HETEROGENEOUS EXPECTATIONS: AGGREGATION BIAS AND THE POOLABILITY OF SURVEY FORECASTS IN TESTS OF THE RATIONAL EXPECTATIONS HYPOTHESIS

by

Carl S. Boham and Richard Cohen

Working Paper No. 96-16
September 1996

Heterogeneous Expectations: Aggregation Bias and the Poolability of Survey Forecasts in Tests of the Rational Expectations Hypothesis

Carl S. Bonham and Richard Cohen*

September 27, 1996

prepared for

Conference on Expectations in Economics:
In Honor of the 50th Anniversary of the Livingston Survey

October 3 & 4, 1996
Philadelphia Federal Reserve Bank, and University of Pennsylvania

* Bonham, Department of Economics, University of Hawaii at Manoa, Honolulu, Hawaii 96822, and Visiting Associate Professor, Kobe University Graduate School for International Cooperation Studies, phone: 808 956 2321, fax: 808 956 4347, e-mail: bonham@hawaii.edu; Cohen, Department of Economics, University of Hawaii at Manoa, Honolulu, Hawaii 96822, phone: 808 949 6728, e-mail: rcohen@hawaii.edu. We thank seminar participants at the University of Hawaii; and the 1996 Western Economic Association Conference. We thank Howard He for very helpful research assistance.

Introduction

At least since Keynes gave "animal spirits" a prominent role in his theory of investment, both consumer and business expectations have been central to macroeconomics. As a result, understanding how expectations are formed and obtaining direct measures of expectations have received considerable attention.[1] Following Muth's (1961) paper introducing the rational expectations hypothesis (REH), a large literature has developed to test that hypothesis.[2] An inherent difficulty with testing the REH is that agents' expectations are unobservable. Nevertheless, survey data on expectations have been used for many years in testing the implications of rationality. Survey data are most commonly used in its "consensus" form, i.e., for each time period the cross-section of survey respondents' forecasts is averaged to form a single time series prediction. However, a number of authors have argued against using consensus forecasts in rationality tests due to aggregation bias. Figlewski and Wachtel (1981, hereafter FW), Urich and Wachtel (1984), Keane and Runkle (1990, hereafter KR), and Ehrbeck (1992) pooled cross-section and time-series survey responses in an attempt to avoid the potential aggregation bias inherent in the use of consensus forecasts. Finally, Urich and Wachtel (1984), and Bonham and Cohen (1992) tested rationality by treating the survey respondents as different individuals who are not to be aggregated or pooled. None of this literature has recognized that for unbiasedness regressions with integrated predictions, if the conditions for pooling are met it is safe to form the consensus as well.

The purpose of this paper is to establish, in the context of cointegrated predictions and realizations, when it is appropriate to pool, aggregate, or use individual forecasts in direct tests of the unbiasedness implication of the REH. In doing so, we demonstrate the close relationship between aggregating and pooling individual forecasts, and we establish

---

[1] See Klein (1954), Haavelmo (1958, pp. 356-7), and more recently Pesaran (1987), Zarnowitz (1984, 1985), and Zarnowitz and Braun (1992).

[2] See Lovell (1986), Pesaran (1987) or Holden, Peel, and Thompson (1985) for recent surveys of the literature on testing the REH.

the importance of testing for micro-homogeneity (i.e. the equality of corresponding unbiasedness coefficients across individuals) before analyzing the unbiasedness of survey forecasts. For various forecasts from the Survey of Professional Forecasters, we show that micro-homogeneity does not hold, and we conclude that unbiasedness should only be tested at the individual level. Furthermore, the failure of micro-homogeneity implies that agents make use of different private and/or public information when forming their forecasts. Such information heterogeneity generally leads to behavioral uncertainty and rules out "a rigorous derivation of the REH from principles of economic optimization . . . [and] casts serious doubt on the validity of the formulation of the REH along the lines proposed, for example, by Hansen and Sargent (1980)" (Pesaran, 1987, p. 50).[3]

The first section of this paper describes the relationship between heterogeneity bias from improper pooling and aggregation bias from the improper use of consensus forecasts. Following Hsiao (1986), we argue that micro-homogeneity is a necessary assumption for estimating pooled unbiasedness regressions. We distinguish between two different types of aggregation bias. One type of bias, which may arise if micro-homogeneity does not hold, we refer to as Pure Aggregation Bias (PAB). We refer to the second type of bias, which may arise when rational individuals rely on both public and private information, as Private Information Bias (PIB). We show that when micro-homogeneity holds, it is appropriate to conduct unbiasedness tests using either consensus or pooled regressions.

On the other hand, a failure of micro-homogeneity in unbiasedness regressions implies that some individuals have biased forecasts. Therefore, consensus unbiasedness parameters are generally inconsistent, or lead to false acceptance of the unbiasedness hypothesis. Also, pooling individual forecasts is not an alternative to consensus regressions, because heterogeneity bias also implies inconsistent parameter estimates. Therefore, micro-homogeneity tests are an important component of any strategy for

[3] Also see Frydman and Phelps (1983).

performing direct tests of the unbiasedness hypothesis. In the second section of the paper we develop such a test.

In section two, we extend Zellner's (1962a) test of micro-homogeneity to the case of Generalized Method of Moments (GMM) estimation. Our GMM test statistic incorporates a weighting matrix suggested by KR (1990) to account for the possibility of contemporaneous and lagged dependencies across respondents' forecast errors. Such dependencies are likely to arise in the presence of information lags and aggregate shocks to the forecast target series. Finally, in section three we apply our tests to various forecasts made by respondents to the Survey of Professional Forecasters. We reject the null hypothesis of micro-homogeneity and conclude that some of the forecasts in this sample are not rational forecasts. Neither pooled nor consensus methods are recommended for testing unbiasedness for this sample of forecasts. Testing for unbiasedness at the individual level may be useful for determining the relative frequency of individual rejections. Section four concludes.

