Connectionist-Based Rules Describing the Pass-through of Individual Goods Prices into Trend Inflation in the United States

Vincent A. Schmidt
Air Force Research Laboratory
Dayton, Ohio USA

Jane M. Binner
Aston University
Birmingham, UK

Richard G. Anderson
Federal Reserve Bank of St. Louis
St. Louis, Missouri USA

Abstract

This paper demonstrates a mechanism whereby rules can be extracted from a feedforward neural network trained to characterize the inflation "pass-through" problem in American monetary policy, defined as the relationship between changes in the growth rate(s) of individual commodities and the economy-wide rate of growth of consumer prices. Monthly price data are encoded and used to train a group of candidate connectionist architectures. One candidate is selected for rule extraction, using a custom decompositional extraction algorithm that generates rules in human-readable and machine--executable form. Rule and network accuracy are compared, and comments are made on the relationships expressed within the discovered rules. The types of discovered relationships could be used to guide monetary policy decisions.[1]

Keywords: Consumer prices, Inflation, Neural Network, Data Mining, Rule Generation

1 Introduction

In recent years the relationship between changes in the prices of individual goods and services and overall consumer inflation has assumed prominence in academic literature and in Central Banks' circles. Some Central Bankers explicitly have accepted that only aggregates of subsets of consumer prices should be objectives of monetary policy, while others prefer overall "headline" inflation as a target. Perhaps the most prominent among the former is the U.S. Federal Reserve, which has accepted the core (excluding food and energy prices) chain-price index for personal consumption expenditures ("core PCE") as its policy objective. Among the latter are the Bank of England and the European Central Bank, which prefer the headline measure because it includes all the products purchased by consumers, including food and energy. Despite their differences, all Central Bankers desire a relatively low and stable trend in inflation. The ability of Central Bankers to attain this goal depends, in part, on their own behavior. Suppose the price of energy increases sharply. What is the reaction of other prices? If most households and firms believe the Central Bank will seek to prevent any increase in the overall rate of inflation, then prices of goods other than energy will tend to fall—policy put in place by the Central Bank will limit aggregate demand in the economy for goods and services until such price decreases occur, thereby achieving their overall objective of stabilizing the trend rate of growth of the overall headline price level. But, suppose instead that most households and firms believe the Central Bank fears slower economic activity and its accompanying higher unemployment and political difficulties. In this case, the Central Bank will not implement policies to prevent the increase in the headline inflation rate—and the firms and households will reasonably believe that no pressure via slower economic activity will arise to force lower prices. The relationships that link movements in individual prices to the aggregate inflation trend are difficult to estimate because they vary through time. Changes in a nation's political leadership may refocus concern on price stability versus more rapid growth of economic activity. In addition, the pattern of price increases across goods and services changes through time. The pattern of external shocks (such as weather patterns and energy prices) affecting the economy varies, as do the levels of economic activity in trading partner countries. Further, the opinions to firms and households

[1] Paper cleared for public release #88ABW-2010-1111

may differ among goods, with some price changes eliciting strong reactions and others little if any reaction. Such variation will affect the strength and pattern of pass-through from changes in individual prices to the overall headline inflation rate. Our goal is to test whether the strength of the pass-through effect is larger for some prices than for others. Energy prices often are regarded as the most likely to have large pass-through effects because energy is purchased frequently and, via transport, is an essential input to the availability of many other products. Prices of consumer durable goods, such as cars and home furnishings, are expected to have the weakest pass-through because these purchases are more readily deferred. Intermediate are prices for foods because less expensive food products may be substituted when prices increase sharply.

Our work is related to a large literature on the broader economic effects of oil price shocks, most notably Hamilton [1], which documents a statistical link between oil price shocks and post war recessions in the United States. This study, along with those of Hooker [2] and Barsky and Kilian [3] debates the robustness of important real and inflationary effects of oil price shocks to different price specifications, assumptions of exogeneity, and, importantly, evidence of a weakening of the effects of oil prices in more recent data. The sustained increase in oil prices spanning the end of 2001 until the summer of 2008 has generated a plethora of papers reporting a reduction on the pass-through of energy prices to broader inflation in the United States and elsewhere (see for example [4–9]). Recent surveys of the economic literature on energy prices can be found in Segal [10] and Kilian [11]. In general, the appreciation of the domestic currency, a more active monetary policy in response to inflation, and a higher degree of trade openness are found to explain the decline in oil price pass-through, although the evidence has largely relied on relatively simple methods of assessing changes over time.

