ABSTRACT

HOUCK, ADAM CHRISTOPHER. Evaluating the Rationality of the Forecasts of *The Wall Street Journal's* Panel of Economists. (Under the direction of Douglas Pearce.)

This paper will explore a methodology that will examine the difference between average and individual forecasts, concentrating on whether individual WSJ forecasters are unbiased and efficient. This result is important because the past literature has examined the accuracy of average forecasts, not individuals. In addition, a brief evaluation of Lamont's (2002) hypothesis will follow. Lamont determined that as forecasters become older and more established, in many instances deviations from the consensus forecast grew with time. The method adopted will allow for the testing of whether individual forecasts are unbiased and rational, telling more about how individuals, not averages, behave in broader contexts.

Evaluating the Rationality of the Forecasts of

The Wall Street Journal's Panel of Economists

by

Adam Christopher Houck

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Arts

ECONOMICS

Raleigh

2002

APPROVED BY:

Chair of Advisory Committee

BIOGRAPHY

Adam Christopher Houck was born on November 19, 1978 in Kettering, Ohio, where he lived until the age of four when he moved to Lynchburg, Virginia. He attended James Madison University in Harrisonburg, Virginia and decided during his junior year to pursue graduate work in economics, mainly from the suggestion of Dr. Robert Holland. He graduated in May 2001 with a Bachelor of Science degree in Financial Economics and enrolled in the Ph.D. program at North Carolina State University in the fall of 2001. He switched degree options his second semester, earning a Master of Arts in Economics in December 2002.

TABLE OF CONTENTS

LIST OF TABLES

I. Introduction

Rationality is the cornerstone of mainstream neoclassical economic theory. It underlies the principle of utility maximization among individuals and tells us that forecasters employ all relevant information to generate the best possible forecasts. To reject the assumption of rationality in economic theory would be to undermine much of the literature of the past century.

The effectiveness of policy relies strongly upon agents' expectations. Ignoring agents' beliefs and informational restrictions could result in policy that would not only be ineffective, but could prove detrimental to the economy itself. Thomas (1999) stated "In fact, inflation expectations have become more important to the FED given the diminished stability of the link between the monetary aggregates and GDP expenditures since the early 1980's." If expectations are in fact rational, a government policy based on surprising or fooling the public in any systematic way will be unsuccessful, as individual's expectations will differ by only some unforeseeable surprise, some information that was not available in the previous period.

In the financial context, expectations play an important role in determining the prices of assets. Many asset pricing models affirm that prices today are closely linked to prices expected to occur in the future. Therefore, it is crucial to look at expectations and the implications of how expectations are formed. If these expectations are rational, an asset's price is an accurate reflection of expectations of future prices. This result reaffirms one major assumption of the Efficient Markets Hypothesis; in the absence of transactions costs, risk-free arbitrage opportunities do not exist.

If agents are rational, can we assume the forecasts they construct exhibit rationality? In an incentive-driven economy, individuals make choices to maximize their respective utility. In addition, negative incentives are in place to keep firms from colluding and engaging in anticompetitive behavior. However, it is difficult to see whether incentives exist for forecasters to construct rational forecasts. If we assume forecasters act to maximize their reputation, we can believe they use all information sets available while making forecasts. Therefore, the principle of maximizing one's reputation is an implicit reward system predicated mainly on name recognition. This fact could play a major role in determining how accurate forecasts are, particularly between the Livingston and *Wall Street Journa*l (WSJ) Surveys. It is difficult to test whether the anonymity provided by the Livingston Survey affords the forecasters the luxury of making less rational predictions. The WSJ survey does identify each forecaster so they may have a stronger incentive to make the most accurate forecasts possible. Therefore, if we assume survey respondents report the same forecasts they sell on the market, their forecasts should provide a reasonable measure of their expectations.[1]

Most of the preceding literature has used measures of central tendency in the evaluation of forecast accuracy. However, using an average or median value can hide the variation in individual forecaster's results and is not appropriate in testing for rationality, as these statistics can be misleading and can easily hide the irrationality and bias contained within individual forecasts. In the simplest case, every individual's forecast can be biased, irrational, and wrong while the average remains unbiased, rational, and correct. Bonham and Cohen warn "Conclusions about individual behavior based on consensus regressions are likely to be misleading even in the unlikely event that the consensus (average) parameters are consistently estimated" (2001) Most surveys look at how accurate average forecasts are; however, determining whether each individual forecaster's predictions are unbiased and rational can tell much more about the underlying data.

[1] For more elaboration, see Keane and Runkle (1990)

Using the WSJ panel data there are numerous hypotheses that can be tested. This paper will explore a methodology that will examine the difference between average and individual forecasts, concentrating on whether individual WSJ forecasters are unbiased and efficient. This result is important because the preceding literature has looked at the validity of average forecasts, not individuals. In addition, a brief evaluation of Lamont's (2002) hypothesis will follow. Lamont determined that as forecasters become older and more established, in many instances deviations from the consensus forecast grew with time. The method adopted will allow for the testing of whether individual forecasts are unbiased and rational, telling more about how individuals, not averages, behave in a broader context.[2]

[2] For an exhaustive analysis of the evaluation of WSJ forecasters, see Eisenbeis, Waggoner, and Zha (2002)

II. Literature Review

Survey measures of expected inflation have been collected and compiled by three main bodies: the Livingston Survey (Livingston), the Survey of Professional Forecasters (SPF), and the Institute of Social Research Michigan Survey of Households (Michigan). The Livingston Survey was created in 1946 by Joseph Livingston, a regular columnist in the *Philadelphia Inquirer,* for journalistic, not scientific purposes (Thomas, 1999). The survey initially asked forecasters to predict twelve key macroeconomic variables to capture the overall economy's performance. Since 1950, between 37 and 64 survey respondents have been included with a myriad of public, private, and academic affiliations. The survey methods and number of forecasted variables have evolved from the survey's inception, but the underlying theme has not.[3] The Philadelphia FED collects individual forecasts from respondents and publishes mean and median forecasts in June and December of each year.[4]

The Survey of Professional Forecasters (SPF) was first conducted in 1968 to forecast the GNP deflator, among other variables; inflation forecasts were not officially instituted until the third quarter of 1981. Its forecasters come predominantly from the business sector, including chief economists at several Fortune 500 companies. Similar to Livingston, SPF collects individual forecasts and publishes mean and median forecasts.[5] Overall, the survey is regarded as extremely accurate in predicting key variables in the economy among constituents. (Thomas, 1999)

---

[3] For more information regarding the improvement and evolution of the Livingston Survey over time, see Thomas, 1999.

[4] The Federal Reserve Bank of Philadelphia assumed responsibility of conducting the survey after Livingston's death on Christmas Day 1989.