1. To Aggregate, Pool, or Neither

"[E]xpectations, since they are informed predictions of future events, are essentially the same as the predictions of the relevant economic theory," (Muth, 1961, p. 316). The Rational Expectations literature has generally taken this well known quote to mean "that the unobservable subjective expectation of *individuals* are exactly the *true* mathematical conditional expectations implied by the model itself." (Begg, 1982, p. 30, emphasis added).

Let ${}_kP_{i,t} \equiv E(A_{t+k}|\phi_{i,t})$ denote the k-step-ahead prediction made by individual i in period t. When forming this forecast, individual i makes use of information contained in (or some subset of) $\phi_{i,t} = \{A_{t-1}, A_{t-2}, \ldots ; {}_k\eta_{i,t-1}, {}_k\eta_{i,t-2}, \ldots ; X_t, X_{t-1}, \ldots\}$, where X_t is a vector of time series available at time t, and ${}_k\eta_{i,t} = A_{t+k} - {}_kP_{i,t}$ is individual i's k-step-ahead forecast error. Individual rational expectations, which are equal to the true mathematical conditional expectations implied by the data generating process (DGP) for A_{t+k}, will possess certain optimal properties. Such expectations will be unbiased—the unconditional mean of forecast errors will be zero, i.e., $E({}_k\eta_{i,t}) = 0$. In addition, rational forecasts will be informationally-efficient—forecast errors will be uncorrelated with information contained in $\phi_{i,t}$, i.e., $E({}_k\eta_{i,t}|\phi_{i,t}) = 0$. The unbiasedness and informational efficiency properties of rational forecasts are often tested in the literature.[4] For example, unbiasedness is often tested through OLS estimation of equations such as:

$$(1) \qquad A_{t+k} = \alpha_i + \beta_i \cdot {}_kP_{i,t} + {}_k\varepsilon_{i,t} \qquad i=1, \ldots N.$$

The null hypothesis of unbiasedness is H_0: $\alpha_i = 0$, $\beta_i = 1$.

Survey data have been used extensively to perform direct tests of the implications of rational expectations as well as other theories of expectations formation. Most of this research has relied on consensus measures of expectations. The Livingston survey has been studied in its consensus form by Turnovsky (1970), Pesando (1975), Carlson (1977), Mullineaux (1978, 1980), Pearce (1979), Brown and Maital (1981), Dietrich and Joines (1983), and Gramlich (1983). Zarnowitz (1984, 1985), Hafer and Hein (1985), and Bonham and Dacy (1991) used the mean response to the Survey of Professional Forecasters (SPF). Other survey forecasts studied in consensus form are represented in the work of Leonard (1982), deLeeuw and McKelvey (1981), Frankel and Froot (1987), Friedman (1980), and Ito (1990).

[4] In this paper, the term rational is used to refer to forecasts which meet all of the optimality properties of the strong version of the REH due to Muth. Unbiasedness refers specifically to a forecast which generate

While consensus forecasts have been used extensively, as early as Kane and Malkiel (1967, p. 343), researchers have asked whether "the expectations of market participants tend to be identical or does there tend to be a wide dispersion of forecasts?" They concluded, based on surveys of interest rate expectations of 200 banks, life insurance companies and nonfinancial corporations, that the data did not support "[t]he hypothesis that investors hold a uniform set of expectations of future interest rates . . ." More recently, Zarnowitz (1985, p. 294) echoed this view.

> Valuable information is limited and unevenly diffused, and its costs and returns to individuals with different skills and interests vary. There is uncertainty and disagreement as to the data and models to be used . . . For all of these reasons, one would not expect individual forecasts to be generally uniform, unbiased, or self-fulfilling. Studies of survey data confirm the existence of substantial differences among concurrent and corresponding forecasts from different sources.

Zarnowitz and Braun (1992, p. 8) were even more forceful, "[t]he premise of a generally shared belief and confidence in a commonly held forecast is so unrealistic as to deprive the theoretical exercises based on it of much practical interest." Given the widespread observation that survey respondents possess very diverse views of the future[5], it is surprising that so many researchers still choose to use "consensus" forecasts in their tests of expectations models.

Even before the REH was popular in macroeconomics, economists had recognized the problems associated with use of consensus forecasts. Bierwag and Grove (1966) showed that if individuals form expectations adaptively with parameters that differ across individuals, then the aggregate "market expectation" is not adaptive, but follows a general Pascal function. Kane and Malkiel (1976) argued that averaging individual forecasts introduces aggregation bias in coefficient estimates. To address this problem, they used a

---

(unconditional) mean zero forecast errors. A rational forecast is always unbiased, but an unbiased forecast is not necessarily fully rational.

[5] See also Urich and Wachtel (1984), Jonung (1981), Zarnowitz (1984), Lovell (1986), and Ito (1990). The view that survey participants are heterogeneous is also reflected in the literature on forecast combination. See Figlewski (1983), Figlewski and Urich (1984). See Clemen (1989) for an early survey of the combination literature.

cross-section of survey forecasts of inflation to test models of the expectations formation mechanism. Zarnowitz (1984) recognized the possibility of serious aggregation errors due to the neglect of cross-sectional and distributional aspects of survey data. He also pointed out the general lack of discussion of aggregation problems in the literature which studies survey forecasts.

The problem of aggregation bias in consensus tests of the REH has been recognized by FW (1981, 1983), Urich and Wachtel (1984), KR (1990), Bonham and Cohen (1992), and Ehrbeck (1992). KR (1990, p. 717) summarized the problems associated with using consensus forecasts to test rationality.

> First, doing so causes serious specification bias. If forecasters are rational, their forecasts will differ only because of differences in their information sets. The mean of many individual rational forecasts, each conditional on a private information set, is not itself a rational forecast conditional on any particular information set (see Stephen Figlewski and Paul Wachtel, 1983). This seemingly minor issue can produce severe bias.
>
> A second problem with using consensus forecasts is that this approach can mask individual deviations from rationality [e.g.] ... some firms are consistently optimistic about future sales while others are consistently pessimistic. Averaging expectations, however, can cancel these biases across firms so that industry mean expectations show no bias.

