,BLS} + s_{t,RPD} + e_{t,BLS}$$

$$(1c) \quad y_{t,SWS} = s_{t,SWS} + s_{t,RPD} + e_{t,SWS}$$

In equations 1a-c, y is an observed price, s is the state variable, and e is the transient error term. All the terms in equation 1a-c have two subscripts; the "t" subscript stands for a particular month and the state or observed variable defined above. The WPD state affects the wholesale price; the RPD state drives the two retail prices.

The dynamic adjustment of prices takes place in the first five states (WHL, BMRP, BLS, SWS, and SOP). In the dynamic-adjustment model, the current states are a function of this month's full-adjustment values, the previous month's states, and random components. ERS specified the full-adjustment values as linear functions of observed variables "x." Appendix table 1 provides the variables' names and descriptions.

This study uses some of the same features in the dynamic adjustment model from previous research on beef and pork price spreads (Hahn, 2004). One of the commonalities is an assumption about the relationship between the full-adjustment wholesale and retail prices. When fully adjusted, the retail price is the wholesale price plus a price spread, which is independent of the wholesale price. For example, if the full-adjustment wholesale price is \$2 and the full-adjustment price spread is \$1 per pound, then, the full-adjustment retail price is \$3. If the full-adjustment wholesale price rises 10 cents and the full-adjustment price spread does not change, then the full-adjustment retail price also goes up 10 cents. One can compare this type of behavior to the case where the retail price is always some percentage markup over the wholesale price, say 100 percent. If there is a 100-percent markup, doubling the wholesale price will double the retail price.

Appendix table 1

Exogenous variables in the state-space models

Name	Description	Subset
BeefQ	Log of beef production	X_l
ChicQ	Log of chicken production	
PorkQ	Log of pork production	
TurkQ	Log of turkey production	
Corn	Corn price from USDA, Agricultural Marketing Service (AMS) - #2 Yellow, Central Illinois	
Soy	Soy meal price from USDA, AMS - High protein, Central Illinois	
AOG	Consumer Price Index (CPI) for all items less food and shelter	
Fish	CPI for fish and seafood	
Milk	CPI for dairy and related products	
CAN0, CAN1	CAN0 measures the immediate effect of the bovine spongiform encephalopathy (BSE) outbreak in Canada, CAN1 the ending effects	
USA0, USA1	USA0 is the immediate effect of BSE in the United States, USA1 the ending effects CAN0 and USA0 go to CAN1 and USA1 over 6 months	
x0, x1, x2	Intercept, trend, and trend-squared, post scanner (January 2001-August 2005)	X_r
xb1, xb2	Trend and trend-squared, prescanner (January 1998-December 2000)	
JAN, FEB, MAR, APR, MAY, JUN, JUL, AUG, SEP, OCT, NOV, DEC	Twelve monthly dummies for seasonal effects	X_r, X_d

Source: USDA, Economic Research Service.

The assumed wholesale-retail price relationship with full price adjustment makes sense economically if:

- A fixed proportion relationship exists between wholesale meat inputs and retail meat output
- Meat retailing is reasonably competitive with constant returns to scale

The x variables are divided into three subsets. The first set is the “l” (level) subset. It includes beef, pork, chicken, and turkey production, factors that would shift the supply of meats, such as the cost of feedstuffs, and factors that would shift the demand for meat, such as the prices of other foods. This subset also includes variables meant to measure the effects of the discovery of bovine spongiform encephalopathy (BSE), first in Canada and then in the United States. Canada’s BSE cases have had major repercussions on U.S. beef imports, while the discovery of BSE in the United States caused a severe

drop in U.S. beef exports (Mathews et al., 2006). The BSE-related variables are in the “XI” subset and will not affect target price spreads.

The next set of variables is the “r” (“rest” of the variables) subset, which includes the intercept, trend, trend squared, and monthly dummies. As noted previously, it is assumed that the full-adjustment values of the retail prices are to be the wholesale price plus a price spread. Similar to Hahn (2004), it is assumed that the full-adjustment wholesale price is a function of all the x variables and that the full-adjustment price spreads are a function of the first subset.