[5]The anonymous forecasts are available via the Philadelphia FED <http://www.phil.frb.org>

The Michigan survey collects a random monthly sample via telephone of 500 households, not professionals like both Livingston and SPF. The questions asked have evolved over time; from 1948 until the second quarter of 1966, households were simply asked whether prices would increase, decrease, or remain the same. From the second quarter of 1966 until the third quarter of 1977, households were asked within what specified range prices would increase, decrease, or remain the same. Now, households are asked to provide actual values of expected inflation on "things they buy." (Thomas, 1999) The Michigan-mean and Michigan-median forecasts are both published and available via Michigan's website and are strictly anonymous as are the other survey measures.[6]

Historically, testing for rationality in survey measures has involved two necessary conditions of rational expectations. First, forecasts should be unbiased; that is, respondents are correct on average. Second, forecasts are efficient in the sense that respondents employ all relevant information in constructing expectations. The test for bias usually involves running the following regression:

$$A_t = \alpha + \beta P_{t-1} + u_t \tag{1}$$

where A_t is the actual value of the variable in question at time t, P_{t-1} is the forecasted value at time $(t - 1)$, and u_t is the disturbance term. Forecasts are unbiased if $\alpha = 0$ and $\beta = 1$. Testing for efficiency involves running a similar regression:

$$A_t - P_{t-1} = X_{t-1}\theta + u_t \tag{2}$$

where X_{t-1} = information known at time $t - 1$ and θ = vector of coefficients. The null hypothesis of efficiency requires that all $\theta = 0$.

[6]For the anonymous individual Michigan surveys, see <http://www.athena.sca.isr.umich.edu>

The literature has generally used aggregate measures of expectations in attempted tests of rationality. However, Zarnowitz and Braun (1993, p. 66) warn:

> **"The distributions of the [forecast] error statistics show that there is much dispersion across the forecasts…Forecasters differ in many respects and so do their products. The idea that a close 'consensus' persists, i.e., the current matched forecasts are generally all alike, is a popular fiction. The differentiation of the forecasts usually involves much more than the existence of just a few outliers."**

It is therefore important to examine whether the term "consensus" applies to cross-sectional forecast means. If forecasts meet the conditions of unbiasedness and efficiency, they should differ only by the heterogeneous private information sets specific to individual forecasters, which could create potentially large differences in forecasts across the sample. In this context, the notion of "consensus" is misleading and ignores that forecasts can differ greatly, mostly due to forecaster-specific information sets. Therefore, in both tests for unbiasedness and efficiency, cross-sectional means can mask individual deviations from rationality. One must therefore be careful when drawing inferences regarding rationality. If a cross-sectional mean forecast is rational, it does not necessarily follow that all forecasts in the sample are rational as well. Conversely, if all forecasts in a sample are rational, it does follow that the cross-sectional mean forecast will exhibit rationality.

In addition to testing for unbiasedness and efficiency in individual forecasts, one can look at how individuals perform over time in relation to the cross-sectional mean forecast. Lamont (2002) has shown that as forecasters become older, their forecast deviations (with respect to the cross-sectional mean) may increase as well; as a result, forecast accuracy grows worse as forecasters become more experienced. Lamont attributes this behavior mainly to reputation; individuals construct more radical forecasts to generate attention and to gain credibility for themselves and their respective firms in the unlikely event that the forecast turns out to be accurate. A test of his theory appears in Section IV.

III. The Data

The *Wall Street Journal* Forecasting Survey began in 1986 by mailing questionnaires to a select number of economists with varying affiliations in the public, private, and academic sectors. Since that time, WSJ has asked forecasters to provide six- and twelve-month forecasts for specific financial and economic variables as well as answers to more qualitative questions. Quantitative predictions are published the first week of January and July each year. In addition to the published forecasts, many questions the respondents are asked never reach print. For example, the June 2002 questionnaire mailed to forecasters asked for qualitative responses regarding business cycles, the Dow Jones Industrial Index, and approximate growth rates of corporate profits. Personal finance questions were even asked of respondents, namely changes in personal savings and portfolio composition. Finally, forecasters were asked to reply to various questions with short (five words or less) answers regarding issues such as FED policy, the war on terrorism, and accounting scandals. How long individual forecasters remain on the panel varies greatly across the sample. Several forecasters exist on the panel with thirty-two observations; conversely, there are individuals who have only one or two observations.

Before testing the rationality of these forecasters, two additional issues need to be explored. For purposes here, real-time, not revised data will be used for two reasons. First, revised numbers were not available to respondents at the time of forecast and would not have entered into their information sets. Secondly, using revised actual values could alter the results of hypothesis tests for bias and efficiency; this outcome can also be tested. In addition, the WSJ does not provide their survey panel with any past information; under the null hypothesis of efficiency, forecasters will use all available information in constructing forecasts. Each forecaster's information comes predominantly from two areas: public information sets (available to all forecasters) and private information sets (specific to individual forecasters). Theoretically, public information sets should not vary greatly across the sample and should be homogeneous. Conversely, private information sets could vary greatly based on the selection of research of forecasters. Thus, the more heterogeneous the private information sets, the more heterogeneous the individual forecasts. Again, the term “consensus forecast” appears to be misleading and inappropriate.

IV. Tests of Rationality

Using equations (1) and (2) in section II, the rationality of six- and twelve-month forecasts of the annualized percentage change in CPI, the levels of the DLR vs. YEN exchange rate, and the six-month forecast of the annualized percentage change in GDP will be tested. For statistical purposes, only individuals with 20 or more observations will be used; as a result, 26 individual forecasters will be tested. In accord with the preceding literature, the consensus forecasts will be tested first to determine whether they adhere to the null hypothesis of unbiasedness. The results appear in Table 1.

Table 1: Results of Consensus Forecast Unbiasedness Tests							
Variable	Intercept	Slope	D-W	Autocorrelation	R square	F	n
	(s.e.)	(s.e.)					
CPI-6 month	-0.144	1.01	1.587	no	0.446	0.17	32
	0.69	0.206					
CPI-12 month	-0.639	1.08	2.709	no	0.526	2.18	26
	0.773	0.209					
YEN-6 month	14.28	0.873	1.261	yes	0.59	0.49	26
	18.06	0.149					
YEN-12 month	28.07	1.05	1.436	no	0.425	1.04	20
	25.47	0.292					
GDP-6 month	0.602	1.05	1.877	no	0.309	4.3*	31
	0.712	0.292					
note: Autocorrelation refers to a test at 95% of positive autocorrelation in residuals							
note: F statistic represents a joint test that both the intercept = 0 and slope = 1; a star denotes the value is not significant at 95%.							

For the five forecasted variables, the hypothesis that the survey predictions are unbiased can be rejected only for the six-month GDP, as evidenced from the *F*-statistics in Table 1. The R^2 's range from .309 (GDP-six month) to .59 (YEN-six month) over the sample. The only glaring result presented in the above table is the presence of positive autocorrelation in the six-month YEN forecast at the 95% level. One might therefore postulate that the exchange rate series was autocorrelated due to nonstationarity in the sample. An augmented Dickey-Fuller test was run on the series to determine the validity of this claim and the results appear in Table 2.