KR argued that tests of the REH must be conducted using individual data either by estimating pooled time-series cross-section regressions, or by estimating separate regressions for each individual. FW (1981), KR (1990), and Ehrbeck (1992) all used the pooled regression approach in an attempt to avoid aggregation bias. Urich and Wachtel (1984) estimated consensus, pooled, and individual regressions and argued based on analysis of variance tests that the pooled specification is rejected in favor of individual-by-individual tests. Thus, the literature which explicitly addresses the problem of aggregation bias has attempted to avoid such bias by using pooled regression techniques.[6]

[6] Although Zarnowitz (1984) recognized the potential for aggregation bias when using consensus forecasts, Zarnowitz (1985) made use of both consensus and individual regressions in tests of the unbiasedness hypothesis. He argued that pooled parameters were biased due to aggregate shocks. Individual and pooled tests are also found in Hirsch and Lovell (1969) and deLeeuw and McKelvey (1981).

However, the second point made by KR refers specifically to a situation where the individual micro parameters in a typical unbiasedness regression would differ across survey respondents. In this case, pooling is not a straightforward solution to the problem of aggregation bias, as the failure of micro-homogeneity implies that the pooled estimators suffer from heterogeneity bias and therefore false inference may result.

1.1 Heterogeneity Bias

The literature which has explicitly addressed the problem of aggregation bias has attempted to avoid such bias by estimating pooled regressions:[7]

$$(2) \qquad A_{t+k} = \alpha + \beta P_{i,t} + \varepsilon_{i,t}, \qquad i=1, \ldots, N;\ t=1, \ldots, T.$$

A crucial assumption underlying all pooled cross-section time-series techniques is that the true data generating mechanism for the dependent variable, A, can be represented by a single probability distribution function, $P(A\,|\,\bar{\theta})$, which is identical for all individuals at all points in time. If this "representative agent" assumption fails, perhaps because $\bar{\theta}$ is not identical across individuals, least squares parameter estimates from the pooled regression will exhibit heterogeneity bias and false inference may result.[8] Hsiao (1986, pp. 5-8) illustrated several cases in which "a straightforward pooling of all NT observations, assuming identical parameters for all cross-section units, would lead to nonsensical results, because it would represent an average of coefficients that differ greatly across individuals."

For example, consider two forecasters, one who consistently overpredicts and one who consistently underpredicts, i.e. KR's second point. Assume that in equation (1), individual 1 has coefficients $\alpha_1 > 0$, $\beta_1 = 1.0$ and individual 2 has $\alpha_2 < 0$, $\beta_2 = 1.0$. In Figure 1, the pooled coefficients of equation (2) are obtained from the least squares line

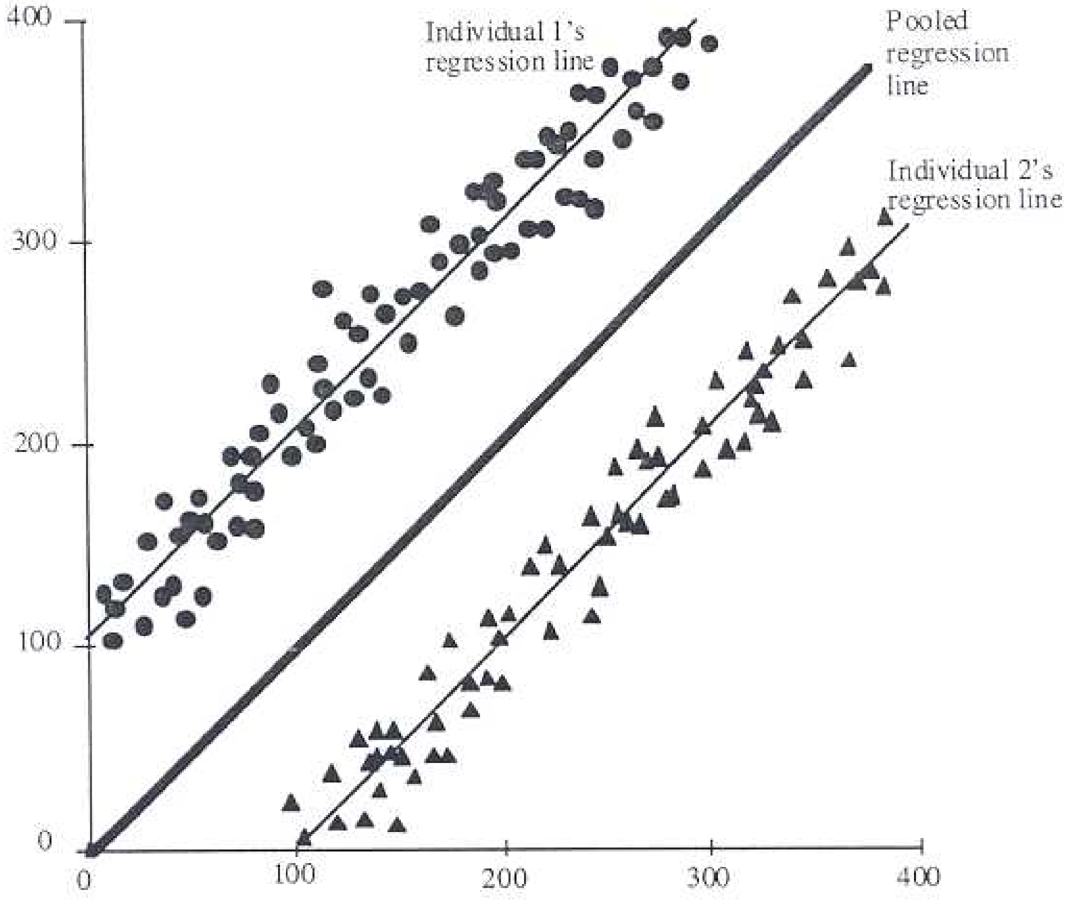
---

[7] For convenience, whenever the forecast horizon is clear we drop the subscript k.