This paper addresses the problem of whether policymakers should respond strongly to changes in individual prices, that is, whether they should anticipate a strong reaction of the headline inflation trend.

Our techniques in this study are similar to ones we have used previously to study the relationship between movements in individual financial assets held by households and firms and economy-wide headline inflation [12]. We discover interesting relationships using simple and standard feedforward connectionist models; using a newly designed and revised rule extraction algorithm further simplified and reduced the number of rules. (These rules are still automatically produced as a collection of MATLAB-based human-readable and machine-executable if-then rules, expressing the discovered relationships in terms of the original data.) Our data are monthly, published by the U.S. Bureau of Labor Statistics.

An explanation of our experimentation follows, to include a description of the identification of individual assets and their encoding, details regarding the selection and training of the neural network architecture, and subsequent rule extraction and representation. The discussion of the rules is expected to add insight that could be used to guide monetary policy decisions.

2 Dataset Preparation

Historical U.S. consumer price index data, for the aggregate index and its components, were obtained in order to investigate the relationship between changes in the component prices and several trend estimators applied to the headline index. Our overall index is the Bureau of Labor Statistics CPI-RS, a research series that has been constructed for historical dates using the same definitions and methods used for newly published data (see [13]). The beginning of our data is limited by the availability of components. The training data used for connectionist model selection included monthly seasonally adjusted values from December 1977 through December 2009, a total of 385 exemplars. Inflation was constructed for each month as year-on-year growth rates of prices. All data are obtained from the web site of the U.S. Bureau of Labor Statistics. Table 1 enumerates the components available within this dataset.

The component data were prepared by calculating the percentage of increase in value for corresponding months in consecutive years. The trend in headline inflation was prepared as a 12-period lagged moving average of the headline monthly rate. This reduced the dataset to 373 exemplars. The automated clustering algorithm we've used in previous studies (there, to examine quarterly Di-

Table 1: Component Data

Variable	Description
urs	all items, sa
ursxfdg	all items except food and energy, sa
ursn	all items, nsa
ursfn	food & bev
ursfdn	food
urshn	housing
ursan	apparel
urstn	transportation
ursmn	medical care
ursennn	entertainment and recreation
ursxfdgn	all items less food and energy
ursegn	energy

Table 2: Inflation % change ranges

Category	Inflation % change Range
1	$< 1.25\%$
2	(1.25%, 2.0%)
3	(2.0%, 3.0%)
4	(3.0%, 5.0%)
5	(5.0%, 9.0%)
6	$> 9.0\%$

visia data) was used again this time to discretize the monthly component data.

Components were represented using thermometer encoding. After inspection, the headline inflation trend was manually discretized (using mutex, or 1-of-N, encoding) into the six distinct ranges indicated in Table 2. These encoding schemes were selected based on the successful training of quarterly Divisia data in our past work. (Mutex and thermometer encoding schemes are commonly used to prepare discretized data for neural network consumption.)

The dataset was visually inspected before considering any specific neural network architecture in order to select the input components most likely to influence inflation. (Reducing the size of the input dataset results in a desirable reduction of network and rule complexity, as long as accuracy is not sacrificed with such a decision.) Based on this initial inspection, it was determined that transportation (URSTN) and energy (URSEGN) seemed to be the candidates most capable of influencing inflation (URSN). The graph shown in Figure 1 (top of next page) depicts the recalculated values of these

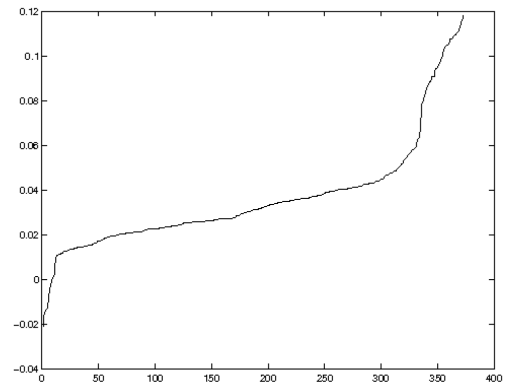

Figure 2: Inflation (y) vs sorted by % increase (x)

categories of data, independently scaled to emphasize the relationship of transportation and energy to inflation. Although the details of this graphic are nearly unreadable at this size, the important point of this figure is to note how closely the inflation trend is tracked by transportation and energy trends.