Based on visual examination, seasonality is important in the turkey price spreads. November invariably has the year’s lowest retail prices, highest wholesale prices, and lowest price spread. Other meats may have seasonal price spreads; steaks, chops, and ribs tend to be featured heavily during the summer grilling season, as is ham during Christmas and Easter. Therefore, seasonality variables are added to the previous Hahn model. The seasonal dummies are added in the “r” subset and a “d” subset is included for the seasonal dummies exclusively.

The equations for the full-adjustment values of the states can be written:

$$(2a) \quad f_{t,WHL} = \mathbf{b}_l \mathbf{x}_{t,l} + \mathbf{b}_r \mathbf{x}_{t,r}$$

$$(2b) \quad f_{t,j} = \mathbf{b}_l \mathbf{x}_{t,l} + (\mathbf{b}_r + \mathbf{c}_{r,j}) \mathbf{x}_{t,r} \text{ for } j=\{\text{BLS, SWS, SOP, BMRP}\}$$

In equations 2a and 2b, the “f”s are the full-adjustment values for the first five states. The term “b” is a vector of coefficients that make the wholesale price’s full adjustment value, and “x_t” is the vector of exogenous variables for month “t.” Both b and x are split into their “l” and “r” subsets. The “c_{r,j}” is a set of vectors that make the full-adjustment price spreads for the four retail prices. Equation (2b) can be used to explain the wholesale price also by setting the c_{r,WHL} vector to all 0s.

The x_r set contains an intercept and all 12 of the monthly dummies. These 13 variables are perfectly collinear, which means multiple sets of coefficients will give the same ending results. The coefficients are made unique by the requirement that the seasonal dummy coefficients sum to 0. The dummies’ coefficients then represent their month’s typical deviation from the annual average.

The BMRP state is defined as a weighted average of the BLS, SWS, and SOP states. The x_l effects on all five full-adjustment values are the same, so the price spreads are constrained as follows:

$$(3) \quad \mathbf{c}_{r,PBMRP} = \alpha_{BLS} * \mathbf{c}_{r,BLS} + \alpha_{SWS} * \mathbf{c}_{r,SWS} + \alpha_{SOP} * \mathbf{c}_{r,SOP}$$

The Greek letter alpha, α_j is a positive weight used to average the BLS, scanner, and SOP states up to the BMRP state. The three alphas sum to 1.

In the dynamic adjustment model, the current states are a function of the current value of the full-adjustment values, the previous months’ states, and

random components. Equation 4 is a general version of the dynamic-adjustment model:

$$(4) \quad \mathbf{S}_t = \mathbf{A} * \mathbf{S}_{t-1} + [\mathbf{I} - \mathbf{A}] * \mathbf{F}_t + \mathbf{U}_t$$

In equation 4, \mathbf{F}_t is a vector of full-adjustment values, \mathbf{S}_t is the vector of state variables, (\mathbf{S}_{t-1} is the previous month's vector), \mathbf{A} is an adjustment matrix, \mathbf{I} the identity matrix, and \mathbf{U}_t a vector of random error terms. Equation 4 is set up so that if all the errors are 0, and if \mathbf{S}_{t-1} equals \mathbf{F}_t , then \mathbf{S}_t equals \mathbf{F}_t . If the matrix \mathbf{A} implies partial adjustment, then the current state will be somewhere between the previous month's state and the current month's full-adjustment value. In most state-space applications, what is referred to as the "adjustment matrix" is generally called the transition matrix.

The desire is for the transition matrix to be "stable." If it is stable, \mathbf{F}_t is fixed for a long period of time, and all the error terms are made zero, simulating the process, and \mathbf{S}_t will get closer and closer to \mathbf{F}_t .