[8] See Hsiao (1986, p. 11).

labeled "Pooled regression line", while the individual coefficients are obtained from the "Individual regression line" passing through their individual scatter plots. The pooled (or consensus) coefficients would clearly lead to non rejection of the unbiasedness hypothesis, although each individual produces biased forecasts.

Figure 1: Nonsensical Pooling



Other examples could be illustrated (e.g. $\alpha_1 < 0$, $\alpha_2 > 0$, $\beta_1 > 1.0$, $\beta_2 < 1.0$). In all of these examples, the clear implication is that pooling should not be undertaken without first testing for the homogeneity of individual forecasts. Kmenta (1986, p. 633) and Baltagi and Raj (1992, p. 100) emphasized the importance of testing for pooling, and Kmenta cautioned that utilizing the results of such tests in subsequent model specification introduces a pre-test bias. Zellner (1962a) showed that it is possible to test the hypothesis that coefficient vectors are all the same, and recommended performing such tests whenever possible to place the interpretation of coefficient estimates on "firm ground" (Zellner, 1962b, p. 117).

Not all panel data analysis requires prior tests for poolability; in many situations, heterogeneity among agents may be accommodated by allowing for variable intercepts and/or slopes.[9] In the case of pooled tests of unbiasedness, no such parametric allowance for individual heterogeneity is possible. Of the studies which use pooled regressions to test the REH, only Urich and Wachtel (1984) test for poolability. They apply analysis of variance tests for poolability in both unbiasedness and efficiency regressions involving forecasts of changes in the money supply. Their approach differs from ours in two ways. First, although they explicitly note that aggregation bias renders the consensus coefficients inconsistent, they do not address the heterogeneity bias in the pooled specification. Rather, they prefer the disaggregated system over the pooled regression for reasons of efficiency; that is, they interpret their rejections of pooling as conveying information about the incremental explanatory power of the disaggregated system. Second, and more importantly, Urich and Wachtel's variance covariance matrix does not account for cross-correlation of disturbances, so their conclusions may be invalid.

[9] Recent panel data testing of the consumption-based capital asset pricing model illustrates this point. See Zeldes (1989) and Runkle (1991) for examples. See Grossman and Shiller (1982), and Singleton (1990) for discussions of aggregation conditions which permit the construction of "representative agent" utility function for these types of models.

1.2 Aggregation Bias

The classic problem of aggregation bias is to define the relationship between the micro coefficients in:

$$(1) \qquad A_{t+k} = \alpha_i + \beta_i P_{i,t} + \varepsilon_{i,t}, \qquad i=1, \ldots N,$$

and the macro coefficients in:

$$(2) \qquad A_{t+k} = \alpha + \beta P_t + \varepsilon_t, \qquad t=1, \ldots, T.$$

In (1), A_{t+k} is the actual value of the target series in period t+k, $P_{i,t}$ is the prediction made by individual i in period t for period t+k, and $\varepsilon_{i,t}$, the regression residual, is an unbiased k-step ahead forecast error if $\alpha_i = 0$ and $\beta_i = 1$. N is the number of individual respondents in a survey. In (2), $P_t = \frac{1}{N}\sum_{i=1}^{N} P_{i,t}$ is the average or "consensus" prediction.

RE testing aside, the problem of aggregating over micro units has a long tradition in both macroeconomics and econometrics (see Tinbergen (1939), and Klein (1946)). The pioneering work of Theil (1954) showed that coefficients of interest in macro relationships will generally depend on complicated combinations of "corresponding" and "noncorresponding" micro coefficients. For instance, α in equation (2) will be a function of not only the individual α_i, but also a weighted average of the β_i, while β will be a weighted average of the β_i. Theil defined aggregation bias as the difference between the macro coefficient and the average of the "corresponding" micro coefficients and showed that a sufficient condition for zero (pure) aggregation bias is the equality of all individual coefficient vectors (micro-homogeneity), i.e. $\alpha_i = \alpha_j$, and $\beta_i = \beta_j$, for all i, j. Because this sufficient condition for no PAB is also a necessary assumption when pooling the individual forecasts, it is worthwhile exploring the relationship between Theil's definition of aggregation bias and the two problems summarized by KR.

The first problem referred to by KR is that the average of individually rational predictions, based on private information sets, is not rational on any particular information set. They cite FW (1983) who show that individuals producing rational forecasts based on public and private information will make different forecasts.[10] Although individual forecasts differ, because they are rational forecasts, they are unbiased with $\alpha_i = 0$, $\beta_i = 1$, for all i. Hence, micro-homogeneity is a direct implication of individual rationality, and Theil's sufficient condition for no PAB is met. However, because each forecaster cannot use other individuals' private information, FW (1983) argue that the consensus forecast is dependent on the consensus-forecast error and hence is not a rational forecast. Intuitively, the consensus forecast error aggregates the omitted private information in each individual's forecast errors, while the consensus prediction aggregates both public and private information in each individual's forecast. As a result, the consensus prediction and the consensus prediction error are correlated. FW (1983) show that ordinary least squares estimates of the macro parameter in equation (2) are biased because of PIB induced by aggregation when individuals cannot make use of other forecasters' private information.