Table 2: Augmented Dickey-Fuller Tests For YEN series			
Type	Lags	Rho	Pr < Rho
Single Mean	0	-0.2749	0.9363
	1	1.7431	0.9946
	2	-1.5774	0.816
Trend	0	-7.3744	0.5873
	1	-3.7867	0.8862
	2	-12.1991	0.2264
note: ADF tests that a unit root is present (*P* =1)			

The *P*-values demonstrate that for all lags shown in Single Mean and Trend tests, the absence of a unit root cannot be rejected. Therefore, by showing the YEN series was nonstationary over the sample, it might better explain why autocorrelation was present in the errors of the consensus forecasts for the YEN.

Looking at how individuals performed, for the six-month CPI forecasts, 20 of the 26 forecasters constructed unbiased forecasts as evidenced by the joint *F*-statistics at the 95% significance level (see Appendix A). However, the Durbin-Watson statistics showed that 1 forecaster, Alan Reynolds, had autocorrelation present in his forecasts; 2 individuals yielded inconclusive results, while the vast majority (23) had no autocorrelation present in forecasts. The highest R^2 (.70) was attained by Carol Leisenring; she had no autocorrelation present in her forecast. The lowest R^2 (.092) came from the aforementioned Alan Reynolds, who had autocorrelation present in his forecasts.

The one-year CPI forecasts yielded slightly different results. Fewer individuals (16) constructed unbiased forecasts at the 95% level (see Appendix B). Again, Alan Reynolds had autocorrelation present in his forecasts and yielded the lowest R^2 (.002) for the sample. In addition to Reynolds, A. Gary Shilling also failed the test for autocorrelation and his R^2 value was only marginally better (.003) than Reynolds. Carol Leisenring once again yielded the highest R^2 value for the sample (.694) and had no autocorrelation present in her forecasts. Comparing the results of the six-month and one-year CPI forecasts, it follows that the greater the time horizon of the forecast, the greater the spread of R^2 values over the sample, which makes intuitive sense. The longer time horizon could yield greater uncertainty for individuals and their six-month forecasts might have a tighter spread in comparison with their comparable one-year forecasts.

The six-month YEN forecasts yielded results that were similar to the 6-month CPI forecasts in some respects and very different in others. Of the 26 individuals, 20 had significant F values at the 95% level (see Appendix C). However, 10 forecasters had autocorrelation present in their forecasts, perhaps a reflection of the nonstationary problem with the YEN previously aforementioned. Ten forecasters yielded inconclusive Durbin-Watson Statistics while only 8 individuals had no autocorrelation present in their forecasts. The R^2 's ranged from a low of .356 from James Smith, who had an inconclusive D-W statistic, to a high of .706 from Donald Straszheim, who had no autocorrelation present in his forecasts.

The one-year YEN forecasts again yielded slightly different results from the 6-month. At the 95% level, 19 of the 26 forecasters had significant F values, one fewer than the 6-month (see Appendix D). Thirteen forecasters had no autocorrelation present in their forecasts, while 5 had autocorrelation present and 8 individuals yielded inconclusive results. The R^2 's ranged from a low of .20 from Gail Fosler to a high of .61 from Robert Dederick. Both Fosler and Dederick yielded inconclusive D-W statistics.

The six-month GDP forecasts yielded the worst results of the five variables tested. Only 13 of the 26 forecasters had significant F values (see Appendix E). Five forecasters had autocorrelation present in their results, while 15 did not and 6 were inconclusive. The $\bar{R}^2$'s had the largest range of any variable, from a low of .000 from Edward Yardeni to a high of .468 from Philip Braverman. Braverman had no autocorrelation present in his results while Yardeni failed the test of first-order autocorrelation. Overall, it seems GDP was the most difficult variable to forecast, evidenced by low $\bar{R}^2$'s and insignificant F values.

The second necessary condition for rationality is the existence of efficiency in estimates. Under the null hypothesis of rationality, a regression of the forecast errors on known information at the time of forecast should not yield any statistically significant coefficients, as evidenced by equation (2). Although the information set may be difficult to specify completely, leaving out relevant variables from X_{t-1} does not invalidate this test. This follows from the fact that, under the null hypothesis, any variable on the right-hand side of (2) that is known at time $t - 1$ has a zero coefficient. (Pearce, 1984)

The information set assumed to be known at the time of forecast includes the six-month annualized percentage changes in CPI and GDP, measured from the previous survey date to the current survey date. In addition, the value of the YEN exchange rate on the last day of the month before surveys were published will serve as the last known information respondents could have known about the exchange rate; this is convenient because the WSJ survey asks for the actual value of the exchange rate, not an annualized percentage change in the exchange rate. These variables have been chosen because respondents are asked for their forecasts of these series and are very likely to have this information. The information set contains values of variables that were available to respondents at the time of forecast, not revised data that would have been unavailable at the time of forecast.

As before, the consensus forecasts will be tested first to determine whether they adhere to the null hypothesis of efficiency. The results appear in Table 3.

Table 3: Efficiency Tests for Consensus Forecasts						
Variable	*Intercept*	*GDP*	*CPI*	*YEN*	*F*	*n*
	(s.e.)	*(s.e.)*	*(s.e.)*	*(s.e.)*		
CPI-6 month	0.228	0.009	0.173	-0.007	0.86	32
	1.1	0.106	0.149	0.006		
CPI-12 month	0.858	-0.141	-0.109	-0.003	0.66	26
	1.08	0.11	0.151	0.006		
YEN-6 month	25.85	0.21	3.22	-0.317	2.23	26
	17.56	1.28	1.99	0.134		
YEN-12 month	0.098	2.51	4.7	-0.207	1.13	20
	24.24	1.94	3.16	0.201		
GDP-6 month	0.665	0.296	-0.305	0.001	3.24*	31
	1.453	0.142	0.195	0.008		
note: F statistic represents a test that all slope coefficients = 0 and has 3 and (n - 4) degrees of freedom; a star represents the statistic is not significant at 95%						

Table 3 presents estimates of equation (2) for the consensus forecasts along with the *F*-statistics for the null hypothesis that the forecast errors are not linearly related to known changes in the information set X_{t-1}. These results indicate that the null hypothesis for efficiency cannot be rejected at the 95 percent level except for the six-month GDP. One must therefore conclude that the consensus forecasts, with the exception of GDP, displayed characteristics in accord with rational expectations.

Once again turning attention to individual forecasters, efficiency tests were run on the selected 26 survey respondents. Six-month surveys were tested for efficiency because they contained the greatest number of observations. For the six-month CPI forecasts, twenty-five of the twenty-six forecasters passed the test for efficiency at the 95 percent level, with only Alan Reynolds failing with an F-statistic of 4.54 (see Appendix F). For the six-month YEN forecasts, twenty-two of the twenty-six forecasters passed the efficiency tests with Mickey Levy, Saul Hymans, A. Gary Shilling, and Edward Yardeni not constructing efficient estimates at the 95 percent level (see Appendix G). The six-month GDP forecasts again yielded the worst results, with only six of the twenty-six forecasters producing efficient estimates. These results (which appear in Appendix H) demonstrate the difficulty survey respondents had in efficiently using the known information in making forecasts.