The averaging constraint from equation 3 imposes restrictions on the adjustment matrix, \mathbf{A} , and on the random components, the \mathbf{U}_t . These issues are addressed later. Now, equation 4 is made simpler and the observation equations (1a-c) are made more complicated. Equation 4 is simple enough, but it is complicated first by \mathbf{F}_t being replaced with the function of the exogenous variables. These are split into two groups:

$$(5) \quad \mathbf{S}_t = \mathbf{A} * \mathbf{S}_{t-1} + [\mathbf{I} - \mathbf{A}] * [\mathbf{1}] * \mathbf{b}_t * \mathbf{x}_{t,l} + \mathbf{C} * \mathbf{x}_{t,r} + \mathbf{U}_t$$

In equation 5, the term $[\mathbf{1}]$ is a five-row, one-column vector of ones, and \mathbf{C} is a matrix of coefficients defined by stacking up the \mathbf{b}_r and \mathbf{c}_r :

$$(6) \quad \mathbf{C} = \begin{bmatrix} \mathbf{b}_r \\ \mathbf{b}_r + \mathbf{c}_{r,BMRP} \\ \mathbf{b}_r + \mathbf{c}_{r,BLS} \\ \mathbf{b}_r + \mathbf{c}_{r,MPR} \\ \mathbf{b}_r + \mathbf{c}_{r,SOP} \end{bmatrix}$$

Equations 4 and 5 are linear difference equations. One of the strategies in solving these types of systems is to strategically separate them. Baumol shows several examples on how this is done using the linear structure of the model. The strategy here is to separate those parts associated with the \mathbf{x}_r variables.

Dropping the random portion and \mathbf{x}_l from equation 5 leaves 7:

$$(7) \quad \mathbf{S}_t = \mathbf{A} * \mathbf{S}_{t-1} + [\mathbf{I} - \mathbf{A}] * \mathbf{C} * \mathbf{x}_{t,r}$$

Baumol shows specifically in cases where the \mathbf{x} consists of intercepts, trends, and squared trends that a solution for \mathbf{S}_t will take the form:

$$(8) \quad \mathbf{S}_t = \mathbf{D} * \mathbf{x}_{t,r},$$

where \mathbf{D} is a matrix of coefficients whose values depend on the \mathbf{A} and \mathbf{C} . This type of procedure works because the \mathbf{x}_t variables are simple functions of time. The X effects are much more complicated, and simple solutions such as equation 8 do not exist.

The set of exogenous variables in equation 8 also contains monthly dummies for each month t . Baumol's procedures can be generalized to include monthly dummies. Solving by parts, as is done here, doubles (at least) the state variables.

Returning to equation 5, some terms are not used in equation 6, such as:

$$(9) \quad \mathbf{S}_t = \mathbf{A} * \mathbf{S}_{t-1} + \mathbf{a}_v * \mathbf{b}_l * \mathbf{x}_{t,l} + \mathbf{U}_t \text{ where}$$

$$(10) \quad \mathbf{a}_v = [\mathbf{I} - \mathbf{A}] * [\mathbf{1}]$$

A specification like equation 10 has the full-adjustment values of all five state variables with the same value, $\mathbf{b}_l * \mathbf{x}_{t,l}$. In the model, that common, full-adjustment value is called the "target" value. Having the same full-adjustment value for these five states makes programming the model easier. The "split" state variables can be incorporated into the observation equation by defining the $s_{t,i}$ as in equation 5e and writing equation 1a-c as:

$$(11a) \quad y_{t,WHL} = s_{t,WHL} + s_{t,WPD} + \mathbf{d}_{WHL} * \mathbf{x}_{t,r} + e_{t,WHL}$$

$$(11b) \quad y_{t,BLS} = s_{t,BLS} + s_{t,RPD} + \mathbf{d}_{BLS} * \mathbf{x}_{t,r} + e_{t,BLS}$$

$$(11c) \quad y_{t,MPR} = s_{t,MPR} + s_{t,RPD} + \mathbf{d}_{MPR} * \mathbf{x}_{t,r} + e_{t,MPR}$$

The term \mathbf{d}_i is the appropriate row from the matrix \mathbf{D} . The states, $s_{t,i}$ now just account for the \mathbf{x}_l and random effects, while the $\mathbf{d}_i * \mathbf{x}_{t,r}$ account for \mathbf{x}_r effects. Because the states share the same full-adjustment value, dynamic adjustment tends toward making the states similar to one another, which makes the prices similar. Some may adjust more quickly. The $\mathbf{d}_i * \mathbf{x}_{t,r}$ part makes the observed prices different from one another.