In Appendix II, we present a version of FW's (1983) model of rational individuals who make use of both public and private information. Our model explicitly addresses the stochastic properties of the DGP for the target series A_{t+k} and its predictions, $P_{i,t}$. For target series that are integrated, rational agents will have integrated predictions that are cointegrated with the actual series. Rational forecast errors will be stationary, and therefore omitted information is stationary and asymptotically uncorrelated with the consensus

---

[10] It is important to note that individual, rational forecasts differ only due to the use of different private-information variables. In other words, the strong version of REH due to Muth requires full knowledge of the structure of the economy, including knowledge of the structural parameters. See Pesaran (1987, pp. 21-30).

forecast.[11] As a result, consensus estimators are super consistent; for the majority of survey data forecasts of integrated time series, PIB is unlikely to pose much of a problem in large samples.[12] This conclusion obviates the first problem discussed by KR. Unfortunately, this same conclusion does not hold for their second point.

The second problem is that aggregation may mask systematic individual differences. For instance, some individuals may systematically underpredict the target series, while others systematically overpredict. By averaging the individual forecasts, the consensus will hide such individual differences and could lead to a failure to reject the REH. (This is the same conclusion reached in the case of the pooled regression in Figure 1.) These systematic over or underpredictions imply individual bias and micro coefficients, in equation (1), that differ across individuals. While micro-homogeneity always holds if all forecasters are rational, when some individuals produce biased forecasts micro-homogeneity fails, except in the unlikely event that all forecasters are biased in exactly the same way.

Another implication of irrationality is that both individual and consensus coefficient estimates are generally inconsistent.[13] This inconsistency arises regardless of whether forecasters omit other individuals' stationary private information. PIB is ruled out by our assumptions regarding the stochastic properties of the DGP. However, if the assumption of stationary private information is relaxed, along with the assumption of rationality, the obvious implication is that consensus estimates are inconsistent due to PIB as well.[14]

---

[11] See Granger (1991, pp. 69-70). ARIMA models can be used to produce cointegrated predictions with stationary prediction errors.

[12] See Appendix II for a proof of the lack of PIB under our assumptions regarding the stochastic processes generating the target series and the predictions. See Stock (1987) for discussion of super consistency.

[13] In Appendix III, we relax the assumption of complete knowledge of the structure of the economy, and demonstrate the difficulties for consensus and individual regressions when some individuals are irrational.

[14] If individuals omit their own mean zero, stationary, private information, or publicly available mean zero, stationary information variables, individual and consensus unbiasedness regressions are unaffected since the unconditional mean of forecast errors is still zero. However, individual and consensus forecasts would fail tests of informational efficiency, i.e., $E(\varepsilon_{i,t}|\phi_{i,t}) \neq 0$, and $E(\varepsilon_t|\Omega_t) \neq 0$, where Ω_t is the information set consisting of all public information. (See Appendix III.)

Finally, if all individual biases average to zero, then although micro-homogeneity does not hold, consensus parameter estimates are consistent and would lead to acceptance of the REH.[15] This is KR's second point; it occurs because averaging eliminates the equal and opposite bias across forecasters. The true consensus parameter values imply unbiasedness, although all individuals are producing biased forecasts. However, contrary to KR, pooling is not a straightforward solution to this problem since pooled estimators suffer from heterogeneity bias, and false inference may result. (See figure 1.)

In summary, individuals producing unbiased predictions of integrated variables, using both public and private information, have different forecasts but identical micro coefficients in unbiasedness regressions. Hence, consensus forecasts do not suffer from PAB or PIB, and neither of KR's two concerns regarding consensus forecasts is correct.[16] Micro-homogeneity is an implication of the REH. While individual forecasts may differ due to stationary private information, micro coefficients in unbiasedness regressions will differ only if some individuals are irrational in the sense that they do not make efficient use of integrated public or private information or are unaware of the structure of the economy or both. Micro-homogeneity is a sufficient condition for no PAB in consensus regressions, a necessary assumption for pooling individual forecasts, and an important component of any strategy for performing direct tests of the unbiasedness hypothesis. We are unaware of any work which recognizes this close relationship between aggregation bias and heterogeneity bias, or which correctly conducts pre-tests for micro-homogeneity before pooling or aggregating individual forecasts.

[15] The conclusion that micro-homogeneity fails, but consensus parameter estimates are consistent also holds if individuals' forecasts differ from unbiased predictions by only a constant. Many individual forecasts could be biased, yet individual and consensus estimates may be consistent. However, it is unclear why one would choose to form a consensus forecast and test for unbiasedness once micro-homogeneity has been rejected, establishing the existence of some irrationality.

[16] If the actual and predicted series are not I(1), then as shown in FW (1983), the consensus forecast will suffer from PIB in the presence of private information.

2. Testing Micro-Homogeneity in Unbiasedness Regressions

We extend Zellner's (1962a) test of micro-homogeneity to the case of GMM estimation. Our GMM test statistic incorporates a weighting matrix suggested by KR (1990, pp. 720-22) to account for the possibility of contemporaneous and lagged dependencies across respondents' forecast errors. Such dependencies are likely to arise in the presence of information lags and aggregate shocks to the forecast target series.

We have two motives for conducting pooling tests in the SUR framework.[17] First and foremost, we can determine whether pooled estimates suffer from heterogeneity bias. Second, because Zellner's test of micro-homogeneity was originally conceived as a test of the sufficient conditions for no aggregation bias, not as a pre-test for pooling, we can evaluate the significance of the common assertion that pooling of data eliminates the "aggregation problem" associated with the use of consensus forecasts. If individuals form expectations using different (incorrect) models or based on different integrated public or private information, i.e., micro-homogeneity fails, then neither the consensus forecast nor the pooled specification are recommended for testing the rationality hypothesis.[18] If rationality testing is possible, it should be done at the individual level.[19]

Below we report pooling tests for a variety of forecasts from the SPF. The null hypothesis of micro-homogeneity is that the slope and intercept coefficients are equal across individual unbiasedness regressions; that is, for each individual regression

---

[17] Other approaches to testing for pooling extend cusum and cusum of squares tests from structural change to panel data. See Baltagi and Raj (1992).