The final empirical test was run to determine whether Owen Lamont's assertion that as forecasters become older, they construct more radical forecasts was evident for this sample. For purposes here, the five individuals with the largest number of observations for each variable were selected and a regression was run of the following form:

$$\left| P_{i,t-1} - P_{c,t-1} \right| = \alpha + \beta t + u_t \tag{3}$$

Where $P_{i,t\text{-}1}$ is an individual's prediction at time $t - 1$, $P_{c,t\text{-}1}$ is the consensus forecast prediction at time $t - 1$, t is a time trend, and ε_t is a disturbance term. A statistically positive coefficient on t would show that for the sample space, the absolute value of the deviation from the consensus prediction for any forecaster would increase over time. Therefore, respondents with positive slope coefficients would adhere to Lamont's hypothesis: that as forecasters becomes older, the construct more radical forecasts with respect to the consensus.

After running the above regressions, no statistically significant positive slope coefficients were found in either the six-month or one-year CPI forecasts (see Appendix I). David Resler was the only forecaster who yielded a statistically significant positive coefficient for the six-month YEN forecasts, but neither he nor any other forecaster yielded significant results for the YEN one-year. For the six-month GDP forecasts, only Mickey Levy yielded a statistically positive slope coefficient. From these results, it would appear that no forecaster deviated from the consensus forecast in any systematic way, or that the absolute value of the difference between their individual forecasts and the consensus grew over time. The difference in results from Lamont's experiment can come from two predominant areas. First, Lamont used a data set comprised of the *BusinessWeek* annual year-end outlook issue from 1971 to 1992, which could yield different results. Also, Lamont's method of estimation varied slightly from the method used in this experiment; he excluded the individual's forecast being tested from the consensus used. Even though this difference is negligible, this simple variation in method could lead to the different results found in this experiment.

V. Conclusions

This paper employed the individual responses from the *Wall Street Journal's* Survey of Economists to explore the potential differences between individual and average (consensus) forecasts. Empirical tests were performed to determine whether consensus forecasts could lead to false inferences regarding the behavior of individuals, and whether individuals, in fact, exhibit rationality. The evidence presented has shown that even in the event that a consensus forecast behaves in accordance with rationality, the individuals who comprise that consensus may not. In addition, by employing the techniques used by Owen Lamont it has been shown that for the data sample used, forecasters did not necessarily construct more radical forecasts as their tenure on the survey increased.

How well respondents performed seemed to vary from variable to variable; more individuals were able to construct rational forecasts for the CPI than GDP. In summary, the evidence has shown the sample of twenty-six forecasters examined here exhibited characteristics that both adhered to and deviated from rationality. By examining the behavior of these individuals and not consensus estimates, one might gain better insight as to how individuals actually form expectations.

VI. References

Bonham, C., and Cohen, R. (2001), "To Aggregate, Pool, or Neither: Testing the Rational-Expectations Hypothesis Using Survey Data," *Journal of Business and Economic Statistics,* 19, 278-291.

Champ, B., and Freeman, S. (2001), *Modeling Monetary Economies,* Cambridge, U.K.

Eisenbeis, R., Waggoner, D., and Zha, T. (2002), "Evaluating Wall Street Journal Survey Forecasters: A Multivariate Approach," Working Paper 2002-8, Federal Reserve Bank of Atlanta.

Mehra, Y., "Survey Measures of Expected Inflation: Revisiting the Issues of Predictive Content and Rationality," *Economic Quarterly* (Summer), 88, 17-36.

Pearce, D.K., (1984), "An Empirical Analysis of Expected Stock Price Movements," *Journal of Money, Credit and Banking*, 16, 317-327.

Lamont, O., (2002), "Macroeconomic Forecasts and Microeconomic Forecasters," *Journal of Economic Behavior and Organization,* 48, 265-280.

Thomas, L.B., (1999), "Survey Measures of Expected U.S. Inflation," *Journal of Economic Perspectives* (Fall), 13, 125-44.

Zarnowitz, V., and Braun, P. (1993), "Twenty-Two Years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisons of Forecasting Performance," in *Business Cycles, Indicators, and Forecasting,* eds. J.H. Stock and M.W. Watson, Chicago: University of Chicago Press, pp. 11-93.

Federal Reserve Bank of Philadelphia <http://www.phil.frb.org> Accessed August 12, 2002.

Institute of Social Research Michigan Survey of Households <http://www.athena.sca.isr.umich.edu> Accessed August 13, 2002.

Appendix A: CPI 6-month Forecasts

Forecaster	*Intercept (s.e.)*	*Slope (s.e.)*	*D-W*	*Autocorrelation*	*R square*	*F*	*n*
Resler	0.547 0.754	0.832 0.234	1.627	no	0.297	0.27	32
Hoffman	0.368 0.595	0.836 0.166	1.83	no	0.502	1.17	27
Levy	0.81 0.804	0.735 0.245	1.433	inconclusive	0.231	0.6	32
Wyss	0.197 0.897	0.91 0.269	1.8	no	0.305	0.5	28
Harris	0.784 0.754	0.776 0.235	1.717	no	0.28	0.56	30
Sinai	0.776 0.676	0.746 0.205	1.397	inconclusive	0.307	0.51	32
Ratajczak	-0.376 0.66	1.15 0.2	2.035	no	0.55	0.65	29
Berson	-0.49 0.732	1.16 0.239	2.005	no	0.52	0.23	24
Daane	-0.08 0.615	0.91 0.166	2.096	no	0.55	3.21	27
Hymans	0.704 0.445	0.73 0.122	1.825	no	0.55	3.54*	31
Rippe	0.168 1.03	0.953 0.344	1.881	no	0.258	0.02	24
Hyman	1.56 0.586	0.59 0.206	1.231	yes	0.215	4.14*	32
Smith	1.64 0.685	0.608 0.245	1.903	no	0.192	5.31*	28
Fosler	1.38 0.88	0.42 0.259	2.38	no	0.116	6.95*	22
Shilling	1.506 0.589	0.644 0.235	1.688	no	0.21	6.97*	30
Moskowitz	1.49 0.684	0.525 0.2	1.975	no	0.196	0.99	30
Yardeni	1.55 0.663	0.63 0.237	1.657	no	0.213	5.02*	28
Platt	0.391 0.61	0.854 0.174	2.205	no	0.547	0.48	22
Reynolds	1.33 1.24	0.625 0.393	1.165	yes	0.092	0.71	27
Braverman	0.098 0.587	1.01 0.185	2.134	no	0.546	0.34	27
Boltz	-0.893 0.754	1.15 0.199	2.118	no	0.602	2.37	24
Leisenring	0.108 0.502	0.978 0.141	2.267	no	0.7	0.04	22
Straszheim	0.032 0.699	0.958 0.188	2.268	no	0.576	0.27	21
Kellner	1.53 0.88	0.578 0.258	1.515	no	0.193	1.52	23
Dederick	-0.326 0.837	1.09 0.224	2.546	no	0.57	0.1	20
Robertson	-0.214 0.975	1.01 0.256	1.937	no	0.463	0.36	20

note: Autocorrelation column represents a test of positive autocorrelation on residuals at 95% using appropriate Durbin-Watson statistic and number of observations

note: F-Statistic represents a joint test that both the intercept = 0 and the slope = 1; a star denotes the value is not significant at 95%. The statistic has 2 and (n - 2) degrees of freedom.