The \mathbf{D} matrix is based on the \mathbf{C} matrix, which has five rows, one a function of the other three. Thus, there are only three rows of the \mathbf{D} matrix in the observation equations. However, the coefficients of the \mathbf{C} matrix cannot be calculated given what is observed from the \mathbf{D} coefficients. This is the other reason the intercept, trend, and seasonal terms are removed from the state equations. Even if the intercept, trends, and seasonal factors are kept in the state equations, there would have been multiple versions of the \mathbf{C} matrix consistent with the observed data.

In equations 11a-c, observed wholesale-retail price spreads are driven by four factors: the difference in adjustment speed of the three prices ($s_{t,i}$), price spread adjustment ($\mathbf{d}_i * \mathbf{x}_{t,r}$), the difference between WPD & RPD, and the transient effects ($e_{t,i}$). The intercept, trend, and trend-squared parts of price-spread adjustment will give a smooth pattern to interpret. The price-driver states, WPD and RPD, are introduced to allow for some "nonsmoothness," or unpredictability, of price spreads. WPD and RPD are not driven by exogenous variables; they have no explicit full-adjustment values.

Until this point, the description of the structure of the adjustment matrix has been vague. The seven state variables require seven equations. Many of the elements of the \mathbf{A} matrix to zero are restricted. This makes it easier to show the state-equations a few equations at a time. The first two equations are written for the RPD and WPD states. These two states interact only with one another:

$$(12) \quad s_{t,WPD} = a_{1,1} * s_{t-1,WPD} + a_{1,2} * s_{t-1,RPD} + u_{t,WPD}$$

$$(13) \quad s_{t,RPD} = a_{2,1} * s_{t-1,WPD} + a_{2,2} * s_{t-1,RPD} + u_{t,RPD}$$

The double-subscripted, small-case “a” in equations 12 and 13 and following equations are coefficients from the \mathbf{A} matrix. The double-subscripted lower-case “u” is the individual element of the random vector \mathbf{U}_t . Both of these state variables are driven by random factors and their lagged values. If their coefficients imply a stable relationship, these states will tend to adjust toward 0. Remember, RPD affects both retail prices, and WPD affects the wholesale price. The structure of these two terms allows for general interaction between retail and wholesale prices; one price is not specified as a leader nor is the other specified as a follower.

WHL and BMRP states interact with one another exclusively. These two states are also affected by the exogenous variables in \mathbf{x}_t :

$$(14) \quad s_{t,WHL} = a_{3,3} * s_{t-1,WHL} + a_{3,4} * s_{t-1,BMRP} + (1 - a_{3,3} - a_{3,4}) * \mathbf{b}_l * \mathbf{x}_{t,l} + u_{t,WHL}$$

$$(15) \quad s_{t,BMRP} = a_{4,3} * s_{t-1,WHL} + a_{4,4} * s_{t-1,BMRP} + (1 - a_{4,3} - a_{4,4}) * \mathbf{b}_l * \mathbf{x}_{t,l} + u_{t,TNA}$$

Just as in equations 9 and 10, this specification makes no assumptions about either price being leader or following the other. The last three equations determine the BLS, SWS, and SOP states. These share a common structure:

$$(16) \quad s_{t,BLS} = a_{4,3} * s_{t-1,WHL} + a_{4,4} * s_{t-1,BMRP} + a_{5,5} * s_{t-1,BLS} + a_{5,6} * s_{t-1,SWS} + a_{5,7} * s_{t-1,SOP} + (1 - a_{4,3} - a_{4,4} - a_{5,5} - a_{5,6} - a_{5,7}) * \mathbf{b}_l * \mathbf{x}_{t,l} + u_{t,BLS}$$