[18] When micro-homogeneity does not hold, contrary to both KR (1990) and Dietrich and Joines (1983) (using Dietrich and Joines' notation), the pooled estimator, $\hat{\alpha}_{1p}$, need not be equal to $\hat{\alpha}_{1i}$, the arithmetic average of the individual regression coefficients. See Zellner (1962b) and Hsiao (1986, pp. 5-8).

[19] Lee, Pesaran and Pierse (1990) showed that aggregation bias may not be present even when all coefficient vectors are not equal. They derive a test statistic for the "no aggregation bias" hypothesis that does not require micro-homogeneity to hold. However, in the context of unbiasedness regressions, if micro-homogeneity has been rejected, implementation of Lee et. al.'s test for aggregation bias may be precluded, because consensus parameter estimates are generally inconsistent or lead to false conclusions. See Appendix III.

(1) $A_{t+k} = \alpha_i + \beta_i P_{i,t} + \varepsilon_{i,t}$, $i=1, \ldots, N$,

we test H_0: $\alpha_1 = \alpha_2 = \ldots = \alpha_N$, and $\beta_1 = \beta_2 = \ldots = \beta_N$.

Following KR's specification of the residual structure, we assume that each forecaster has the same homoscedastic error structure (with a k-period memory, necessitated by publication lag). Similarly, each pair of forecasters has the same homoscedastic cross-covariance (also with a k-period memory).

For each forecaster i,

$$
\begin{aligned}
&E(\varepsilon_{i,t}\,\varepsilon_{i,t}) = \sigma_0^2 && \text{for all i, t;}\\
(4)\quad &E(\varepsilon_{i,t}\,\varepsilon_{i,t+\ell}) = \sigma_\ell^2 && \text{for all i, t, } \ell\text{, such that } 0<\ell\leq k;\\
&E(\varepsilon_{i,t}\,\varepsilon_{i,t+\ell}) = 0 && \text{for all i, t, such that } \ell>k.
\end{aligned}
$$

Every forecaster has the same TxT variance-covariance matrix Q, with elements $q(m-\ell,m) = q(m,m-\ell) = \sigma_\ell^2$ for $\ell \leq k$; 0 otherwise.

For each pair of forecasters i and j,

$$
\begin{aligned}
&E(\varepsilon_{i,t}\,\varepsilon_{j,t}) = \delta_0^2 && \text{for all i, j, t;}\\
(5)\quad &E(\varepsilon_{i,t}\,\varepsilon_{j,t+\ell}) = \delta_\ell^2 && \text{for all i, j, t, } \ell\text{, such that } 0<\ell\leq k;\\
&E(\varepsilon_{i,t}\,\varepsilon_{j,t+\ell}) = 0 && \text{for all i, j, t, such that } \ell>k.
\end{aligned}
$$

Thus, every pair of forecasters has the same TxT cross-covariance matrix R with elements $r(m-\ell,m) = r(m,m-\ell) = \delta_\ell^2$ for $\ell \leq k$; 0 otherwise. The complete variance-covariance matrix of the residuals, denoted Ω, has dimension NTxNT, with matrices Q on the main diagonal and R off the diagonal.[20]

[20]The assumption of $2 \cdot k$ second moments, which are common to all forecasters, is made for analytical tractability and for increased reliability. For example, a large number of missing observations for a given forecaster implies that many individual cross-covariances are calculated with relatively few pairs of residuals.

Stack all N individual regressions into the Seemingly Unrelated Regression system:

$$(6) \qquad A = Xb + \varepsilon,$$

where A is the NTx1 stacked vector of realizations, and X is an NTx2N block diagonal data matrix,

$$(7) \qquad X \equiv \begin{bmatrix} X_1 & & \\ & \ddots & \\ & & X_N \end{bmatrix}.$$

Each $X_i = [i\ P_i]$ is a Tx2 matrix of ones and individual i's forecasts, $b = [\alpha_1\ \beta_1\ \alpha_2\ \beta_2 \cdots \alpha_N\ \beta_N]'$, and ε is an NTx1 vector of stacked residuals. The vector of restrictions, $Rb = r$, corresponding to the null hypothesis of micro-homogeneity is normal distributed, with $Rb - r \sim N[0, R(X'X)^{-1}X'\Omega X(X'X)^{-1}R']$, where R is the 2(N-1)x2N matrix

$$(8) \qquad R \equiv \begin{bmatrix} 1 & 0 & -1 & 0 & \dots & & 0 \\ 0 & 1 & 0 & -1 & 0 & \dots & 0 \\ \vdots & & & & & \dots & \vdots \\ \vdots & & & & & & 0 \\ 0 & \dots & & 0 & 1 & 0 & -1 \end{bmatrix},$$

and r is a 2(N-1)x1 vector of zeros. The corresponding Wald test statistic, $(R\hat{b} - r)'\left[R(X'X)^{-1}X'\hat{\Omega}X(X'X)^{-1}R'\right]^{-1}(R\hat{b}-r)$ is asymptotically distributed as a chi-square random variable with degrees of freedom equal to the number of restrictions $l = 2(N-1)$.

These individual cross-covariances are then averaged to produce a more reliable value. Note that our σ_ℓ and δ_ℓ, for $\ell = 0, \dots k$, are computed using residuals from individual unbiasedness regressions, whereas KR's four second moments are computed using residuals from the pooled regression, which imposes a common slope and intercept on all unbiasedness regressions.

3. Heterogeneity in the Survey of Professional Forecasters

As an example, we consider forecasts of two important macro variables, nominal GNP and the GNP deflator, from the Survey of Professional Forecasters for the period from 1968:4 to 1990:1. These forecasts are of interest because of recent papers by KR (1990) and Bonham and Cohen (1995), which reach opposing conclusions about the rationality of GNP deflator forecasts, and because of the findings of Zarnowitz (1985) that GNP forecasts tend to be more accurate and less biased than price forecasts. One possible explanation is that micro-homogeneity fails because some individual price forecasts are not rational, and pooling such forecasts results in heterogeneity bias and false conclusions, whereas GNP forecasts are homogeneous and rational. However, Zarnowitz (1985) also found greater dispersion of the individual intercepts in unbiasedness regressions using nominal GNP than in GNP deflator inflation regressions.