Appendix B: CPI 1 Year Forecasts

Forecaster	*Intercept (s.e.)*	*Slope (s.e.)*	*D-W*	*Autocorrelation*	*R square*	*F*	*n*
Resler	0.394	0.859	2	no	0.3	0.2	26
	0.924	0.271					
Hoffman	-0.61	1.05	2.632	no	0.59	2.61	21
	0.761	0.202					
Levy	1.44	0.532	1.65	no	0.15	1.84	26
	0.924	0.26					
Wyss	-0.76	1.15	2.214	no	0.55	1.19	24
	0.79	0.22					
Harris	0.077	0.956	1.358	inconclusive	0.42	0.09	24
	0.84	0.239					
Sinai	0.872	0.706	2.117	no	0.204	0.61	24
	1.02	0.298					
Ratajczak	-1.29	1.21	2.644	no	0.69	6.86*	24
	0.67	0.172					
Berson	-1.27	1.25	2.616	no	0.63	3.86*	18
	0.822	0.24					
Daane	-0.644	0.948	2.436	no	0.61	11.91*	21
	0.72	0.173					
Hymans	0.526	0.737	2.686	no	0.6	6.72*	25
	0.505	0.126					
Rippe	-0.127	0.924	2.327	no	0.21	1.17	18
	1.5	0.449					
Hyman	2.26	0.339	1.531	no	0.076	4.7*	24
	0.75	0.251					
Smith	0.999	0.838	2.612	no	0.48	4.79*	22
	0.557	0.196					
Fosler	0.264	0.598	2.094	no	0.26	54.02*	16
	1.04	0.267					
Shilling	2.87	0.126	1.092	yes	0.003	7.99*	24
	1.15	0.504					
Moskowitz	1.31	0.543	1.849	no	0.186	2.48	25
	0.841	0.237					
Yardeni	0.825	0.919	2.3	no	0.447	5.89*	24
	0.602	0.218					
Platt	-0.665	1.1	2.4	no	0.622	1.36	19
	0.758	0.209					
Reynolds	3.06	0.085	0.963	yes	0.002	2.17	24
	1.65	0.463					
Braverman	0.882	0.775	2.375	no	0.332	0.97	26
	0.715	0.224					
Boltz	0.321	0.766	2.048	no	0.321	4.72*	24
	0.967	0.237					
Leisenring	-0.165	1.006	1.995	no	0.694	0.48	22
	0.537	0.149					
Straszheim	0.547	0.792	2.377	no	0.473	1.38	21
	0.727	0.192					
Kellner	2.08	0.416	1.694	no	0.135	3.2	22
	0.825	0.235					
Dederick	-0.256	0.992	3.028	no	0.505	1.21	19
	0.956	0.238					
Robertson	0.937	0.677	2	no	0.29	1.82	19
	1.06	0.259					

note: Autocorrelation column represents a test of positive autocorrelation on residuals at 95% using appropriate Durbin-Watson statistic and number of observations

note: F-Statistic represents a joint test that both the intercept = 0 and the slope = 1; a star denotes the value is not significant at 95%. The statistic has 2 and (n - 2) degrees of freedom.

Appendix C: DLR vs. YEN 6-month Forecasts

Forecaster	*Intercept (s.e.)*	*Slope (s.e.)*	*D-W*	*Autocorrelation*	*R square*	*F*	*n*
Resler	12.5 17.42	0.882 0.143	1.453	no	0.624	0.72	25
Hoffman	29.98 17.27	0.75 0.14	1.205	yes	0.56	1.69	24
Levy	33.48 14.84	0.696 0.119	1.015	yes	0.599	5.11*	25
Wyss	33.48 24.06	0.732 0.201	1.184	yes	0.398	1.06	22
Harris	19.16 21.2	0.816 0.171	1.413	inconclusive	0.496	1.59	25
Sinai	27.06 23.09	0.774 0.192	1.085	yes	0.43	0.69	24
Ratajczak	23.17 18.37	0.83 0.156	1.471	no	0.56	1.49	24
Berson	18.78 15.19	0.821 0.124	1.522	no	0.67	2.09	23
Daane	33.08 23.15	0.734 0.198	1.097	yes	0.38	1.56	25
Hymans	16.01 17.91	0.873 0.149	1.443	no	0.629	0.44	22
Rippe	5.64 20.58	0.956 0.174	1.523	no	0.59	0.06	23
Hyman	32.58 14.52	0.684 0.114	1.293	inconclusive	0.63	9.38*	23
Smith	40.51 23.03	0.602 0.173	1.281	inconclusive	0.356	12.46*	24
Fosler	45.17 15.05	0.629 0.134	1.382	inconclusive	0.538	5.42*	21
Shilling	43.13 16.53	0.572 0.125	1.399	inconclusive	0.5	19.43*	23
Moskowitz	37.63 16.98	0.671 0.137	1.198	yes	0.52	3.39	24
Yardeni	40.43 15.27	0.632 0.12	1.141	yes	0.556	7.55*	24
Platt	2.55 21.49	0.98 0.176	1.158	yes	0.618	0.01	21
Reynolds	24.96 15.27	0.8 0.125	1.387	inconclusive	0.673	1.35	22
Braverman	19.2 19.29	0.851 0.159	1.551	no	0.6	0.57	21
Boltz	20.14 21.14	0.845 0.177	1.238	inconclusive	0.586	0.6	18
Leisenring	8.731 21.16	0.928 0.174	1.225	inconclusive	0.627	0.09	19
Straszheim	2.89 19.73	0.989 0.165	1.452	no	0.706	0.21	17
Kellner	0.717 31.68	1.03 0.271	1.372	no	0.509	0.84	16
Dederick	12.8 27.89	0.911 0.234	1.071	yes	0.519	0.28	16
Robertson	9.232 26.34	0.93 0.215	0.719	yes	0.609	0.08	14

note: Autocorrelation column represents a test of positive autocorrelation on residuals at 95% using appropriate Durbin-Watson statistic and number of observations

note: F-Statistic represents a joint test that both the intercept = 0 and the slope = 1; a star denotes the value is not significant at 95%. The statistic has 2 and (n - 2) degrees of freedom.