$$(17) \quad s_{t,SWS} = a_{4,3} * s_{t-1,WHL} + a_{4,4} * s_{t-1,BMRP} + a_{6,5} * s_{t-1,BLS} + a_{6,6} * s_{t-1,SWS} + a_{6,7} * s_{t-1,SOP} + (1 - a_{4,3} - a_{4,4} - a_{6,5} - a_{6,6} - a_{6,7}) * \mathbf{b}_l * \mathbf{x}_{t,l} + u_{t,MPR}$$

$$(18) \quad s_{t,SOP} = a_{4,3} * s_{t-1,WHL} + a_{4,4} * s_{t-1,BMRP} + a_{7,5} * s_{t-1,BLS} + a_{7,6} * s_{t-1,SWS} + a_{7,7} * s_{t-1,SOP} + (1 - a_{4,3} - a_{4,4} - a_{7,5} - a_{7,6} - a_{7,7}) * \mathbf{b}_l * \mathbf{x}_{t,l} + u_{t,SOP}$$

Equations 16-18 “recycle” the $a_{4,3}$ and $a_{4,4}$ coefficients from equation 15. The coefficients for the previous month’s WHL and BMRP states are the same as

those in the BMRP equation. This keeps all four retail prices relatively close to one another. Also, as noted earlier, the desire if for the TNS state to be a weighted average of the other three “retail” price states. Reusing the $a_{4,3}$ and $a_{4,4}$ coefficients means that the lagged WHL and RET affects the average automatically. In addition, the following side constraints are used on the “a” coefficients and “u” random errors:

$$(19) \quad 0 = \alpha_{BLS} * a_{5,i} + \alpha_{MPR} * a_{6,i} + \alpha_{SOP} * a_{7,i}, \text{ for } i=(5, 6, \text{ and } 7)$$

$$(20) \quad u_{t,RET} = \alpha_{BLS} * u_{t,BLS} + \alpha_{SOP} * u_{t,SOP} + \alpha_{MPR} * u_{t,MPR}$$

All the models are estimated using maximum-likelihood estimation. The general model is specified earlier, and more constrained models, are specified later. The special cases impose constraints on the models, and these constraints will not make the likelihood-value higher. Testing hypotheses involves comparing the estimated likelihood of constrained and unconstrained models. Twice, the likelihood difference generally has an asymptotic chi-square distribution. Some of the hypothesis tests will have odd properties, as discussed later.

Model Restrictions: Restrictions on the Averaging Parameters

The averaging parameters—the alphas—determine how the individual retail prices add up to the BMRP state. Tests determine if either the BLS-based price or the store-weighted, scanner-based price represents the true national average retail price. Also, data are checked to determine if an average of only the BLS and SWS prices (not “some other price,” SOP) approximates the national average. The first two cases mentioned are a special case of this last one. Making the RET *not* include SOP means setting the alpha for SOP equal to 0. This imposes one restriction on the model. Moving from the model where SOP is not part of BMRP to either the BLS-is-BMRP or the SWS-is-BMRP cases requires one more restriction. The three “alpha” must sum to 1, so if two are fixed, the value of the remaining parameter is known.

What would SOP do in the model if its alpha coefficient is 0? Making alpha equal to 0 will not change the basic structure of equations 16-18, just how the constraint operates. The previous month’s value of SOP will affect both BLS and SWS, however, so the averaging constraint will cancel out these effects in BMRP and will prevent SOP from affecting the next month’s wholesale price.

The alpha parameters are required to be positive, which complicates the tests somewhat. Most statistical theory is based on testing parameters that are *not* constrained in size or sign. For example, suppose the true alpha for one of the states is one. In unconstrained estimation, some of the estimates may be larger than one; however, the chances that an estimate will be exactly one are practically zero. Once the estimates are constrained to be no larger than one, there is a chance that an estimate will be exactly one. In models without bounds on variables, imposing a constraint always decreases the likelihood. Putting the upper level of all the “alphas” at one means that occasionally an “alpha” is estimated to be one in the otherwise unconstrained model. In this case, the more- and less-constrained likelihoods are the same. The same

process applies on the lower bound of zero. An unconstrained, estimated alpha could be negative; the chances of an unconstrained estimate being exactly zero are also effectively zero. Because all the estimated alphas are required to be positive, there is a chance that some of the estimates are zero.