Before performing micro-homogeneity tests, we address two data related questions. First, what data was available to survey respondents when they formed their forecasts? Second, what are the stochastic properties of the forecasts and realizations, and how do these properties affect our tests?

The Survey of Professional Forecasters, formally known as the ASA-NBER Economic Outlook Survey, was administered by the National Bureau of Economic Research between 1968 and 1990. Beginning in the second quarter of 1990, the survey has been conducted by the Federal Reserve Bank of Philadelphia. Each quarter, the survey is mailed to members of the Business and Economic Statistics Section of the American Statistical Association. Survey respondents are asked to provide forecasts of the current quarter as well as the next four quarters for GNP, the GNP deflator, the major components of GNP, and a number of other macroeconomic indicators. Beginning in 1981, survey respondents were also asked to forecast the current and the following year's annual values. While the survey was administered by the NBER, it was mailed during the second month of each quarter, after the preliminary 45-day release of the previous quarter's

GNP data became available from the Bureau of Economic Analysis. In fact, the 45-day release of the GNP figures were often mailed along with the survey questionnaire, and the survey was due back to the NBER by the end of the second month in the same quarter.[21] When the Philadelphia Federal Reserve Bank took over the survey in 1990:2, they began to mail the surveys in the fifth week of the quarter. More importantly, they required that forecasters return their surveys by the start of the seventh week of the same quarter. There are likely to be a large number of quarters in which the preliminary release of the GNP data was not available when respondents finalized their forecasts. Therefore, we use forecasts from the 1968:4-1990:1 sample period. We only lose two years of forecasts, since the target series were changed to a GDP concept in the first quarter of 1992.

If we refer to the current survey period as quarter t, then survey respondents submit forecasts for quarters t, t+1, . . ., t+4. The most recent quarter for which data would have been available is the quarter preceding the survey period, i.e. quarter t-1. Thus, forecasters knew their own errors for quarter t-1 (at least the errors based on the 45-day release) when forming forecasts for quarter t and beyond. Rational forecasts of the current quarter (we refer to these as zero-quarter-ahead forecasts) have serially independent forecast errors, while k-quarter-ahead forecast errors may follow a moving average process of order k.[22] We allow for these expected serial correlation patterns in our GMM variance covariance matrix estimates.

To insure that our time series of realizations have not been revised using data unavailable to survey respondents when they formed their forecasts, we collected the unrevised 45-day release of each series from back issues of the *Survey of Current*

[21] See Su and Su (1975), and Zarnowitz and Braun (1992) for a description of the scope and timing of the original ASA-NBER Economic Outlook Survey. See Croushore (1993) for a description of the Survey of Professional Forecasters.

Business. We delete the fourth quarter of 1975, 1980, and 1985, as these are quarters in which potentially unforecastable major benchmark revisions occurred. We also rebase the realizations and each forecast of the GNP deflator to a single (1958=100) base year. Finally, because individual forecasters responded to the survey sporadically, a decision must be made regarding the minimum number of responses required for a forecaster to be included in our tests. We work with only those forecasters with at least twenty survey responses, not necessarily consecutive.

It is well known that nominal GNP and the GNP deflator are integrated processes.[23] To conduct inference based on regressions containing integrated series, we rely on the results of West (1988). West showed that standard inference (i.e., conventional normal asymptotic theory) can be used in cointegrating regressions which can be written in a form containing a single integrated regressor with drift; any additional regressor must be stationary.[24] We assume that first differences of GNP, the GNP deflator and the forecasts of these series are stationary with a nonzero unconditional mean; we also assume that the realizations and the forecasts are cointegrated.[25] Under these assumptions, both individual unbiasedness regressions and pooled unbiasedness regressions will meet the conditions of West and allow for standard inference. Unfortunately, many survey participants have

---

[22] KR (1990) test whether current quarter (they refer to them as one-quarter-ahead forecasts) predictions are rational and assume that the previous quarters realization (and forecast errors) are not known when the prediction is made in quarter t. Therefore, they assume an MA(1) error structure for current quarter predictions while we assume an MA(0) error structure. Our results do not change if we assume an MA(1) error structure, and our estimates of σ_1^2 and δ_1^2 are very close to zero.

[23] See Nelson and Plosser (1982).

[24] Hylleberg and Mizon (1989) have shown in simulation studies that the drift term needs to be fairly large for it to dominate the effects of the integrated part of the DGP. Conclusions from our pooling tests are unaffected by whether we use levels or differences of forecasts and realizations.

[25] Using SPF data Bonham and Cohen (1992) tested for cointegration between realizations and both zero-step-ahead and one-step-ahead forecasts of the GNP deflator. Given the large number of respondents with missing observations, their tests were conducted on a subset of the entire sample of respondents. Out of thirty one cases (16 one-step-ahead regressions and 15 zero-step-ahead regressions), they rejected the null of no cointegration in 22 and 25 cases at the 5% and 10% levels respectively. They conclude that cointegration is likely to hold for all 80 respondents in their sample. Lahiri and Chun (1989) provide evidence of cointegration between consensus GNP and GNP deflator forecasts and their corresponding realizations for a variety of forecast horizons.

sporadic response records and thus a large number of missing observations. The fact that some respondents will have as few as twenty forecasts should be kept in mind when interpreting our results.

Table 1 presents test results for the null hypothesis of micro-homogeneity for nominal GNP using various forecast horizons and samples of individuals with a minimum of 20, 40, and 50 survey responses. (Results based on a minimum of 30 responses do not differ and are available from the authors on request).