Once the categories were selected and recomputed, the clustering algorithm originally developed by Schmidt [14] was used to automatically bin the inputs (transportation and energy) into 12 and 17 bins, respectively. Table 3 identifies the input components selected for training, and enumerates the breakpoints between the bins for these inputs. The bins are used to recode the original inputs into thermometer-encoded binary vectors, which are presented to the network for training and for testing.

The same algorithm was originally used to cluster the inflation values, but inspection of the values and a small test of manually selected breakpoints resulted in the binning for inflation indicated in Table 2. The inflation output was recoded as a 1-of-N binary vector of length 6, where a single true value indicates the desired representative bin. Figure 2 shows the inflation values, sorted by percentage of change, for all data points. It is evident that reasonable breakpoints happen near 0.0125, 0.02, 0.03, 0.05, and 0.09 on the Y axis (increase ratio), as indicated by slope changes in the graph. (The points near 0.0125 and 0.05 are the most evident; others are reasonable rough estimates based on slight changes in the slope. More detailed future analysis should yield more precise breakpoints.)

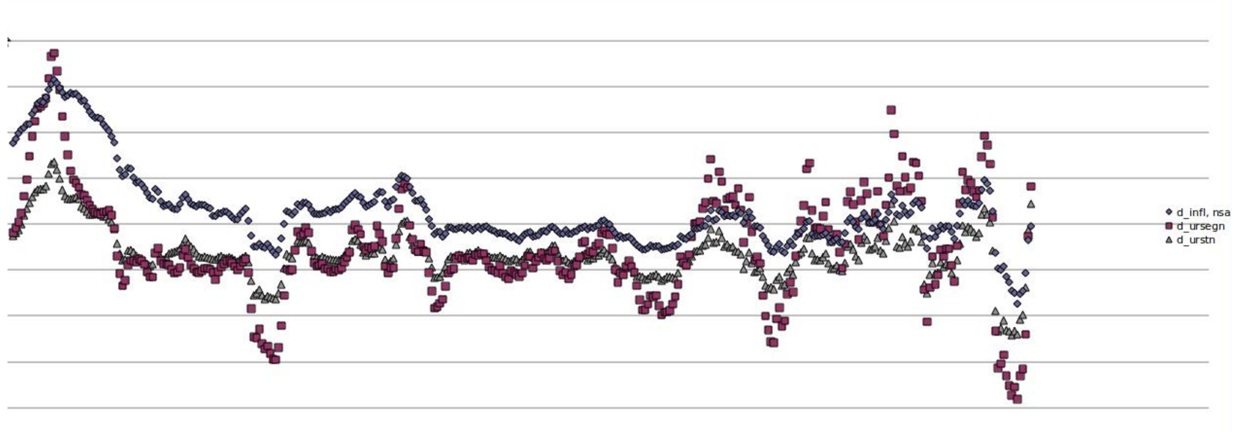

Figure 1: URSN, URSTN, URSEGN values

Table 3: Transportation & Energy encoding

Category (Attribute)	Symbol	Levels	Breakpoints Between Bins			
Transportation	URSTN	12	-0.1213	-0.1012	-0.0707	-0.0325
			-0.0111	0.0147	0.0746	0.1358
			0.1666	0.1925	0.2209	
Energy	URSEGN	17	-0.2652	-0.2048	-0.1074	-0.0221
			0.0508	0.1067	0.1390	0.1637
			0.1857	0.2151	0.2535	0.2825
			0.3214	0.3716	0.4091	0.4470

The input and output encoding suggests the neural architecture will have 29 (12 + 17) binary-valued inputs and six binary outputs. Although the neural model uses these binary vectors for training and testing, the final extracted rules are expressed in terms of the original continuous-valued data (ratio of increase, year-on-year, for each component). This is the same data preparation and encoding style we've used for UK Divisia and US MSI experiments performed in previous years [15, 16].