The alpha-related tests will have a chance of being exactly 0 if the hypothesis is true (or close to being true). What the statistical tables call a 5-percent value would occur less frequently than 5 percent of the time. Rejecting alpha-related restrictions with large test statistics is a somewhat more comfortable choice than accepting those with small test statistics.

Model Restrictions: Requiring BLS-based and Scanner Prices To Have the Same Speed of Adjustment

One of the special models shows the BLS-based and scanner prices adjusting at the same rate, which requires additional side restrictions to equations 13 and 14. Both the BLS-based and scanner states will have the same reaction to the lagged SOP state by imposing:

$$(21) \quad a_{4,7} = a_{5,7}$$

In addition to equation 21, there is the following constraint:

$$(22) \quad a_{4,4} + a_{4,5} = a_{5,4} + a_{5,5}$$

These two restrictions do not have any associated sign or size constraints. The test-statistic for the “same-speed-of-adjustment” hypothesis ought to be asymptotically chi-square.

Model Restrictions: Constraining Transient Effects

The random effects (“e”) terms in the observation equations are transient effects. Forcing BLS and scanner data to have the same-sized transient effect means that their e variances have to be constrained to be equal. This is one restriction. Making all the e uncorrelated means restricting each of the three covariance terms to 0. All three variances must be positive, which complicates tests for the transient effects in some cases.

All variances of e must be positive. It is possible that the optimal estimate for an e variance is zero. If it is, then the observation equation actually has no e terms. If two or more of the observation equations have no e , then all of them are uncorrelated automatically. Making them uncorrelated when two or all are zero is not restrictive. If the unconstrained BLS and SWS variances are zero, then the “BLS-and-SWS-have-the-same-transient-effects” model also have the same likelihood as the unconstrained variances.

Appendix table 2 provides the results of the tests. Only two of the more restricted models passed. Pork’s BLS and SWS have the same speed of adjustment, and broiler’s BMRP is a weighted average of the BLS and SWS prices. It happens to be the case that broiler’s SOP weight hit its lower bound of zero. Consequently, requiring that the BMRP be an average of BLS and SWS is not actually restrictive.

Appendix table 2

Significance levels for the hypothesis test results based on an asymptotic chi-square distributions

Hypothesis	Model degrees of freedom	Choice beef composite	Pork composite	Whole chickens	Whole, frozen turkeys	Broiler composite
<i>Significance level (percent)</i>						
BLS-based price and scanner-based price have same speed of adjustment	2	0.0	21.3	0.0	0.1	0.0
BLS-based price equals national average	2	0.0	0.0	0.0	0.0	0.0
Scanner-based price equals national average	2	0.0	0.0	0.0	0.0	0.0
An average of the BLS-based and scanner prices is the national average (SOP2 has a weight of 0.)	1	0.0	0.0	0.0	0.0	100.0
BLS-based and scanner prices have the same transient variance	1	0.0	0.0	0.0	0.0	0.0
All three prices' transient effects are uncorrelated	3	0.0	0.0	0.0	0.0	0.0

Note: Highlighted cells are not statistically significant.

Source: USDA, Economic Research Service comparison of results from alternative state-space models.

Using State-Space Estimates To Measure the Value of Scanner Data

The dynamic-adjustment model was a state-space model in which the accuracy of predicting the states improved as the data set lengthened. There is a limit to the accuracy of these predictions—the measurement/forecast accuracy of the states/observed variables approaches what is called a steady-state value. The estimates tend to move toward their steady-state values fairly quickly; however, side analysis helps ensure that the steady-state values are actually calculated. These calculations are the basis of the “all-data-released-at-the-same-time” results.

To estimate the “actual-release-pattern” results for instances in which there are only BLS and WHL data, the analysis starts with the steady state results and then calculates the effects of adding just the wholesale price. This provides the result when there is a long series of both prices, but one more observation on the wholesale price. For the “see-all-three-prices” results; the analysis starts with the steady state, adds a month’s worth of BLS and WHL, and then adds a new month of WHL exclusively.