Appendix D: DLR vs. YEN 1-year Forecasts

Forecaster	*Intercept (s.e.)*	*Slope (s.e.)*	*D-W*	*Autocorrelation*	*R square*	*F*	*n*
Resler	39.84	0.648	1.252	inconclusive	0.412	2.86	21
	22.18	0.178					
Hoffman	39.36	0.66	1.781	no	0.501	3.2	20
	19.26	0.154					
Levy	44.02	0.602	1.752	no	0.556	8.3*	21
	15.79	0.124					
Wyss	37.48	0.717	1.184	inconclusive	0.27	1.41	19
	33.2	0.286					
Harris	43.7	0.619	1.233	inconclusive	0.32	2.2	20
	26.6	0.214					
Sinai	26.29	0.746	1.424	no	0.344	2.04	20
	30.37	0.243					
Ratajczak	38.42	0.698	1.619	no	0.403	1.76	21
	22.96	0.195					
Berson	39.27	0.643	1.573	no	0.294	2.13	19
	29.77	0.242					
Daane	46.24	0.637	1.201	yes	0.256	1.88	21
	29.08	0.249					
Hymans	32.96	0.729	1.534	no	0.437	0.93	19
	24.22	0.201					
Rippe	26.21	0.802	1.563	no	0.37	1.05	19
	29.15	0.253					
Hyman	47.98	0.563	1.774	no	0.518	12.35*	19
	17.39	0.131					
Smith	56.77	0.461	0.843	yes	0.304	20.11*	20
	22.94	0.164					
Fosler	79.94	0.316	1.28	inconclusive	0.2	9.31*	17
	18.56	0.163					
Shilling	79.74	0.318	1.521	no	0.326	60.12*	19
	16.45	0.111					
Moskowitz	50.55	0.552	1.593	no	0.446	6.88*	20
	18.55	0.145					
Yardeni	47.3	0.56	1.68	no	0.431	11.65*	20
	20.02	0.151					
Platt	34.48	0.72	1.707	no	0.349	0.67	18
	29.82	0.246					
Reynolds	38.96	0.665	1.212	inconclusive	0.354	1.54	19
	27.21	0.218					
Braverman	44.72	18.29	2.074	no	0.49	3.45*	20
	0.62	0.148					
Boltz	58.3	0.517	0.995	yes	0.23	2.13	18
	28.25	0.236					
Leisenring	23.92	0.787	1.327	inconclusive	0.58	1.2	19
	20.13	0.163					
Straszheim	17.75	0.863	1.26	inconclusive	0.5	0.33	17
	26.24	0.223					
Kellner	17.14	0.903	0.841	yes	0.53	2.35	16
	26.41	0.229					
Dederick	3.41	0.984	1.117	inconclusive	0.61	0.14	16
	24.63	0.209					
Robertson	30.16	0.748	0.996	yes	0.517	0.74	14
	25.65	0.209					

note: Autocorrelation column represents a test of positive autocorrelation on residuals at 95% using appropriate Durbin-Watson statistic and number of observations

note: F-Statistic represents a joint test that both the intercept = 0 and the slope = 1; a star denotes the value is not significant at 95%. The statistic has 2 and (n - 2) degrees of freedom.

Appendix E: GDP 6-month Forecasts

Forecaster	*Intercept (s.e.)*	*Slope (s.e.)*	*D-W*	*Autocorrelation*	*R square*	*F*	*n*
Resler	1.35 0.608	0.813 0.274	1.724	no	0.227	6.46*	32
Hoffman	1.42 0.628	0.757 0.28	2.092	no	0.226	4.65*	27
Levy	0.948 0.686	0.877 0.274	1.819	no	0.254	3.05	32
Wyss	1.774 0.784	0.454 0.332	1.741	no	0.069	3*	27
Harris	0.889 0.899	0.889 0.361	1.38	inconclusive	0.18	2.17	30
Sinai	0.939 0.488	0.924 0.199	1.95	no	0.43	4.9*	30
Ratajczak	1.28 0.663	0.756 0.253	1.47	inconclusive	0.25	3.5*	29
Berson	0.3 1.21	1 0.474	1.07	yes	0.169	0.34	24
Daane	0.477 0.586	1.03 0.228	1.303	yes	0.45	2.11	27
Hymans	3.68 1.12	-0.297 0.432	1.551	no	0.016	5.46*	31
Rippe	-0.19 1.322	1.29 0.559	1.312	inconclusive	0.194	0.96	24
Hyman	1.22 0.536	0.872 0.231	2.108	no	0.329	6.51*	31
Smith	1.7 1.23	0.397 0.399	1.287	yes	0.036	1.16	28
Fosler	1.758 1.35	0.376 0.408	1.344	inconclusive	0.04	1.3	22
Shilling	2.87 0.297	0.381 0.13	1.806	no	0.233	55*	30
Moskowitz	1.526 0.568	0.693 0.221	1.425	inconclusive	0.266	5.62*	29
Yardeni	3.06 0.6	0 0.235	1.242	yes	0	12.97*	28
Platt	0.796 0.766	0.939 0.297	1.399	inconclusive	0.334	2	22
Reynolds	2.21 0.694	0.368 0.263	1.165	yes	0.072	5.44*	27
Braverman	1.661 0.366	0.85 0.181	2.03	no	0.468	17.76*	27
Boltz	0.145 0.847	1.08 0.327	1.6	no	0.33	0.69	24
Leisenring	0.473 0.895	1.06 0.377	1.494	no	0.285	1.69	22
Straszheim	1.77 0.654	0.416 0.272	1.61	no	0.11	3.68*	21
Kellner	2.093 0.554	0.346 0.243	1.627	no	0.089	7.2*	23
Dederick	1.43 0.804	0.493 0.31	1.598	no	0.123	1.6	20
Robertson	0.643 0.793	0.852 0.324	1.683	no	0.278	0.53	20

note: Autocorrelation column represents a test of positive autocorrelation on residuals at 95% using appropriate Durbin-Watson statistic and number of observations.

note: F-Statistic represents a joint test that both the intercept = 0 and the slope = 1; a star denotes the value is not significant at 95%. The statistic has 2 and (n - 2) degrees of freedom.