3 Model Selection & Analysis

Our model continues to evolve as we learn from the successful (and not quite so successful) studies we've executed with UK Divisia data and US MSI data in the past. The Research Series we use in this year's experiment is (in general) larger than similar data we've had before, allowing for a more robust-sized training and testing dataset. Access to a greater quantity of relevant data is almost always preferable when it comes to training and testing neural networks.

A collection of feedforward neural models was trained, where the number of nodes in the hidden layer was the only parameter changed; networks with 2, 3, 4, 5, 6, 7, 8, 9, and 12 hidden layer nodes were tested. For all networks, the single hidden layer applied a traditional sigmoid activation function (MATLAB's *logsig* function), and the unconstrained linear function (MATLAB's *purelin*) was used for the six nodes at the output layer.

All candidate architectures were provided the identical set of randomly selected input/output pairs. Of the 373 data points available, 65% (242) were randomly chosen for training, and the remaining 35% (131) were used for testing the resultant network. All networks were trained for 2500 epochs, and for each architecture, the single best of 25 instances was chosen for evaluation. Note that the *best network instance*[2] was defined as the

[2] Although 25 network instances were trained, solutions tend to fall into small number of "classes" or "categories." In this case, the *best network* is merely a reasonably random selection of a single solution from the category of networks exhibiting the most desirable behavior. Any network instance within this class would be a suitable selection for rule extraction. Also, we are merely using the network to

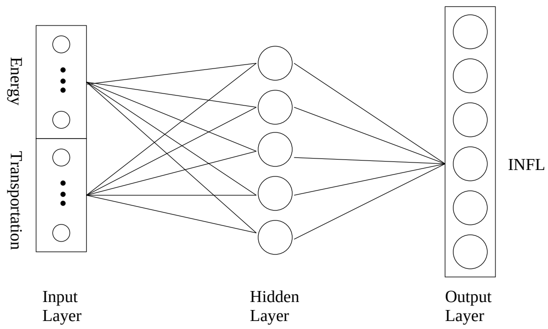

Figure 3: Architecture Selection

instance with the highest training and testing accuracy. All model execution was performed using MATLAB 5.3 (R11) on a Slackware 11.0 Linux-based system running within a VirtualBox virtual machine on an AMD Athlon 64-X2 with 2Gb RAM. No instance of model training, in any configuration, exceeded 9.5 seconds.

In general, we'd like to extract rules from the most accurate network with the least number of nodes in the hidden layer in order to avoid combinatorial explosion during rule extraction. Interestingly, all networks claimed accuracy in the range of 87%–89% for training data, and 82%–84% for testing data. Although this would suggest we select the network model with two nodes in the hidden layer, constraints in the way the rule extractor operates forced us to choose the model with five nodes in the hidden instead; no smaller network architecture was able to produce the intermediate states required for rule extraction. (We do not go into further detail about these constraints here.) A representative drawing of the selected network architecture is shown in Figure 3.

We continue to rely on a decompositional approach to rule extraction, the same approach used in our previous research and described in detail in Schmidt [14], and Schmidt and Chen [17].

There are six nodes in the output layer. For any given input, only a single node should generate *True*, while the remaining five nodes generate *False*. Extracted rules should replicate this behavior: the rules based on a specific output node should return True or False for exactly the same input values as the neural network. Our hope is that the actual set of rules representing an output node is descriptive and can be easily validated by a subject matter expert. (More detail about the actual rules is given in the next section). Table 4 highlights these points. For each output node, the range of outputs in that bin is identified, along with the number of rules covering this node, and the training and testing accuracy for the rules.

The table also includes "Train Targets" and "Test Targets" columns. These columns indicate the number of targets of training and testing data associated with each output node (bin).

discover and describe relationships within the data. The extraction algorithm generates rules, the "real" product of this exercise, which are close, but not exact, representations of the selected neural network. This further decouples the specific network from the end result.