Appendix F: Efficiency Tests for 6-month CPI Forecasts

Forecaster	*Intercept (s.e.)*	*GDP (s.e.)*	*CPI (s.e.)*	*YEN (s.e.)*	*F F(no intercept)*	*n*
Resler	0.37 (1.202)	-0.081 (0.116)	0.267 (0.163)	-0.0075 (0.007)	1.68 (1.71)	32
Hoffman	-0.347 (1.399)	0.013 (0.109)	-0.029 (0.184)	0.0017 (0.0114)	0.02 (0.42)	27
Levy	-0.797 (1.249)	0.1025 (0.1205)	0.39 (0.169)	-0.0063 (0.0073)	2.05 (1.97)	32
Wyss	-0.656 (1.063)	0.087 (0.112)	0.321 (0.149)	-0.0056 (0.006)	1.76 (1.8)	27
Harris	-2.19 (1.36)	0.059 (0.117)	0.386 (0.163)	0.007 (0.008)	2.11 (1.25)	30
Sinai	0.695 (1.22)	0.0078 (0.119)	0.168 (0.166)	-0.009 (0.007)	1.01 (0.93)	30
Ratajczak	-2.04 (0.972)	0.077 (0.086)	0.237 (0.121)	0.009 (0.006)	2.12 (0.82)	29
Berson	-3.3 (1.51)	0.229 (0.109)	0.348 (0.174)	0.013 (0.011)	2.18 (0.49)	24
Daane	-1.1 (1.47)	0.086 (0.106)	0.019 (0.171)	0.003 (0.012)	0.24 (2)	27
Hymans	1.707 (1.16)	-0.026 (0.101)	-0.078 (0.141)	-0.013 (0.007)	1.16 (1.06)	31
Rippe	-3.99 (1.87)	0.241 (0.136)	0.409 (0.216)	0.017 (0.013)	1.99 (0.4)	24
Hyman	0.099 (1.29)	-0.028 (0.126)	0.073 (0.178)	0.001 (0.007)	0.13 (1.83)	31
Smith	-1.94 (1.54)	0.084 (0.136)	0.197 (0.194)	0.013 (0.009)	0.99 (2.94)	28
Fosler	-1.48 (1.96)	0.24 (0.134)	0.186 (0.221)	-0.003 (0.014)	1.27 (3.75*)	22
Shilling	0.62 (1.39)	-0.11 (0.134)	0.112 (0.199)	0.0001 (0.008)	0.53 (4.12*)	30
Moskowitz	1.23 (1.21)	-0.019 (0.119)	0.131 (0.16)	-0.013 (0.007)	1.42 (1.19)	29
Yardeni	-1.003 (1.564)	0.076 (0.139)	0.253 (0.185)	0.004 (0.009)	0.67 (2.91)	28
Platt	-1.67 (1.632)	-0.013 (0.128)	0.034 (0.196)	0.012 (0.013)	0.5 (0.23)	22
Reynolds	-3.37 (1.38)	0.019 (0.127)	0.505 (0.168)	0.015 (0.008)	4.54* (2.31)	27
Braverman	0.802 (1.033)	-0.171 (0.105)	0.051 (0.141)	-0.002 (0.006)	1.21 (1.26)	27
Boltz	-0.793 (0.959)	0.037 (0.101)	0.256 (0.139)	-0.004 (0.006)	1.27 (2.53)	24
Leisenring	-0.673 (1.02)	-0.086 (0.085)	-0.051 (0.171)	0.009 (0.01)	0.81 (0.71)	22
Straszheim	-0.312 (1.16)	-0.017 (0.116)	0.106 (0.151)	-0.001 (0.007)	0.2 (0.34)	21
Kellner	1.587 (1.341)	-0.191 (0.141)	0.244 (0.195)	-0.014 (0.008)	2.26 (1.88)	23
Dederick	-1.47 (0.989)	-0.02 (0.093)	0.136 (0.137)	0.008 (0.006)	1.11 (0.36)	20
Robertson	0.199 (1.24)	-0.215 (0.117)	0.154 (0.166)	-0.003 (0.007)	1.67 (2.06)	20

note: F-Statistic is for the hypothesis that all slope coefficients are jointly zero and has 3 and (n - 4) degrees of freedom; a star denotes the value is not significant at 95%

note: F-Statistic in parentheses represents a joint test with no intercept

note: Independent variables are the prior 6-month changes in GDP and CPI and the actual exchange rate of the YEN at time of forecast

Appendix G: Efficiency Tests for 6-month YEN Forecasts

Forecaster	*Intercept* (*s.e.*)	*GDP* (*s.e.*)	*CPI* (*s.e.*)	*YEN* (*s.e.*)	*F* *F (no intercept)*	*n*
Resler	18.544	0.565	2.11	-0.239	1.15	25
	(18.574)	(1.32)	(2.07)	(0.143)	(1.08)	
Hoffman	33.06	-0.12	2.89	-0.357	1.91	24
	(20)	(1.424)	(2.25)	(0.155)	(0.97)	
Levy	49.12	-0.297	0.631	-0.451	3.51*	25
	(18.89)	(1.346)	(2.105)	(0.146)	(2.03)	
Wyss	22.84	0.96	7.86	-0.412	3.09	22
	(22.99)	(1.743)	(2.76)	(0.189)	(2.91)	
Harris	14.75	0.351	4.27	-0.27	1.6	25
	(21.15)	(1.51)	(2.36)	(0.163)	(2.21)	
Sinai	6.1	-0.266	4.03	-0.154	0.93	24
	(24.17)	(1.72)	(2.71)	(0.186)	(0.95)	
Ratajczak	23.68	0.432	3.44	-0.274	1.4	24
	(20.5)	(1.49)	(2.27)	(0.157)	(1.55)	
Berson	27.42	-0.483	1.097	-0.271	1.26	23
	(19.86)	(1.423)	(2.273)	(0.141)	(1.26)	
Daane	5.87	0.875	5.47	-0.191	1.49	25
	(23.87)	(1.69)	(2.66)	(0.184)	(2.06)	
Hymans	41.11	-0.85	3.94	-0.428	2.28	22
	(19.83)	(1.33)	(2.32)	(0.167)	(0.78)	
Rippe	15.46	-0.734	1.67	-0.151	0.54	23
	(22.38)	(1.603)	(2.56)	(0.158)	(0.41)	
Hyman	54.59	-0.96	0.282	-0.493	4.45*	23
	(18.22)	(1.29)	(2.03)	(0.139)	(1.83)	
Smith	27.81	-1.03	0.907	-0.332	0.87	24
	(26.81)	(1.91)	(3.01)	(0.208)	(5.9*)	
Fosler	58.71	-0.37	-1.78	-0.41	2.97	21
	(21.88)	(1.47)	(2.46)	(0.157)	(1.17)	
Shilling	49.24	0.383	1.08	-0.56	3.86*	23
	(25.3)	(1.81)	(2.9)	(0.18)	(9.67*)	
Moskowitz	43.23	0.197	1.93	-0.44	2.37	24
	(22.12)	(1.61)	(2.45)	(0.17)	(1.27)	
Yardeni	66.04	-0.264	-1.63	-0.49	4.04*	24
	(20.78)	(1.51)	(2.31)	(0.16)	(1.84)	
Platt	10.19	1.15	5.4	-0.26	2.51	21
	(18.18)	(1.46)	(2.16)	(0.14)	(2.5)	
Reynolds	30.34	1.19	1.36	-0.31	3.08	22
	(14.96)	(1.19)	(1.79)	(0.118)	(1.51)	
Braverman	30.93	-1.5	3.61	-0.308	2.57	21
	(18.36)	(1.47)	(2.24)	(0.145)	(1.57)	
Boltz	32.16	0.312	7.04	-0.466	2.37	18
	(21.7)	(1.69)	(3.03)	(0.199)	(1.68)	
Leisenring	29.83	-0.252	3.46	-0.343	1.03	19
	(22.17)	(1.68)	(2.95)	(0.198)	(.41)	
Straszheim	13.09	0.421	7.04	-0.315	2.36	17
	(18.43)	(1.46)	(2.68)	(0.173)	(2.47)	
Kellner	11.74	-0.696	6.18	-0.235	0.8	16
	(28.48)	(2.23)	(4.1)	(0.263)	(1.4)	
Dederick	37.85	-0.781	8.35	-0.53	2.12	16
	(25.44)	(1.96)	(3.61)	(0.237)	(1.43)	
Robertson	30.76	-0.432	7.89	-0.476	2.19	14
	(24.41)	(1.81)	(3.33)	(0.22)	(1.6)	

note: F-Statistic is for the hypothesis that all slope coefficients are jointly zero and has 3 and (n - 4) degrees of freedom; a star denotes the value is not significant at 95%
note: F-Statistic in parentheses represents a joint test with no intercept
note: Independent variables are the prior 6-month changes in GDP and CPI and the actual exchange rate of the YEN at time of forecast