4 Rule Analysis

There are six output nodes, corresponding to the ranges shown in the previous section. The rule generator produces a separate file representing rules describing each range of values. All rules within a file are tested against the input variables (here, % increase for each variable). If any rule in the file "matches" an input data exemplar, that data is said to be represented by the output range defined by that file. For example, all rules in the "node 5" output file describe conditions producing inflation increases in the range 5%–9%.

All of the rules are represented as a collection of expressions in an if-then statement. Each expression is formatted like: (low_value <= attr & attr <= high_value), the mathematical equivalent to: low_value <= attr <= high_value, where "attr" is the value of the attribute being tested. The symbols "&" and "|" are logical "AND" and "OR" operations, respectively, and Inf represents infinity. The logic of the entire rule must evaluate to True for the rule to be true. If a rule does not include an attribute, then that attribute is not required for the given rule. The rules are also numbered for our (human) reference.

Figure 4 shows an example of a rule (rule #3 from node 2's output, in this example) extracted from our trained network. Note the explicit reference to the continuous values of *urstn* (transportation) and *ursegn* (energy). Also note the human-readable format and nature of extracted rules, which makes them ideal for validation by subject-matter experts. Since the rules are also represented as MATLAB code, they can also be executed by computer and applied to new data.

Table 4: Rule Accuracy

Output Node	"Inflation % change" Range (min,max)	# Rules	Train Targets	Test Targets	Net Train Accuracy Correct of 242, %	Net Test Accuracy Correct of 131, %
1	< 1.25%	4	10	9	(237) 97.93%	(127) 96.95%
2	(1.25%, 2.0%)	4	32	16	(223) 92.15%	(116) 88.55%
3	(2.0%, 3.0%)	3	72	41	(174) 71.90%	(85) 64.89%
4	(3.0%, 5.0%)	6	90	44	(182) 75.21%	(84) 64.12%
5	(5.0%, 9.0%)	4	18	12	(230) 95.04%	(117) 89.31%
6	> 9.0%	5	20	9	(240) 99.17%	(123) 93.89%

```
% true if input matches rule 3
if ( (-0.032497 <= urstn & urstn <= -0.011131) ...
   & ( (-Inf <= ursegn & ursegn <= 0.215139) ...
     | (0.253502 <= ursegn & ursegn <= 0.282514) ...
     | (0.321399 <= ursegn & ursegn <= 0.371623)) ...
) return true;
```

Figure 4: Sample Generated Rule

The rules from all six generated files were examined independently by the authors (two are subject-matter experts in econometrics & finance). Despite being executable as computer code, the rules were found to be descriptive and easy to read. Interesting patterns were found merely by examining the rule content.

An initial examination of the rules indicates the rules generated for the 6 nodes of headline inflation vary little in complexity. From Table 4 note that node 3, corresponding to headline inflation between 2 and 3 percent, has the smallest number of rules, three. These rules, as a group, also display the weakest dependence of headline inflation on movements in energy and transport prices. This is completely reasonable: the rules reflect the judgment of many observers that in the absence of large shocks the Federal Reserve generally is satisfied with inflation between 2 and 3 percent per annum. In contrast is node 6, with a monthly headline CPI inflation rate exceeding 9 percent. Inflation that rapid was observed only in one epoch: May 1979 to September 1981, when inflation was consistently greater than a 9 percent annual rate). Interpreted as a stochastic process, headline inflation has never again revisited rates that high. Energy prices also increased at an unusually rapid pace: 20 percent per annum in May 1979, peaking at a 47 percent pace in May 1980, and continuing at more than a 10 percent pace through December 1981. It would be 20 years until energy price inflation once again reached a 20 percent rate in March and June 2000—but then headline inflation was at a 3.7 percent pace. Three of the five rules focus on rapid energy price inflation; one rule includes 16 of the 17 bins for energy price inflation (!), and one rule excludes energy entirely while including only very large increases in transport costs. Considering the unusual type of shocks that generate such inflation, the rules do a reasonable job of capturing the functional linkages.