Appendix H: Efficiency Tests for 6-month GDP Forecasts

Forecaster	*Intercept* (s.e.)	*GDP* (s.e.)	*CPI* (s.e.)	*YEN* (s.e.)	*F* *F (no intercept)*	*n*
Resler	-1.36	0.538	-0.093	0.008	5.14*	32
	(1.564)	(0.151)	(0.212)	(0.009)	(10.89*)	
Hoffman	-3.74	0.592	-0.113	0.027	5.98*	27
	(2.02)	(0.157)	(0.266)	(0.016)	(8.52*)	
Levy	-0.227	0.416	-0.275	0.004	4.09*	32
	(1.59)	(0.153)	(0.215)	(0.009)	(6.95*)	
Wyss	-2.11	0.764	0.037	0.004	6.52*	27
	(1.72)	(0.182)	(0.241)	(0.01)	(7.55*)	
Harris	-0.81	0.635	-0.202	0.001	9.56*	30
	(1.61)	(0.138)	(0.193)	(0.009)	(12.57*)	
Sinai	-0.671	0.36	-0.094	0.005	2.55	30
	(1.52)	(0.149)	(0.206)	(0.009)	(6.53*)	
Ratajczak	-1.51	0.42	-0.23	0.013	3.58*	29
	(1.81)	(0.16)	(0.226)	(0.011)	(6*)	
Berson	0.345	0.474	-0.67	0.006	11.27*	24
	(2.33)	(0.169)	(0.268)	(0.016)	(12.4*)	
Daane	-1.16	0.371	-0.298	0.013	4.1	27
	(2)	(0.144)	(0.233)	(0.016)	(6.13*)	
Hymans	-1.17	0.811	-0.342	0.002	12.31*	31
	(1.843)	(0.161)	(0.224)	(0.011)	(13.53*)	
Rippe	1.754	0.412	-0.764	-0.001	13.45*	24
	(2.18)	(0.158)	(0.251)	(0.015)	(14.98*)	
Hyman	-1.62	0.497	0.104	0.006	3.72*	31
	(1.56)	(0.152)	(0.215)	(0.009)	(8.85*)	
Smith	1.485	0.628	-0.612	-0.011	19.41*	28
	(1.484)	(0.131)	(0.187)	(0.009)	(19.12*)	
Fosler	-0.91	0.28	-0.89	0.022	4.36*	22
	(3.45)	(0.235)	(0.389)	(0.024)	(4.69*)	
Shilling	0.605	0.463	0.465	-0.0052	1.38	30
	(2.62)	(0.251)	(0.374)	(0.015)	(19.41*)	
Moskowitz	-0.989	0.538	0.115	-0.001	4.18*	29
	(1.6)	(0.158)	(0.212)	(0.009)	(8.29*)	
Yardeni	3.36	0.395	-0.635	-0.013	3.82*	28
	(2.72)	(0.242)	(0.322)	(0.016)	(5.28*)	
Platt	-0.335	0.4	-0.327	0.007	4.08*	22
	(2.28)	(0.179)	(0.274)	(0.018)	(6.39*)	
Reynolds	2.49	0.572	-0.265	-0.02	7.07*	27
	(1.95)	(0.179)	(0.236)	(0.012)	(8.74*)	
Braverman	1.06	0.138	-0.108	0.002	0.48	27
	(1.59)	(0.162)	(0.216)	(0.009)	(11.74*)	
Boltz	-0.447	0.481	-0.172	0.001	3.88*	24
	(1.48)	(0.156)	(0.214)	(0.009)	(4.68*)	
Leisenring	-2.55	0.475	-0.46	0.028	4.61*	22
	(1.94)	(0.162)	(0.326)	(0.018)	(5.6*)	
Straszheim	-0.372	0.62	0.219	0.016	2.46	21
	(2.37)	(0.238)	(0.309)	(0.014)	(2.38)	
Kellner	-2.57	0.687	0.365	0.003	3.49*	23
	(2.07)	(0.217)	(0.3)	(0.012)	(5.31*)	
Dederick	-2.2	0.557	-0.021	0.009	2.24	20
	(2.33)	(0.218)	(0.322)	(0.015)	(2.15)	
Robertson	-0.984	0.396	-0.187	0.007	2.01	20
	(1.96)	(0.186)	(0.262)	(0.011)	(2.37)	

note: F-Statistic is for the hypothesis that all slope coefficients are jointly zero and has 3 and (n - 4) degrees of freedom; a star denotes the value is not significant at 95%

note: F-Statistic in parentheses represents a joint test with no intercept

note: Independent variables are the prior 6-month changes in GDP and CPI and the actual exchange rate of the YEN at time of forecast

Appendix I: Results of Lamont Tests

CPI 6-month	intercept (s.e.)	slope (s.e.)	R square	n
Resler	0.237	0.0002	0.0001	32
	0.094	0.005		
Levy	0.221	0.007	0.0764	32
	0.086	0.005		
Sinai	0.451	-0.009	0.075	30
	0.104	0.006		
Hymans	0.32	0.009	0.033	31
	0.166	0.009		
Hyman	0.68	-0.002	0.003	31
	0.144	0.008		

CPI 1-year				
Resler	0.325	0.005	0.011	26
	0.182	0.009		
Levy	0.246	0.012	0.0827	26
	0.164	0.008		
Braverman	0.805	-0.016	0.113	26
	0.15	0.009		
Hymans	0.395	0.011	0.0757	25
	0.159	0.008		
Moskowitz	0.49	-0.004	0.016	25
	0.124	0.007		

YEN 6-month				
Resler	1.66	0.069	0.151	25
	0.61	0.034		
Levy	1.96	0.173	0.099	25
	1.94	0.109		
Harris	3.707	-0.026	0.005	25
	1.23	0.076		
Moskowitz	1.085	0.247	0.203	24
	1.694	0.104		
Ratajczak	4.15	0.026	0.004	24
	1.46	0.089		

YEN 1-year				
Resler	2.89	0.088	0.075	20
	1.35	0.073		
Levy	4.29	0.175	0.069	20
	2.82	0.152		
Ratajczak	5.64	0.01	0.0004	20
	2.16	0.125		
Daane	3.41	0.223	0.092	20
	2.75	0.166		
Yardeni	7.87	0.084	0.008	20
	3.562	0.218		

GDP 6-month				
Resler	0.399	-0.002	0.002	32
	0.124	0.007		
Levy	0.24	0.013	0.104	32
	0.133	0.007		
Hymans	0.68	0.005	0.006	31
	0.226	0.012		
Hyman	0.281	0.015	0.076	31
	0.185	0.01		
Shilling	2.95	-0.044	0.057	30
	0.631	0.034		

note: F statistic represents a joint test that the intercept = 0 and slope = 1; a star denotes the value is not significant at 95%