5 Conclusions & Future Work

The rules are in line with our priori expectations based on common findings in the economics literature. The U.S. Federal Reserve has accepted the core (excluding food and energy prices) chain-price index for personal consumption expenditures ("core PCE") as its policy objective. From a diagram comparing the contents of the rules to the binning of transportation and energy inputs (not included here due to space constraints), it is clear that the ranges used in the rules for energy have wide variability. The series is highly volatile and thus from a policy perspective, energy is nearly useless as a predictor of headline inflation. It appears in a number of rules, yes, but with a wide range of values. This is consistent with the economics literature; energy fluctuates so wildly that it is difficult to infer much from the fluctuations. An interesting future study would be to determine exactly how to use these rules as a means of applying suitable weights to the components of personal consumption expenditure to take account of the volatility such that food and energy prices could be included by monetary policymakers in their decision making. Food and energy prices are clearly important components of households everyday budgeting decisions and ideally should be included for policy purposes. This is the subject of ongoing research. In our future research we hope to test these rules in an out-of-sample forecasting framework to gain futher insights into their validity. A comparative analysis using discrete multivariate statistics or, alternatively, embedding these rules into a DSGE macro model would yield new insights from a policy perspective.

References

[1] J.D. Hamilton. What is an oil shock? *Journal of Econometrics*, 113:363–398, 2003.

[2] M.A. Hooker. What happened to the oil price-macroeconomy relationship? *Journal of Monetary Economics*, 38:195–213, 1996.

[3] R.B. Barsky and L. Kilian. Do we really know that oil caused the great stagflation? A monetary alternative. *NBER Macroeconomics annual*, 16:137–183, 2001.

[4] M.A. Hooker. Are oil shocks inflationary? Asymmetric and nonlinear specifications versus changes in regime. *Journal of Money, Credit and Banking*, 34:540–561, 02.

[5] J. De Gregorio, O. Landerretche, and C. Neilson. Another pass through bites the dust? Oil prices and inflation. Technical report, Central Bank of Chile Working Paper, 2007.

[6] P. Van den Noord and C. Andre. Why has core inflation remained so muted in the face of the oil shock? Technical report, Organisation for economic co-operation and development working paper 551, 2007.

[7] O. Blanchard and J. Gali. The macroeconomic effects of oil price shocks; why are the 2000s so different from the 1970s? Technical report, Massachusetts Institute of Technology Working Paper 07-21, 2008.

[8] Shiu-Sheng Chen. Oil price pass-through into inflation. *Energy Economics*, 31:126–133, 2005.

[9] Todd E. Clark and Stephen J. Terry. Time variation in the inflation passthrough of energy prices. Technical report, The Federal Reserve Bank of Kansas City Economic Research Department. Research Working Papers RWP 09–06, 2009.

[10] P. Segal. Why do oil price shocks no longer shock? Technical report, Oxford Institute for Energy Studies, Working Paper WPM, 35, 2007.

[11] L. Kilian. The economic effects of oil price shocks. *Journal of Economic Literature (forthcoming)*.

[12] Vincent A. Schmidt and Jane M. Binner. Searching for Divisia / inflation relationships with the aggregate feedforward neural network. *Advances in Econometrics*, 19:225–241, 2004.

[13] Kenneth J. Stewart and Stephen B. Reed. Consumer price index research series using current methods, 1978-98. *Monthly Labor Review*, 122:28–38, 1999.

[14] Vincent A. Schmidt. *An Aggregate Connectionist Approach for Discovering Association Rules*. PhD thesis, Wright State University, Dayton, OH, May 2002.

[15] Vincent A. Schmidt and Jane M. Binner. Analyzing Divisia rules extracted from a feedforward neural network. In *Proceedings of the 2006 International Conference on Artificial Intelligence*, Las Vegas, Nevada, 2006.

[16] Vincent A. Schmidt and Jane M. Binner. Analysing MSI rules for the USA extracted from a feedforward neural network. *Advances in Econometrics*, 24:281–294, 2009.

[17] Vincent A. Schmidt and C. L. Philip Chen. Using the aggregate feedforward neural network for rule extraction. *International Journal on Fuzzy Systems*, 4(3), 2002.