TABLE 1
Micro-Homogeneity Tests For GNP Regressions

(1) $A_{t+k} = \alpha_i + \beta_i P_{i,t} + \varepsilon_{i,t}$ $\quad$ $i=1, \ldots, N;\ t=1, \ldots, T.$

H_0: $\alpha_1=\alpha_2=\cdots=\alpha_N,\ \beta_1=\beta_2=\cdots=\beta_N$ $\quad$ 1968:4 - 1990:1

1	2	3	4	5	6	7
Series	k	χ^2	MSL	N	Avg. Resp.	Max/Min
GNP	0	313.71	(0.000)	79	35	73/20
GNP	0	77.72	(0.019)	28	48	73/40
GNP	0	18.42	(0.428)	10	57	73/50
GNP	1	274.53	(0.000)	79	35	73/20
GNP	1	74.11	(0.036)	28	48	73/40
GNP	1	20.03	(0.331)	10	57	73/50
GNP	4	228.46	(0.000)	74	33	68/20
GNP	4	42.80	(0.036)	15	48	68/40
GNP	4	11.28	(0.336)	6	56	68/50

*Column 1 gives the target series: GNP is the 45-day preliminary release of nominal Gross National Product. Column 2 gives number of quarters between when the prediction is made and the realization is available, i.e. the order of the MA error assumed in the test. Column 3 is the value of the Wald-test for H_0. Column 4 is the marginal significance level of the test in column 3. Columns 5 through 7 give the number of forecasters in our tests, the average, and maximum/minimum number of individual survey responses.

Micro-homogeneity is rejected at a marginal significance level of virtually zero when the response limit is 20, and at less than 4% for a response limit of 40. Only when we restrict our sample to those individuals with a minimum of 50 responses do we fail to reject the null of micro-homogeneity. Those tests include forecasts from, at most, 10 of the 79 individuals in our full sample. For the zero-quarter-ahead GNP forecasts of these 10 individuals, the standard deviations of the individual intercept and slope estimates are three and four times smaller (respectively) than for the full sample of survey respondents. Table 2 presents similar results for the GNP deflator forecasts with the exception that we reject micro-homogeneity for all forecast horizons and all samples of survey respondents.

TABLE 2
Micro-Homogeneity Tests for GNP Deflator Regressions

(1) $A_{t+k} = \alpha_i + \beta_i P_{i,t} + \varepsilon_{i,t}$ $i=1, \ldots, N; \ t=1, \ldots, T.$

$H_0: \alpha_1=\alpha_2=\cdots=\alpha_N, \ \beta_1=\beta_2=\cdots=\beta_N$ 1968:4 - 1990:1

1	2	3	4	5	6	7
Series	k	χ^2	MSL	N	Avg. Resp.	Max/Min
PGNP	0	340.43	(0.000)	80	34	72/20
PGNP	0	175.06	(0.000)	28	48	72/40
PGNP	0	62.17	(0.000)	10	57	72/50
PGNP	1	365.69	(0.000)	80	34	72/20
PGNP	1	258.50	(0.000)	28	48	72/40
PGNP	1	101.43	(0.000)	10	57	72/50
PGNP	4	375.88	(0.000)	72	33	67/20
PGNP	4	58.38	(0.0003)	14	49	67/40
PGNP	4	29.40	(0.000)	5	56	67/50

*Column 1 gives the target series: PGNP is the 45-day release of the GNP deflator. See Table 1 for definitions of columns 2-7.

Finally, Tables 3 and 4 present results from splitting the 1968:4 - 1990:1 sample in half.

TABLE 3
Sub Sample Micro-Homogeneity Tests

(1) $A_{t+k} = \alpha_i + \beta_i P_{i,t} + \varepsilon_{i,t}$ $\quad$ $i=1, ..., N; \quad t=1, ..., T.$

H_0: $\alpha_1=\alpha_2=\cdots=\alpha_N, \ \beta_1=\beta_2=\cdots=\beta_N$ $\quad$ 1968:4 - 1979:2

1	2	3	4	5	6	7
Series	k	χ^2	MSL	N	Avg. Resp.	Max/Min
GNP	0	220.07	(0.000)	46	29	41/20
GNP	1	237.09	(0.000)	46	29	41/20
GNP	4	243.52	(0.000)	38	27	36/20
PGNP	0	288.78	(0.000)	45	29	41/20
PGNP	1	367.16	(0.000)	45	29	41/20
PGNP	4	417.02	(0.000)	39	27	36/20

* See Table 1 for definitions of columns 1-7.

TABLE 4
Sub Sample Micro-Homogeneity Tests

(1) $A_{t+k} = \alpha_i + \beta_i P_{i,t} + \varepsilon_{i,t}$ $\quad$ $i=1, ..., N; \quad t=1, ..., T.$

H_0: $\alpha_1=\alpha_2=\cdots=\alpha_N, \ \beta_1=\beta_2=\cdots=\beta_N$ $\quad$ 1979:3 - 1990:1

1	2	3	4	5	6	7
Series	k	χ^2	MSL	N	Avg. Resp.	Max/Min
GNP	0	141.24	(0.000)	24	27	39/20
GNP	0	19.68	(0.235)	9	34	39/31
GNP	1	87.27	(0.000)	24	27	39/20
GNP	1	14.46	(0.565)	9	34	39/31
GNP	4	167.38	(0.000)	22	27	39/20
GNP	4	127.90	(0.000)	8	34	39/31
PGNP	0	232.70	(0.000)	24	27	38/20
PGNP	0	44.24	(0.000)	9	34	38/31
PGNP	1	608.42	(0.000)	24	27	38/20
PGNP	1	266.42	(0.000)	9	34	38/31
PGNP	4	489.36	(0.000)	23	26	38/20
PGNP	4	45.49	(0.000)	8	34	38/31

*See Table 1 for definitions of columns 1-7.

Our results are very robust to sample selection. The failure of the micro-homogeneity hypothesis implies that some individuals have biased forecasts and suggests that parameter estimates in consensus unbiasedness regressions will either be inconsistent or lead to false conclusions, or both.[26] Contrary to the claims of the literature, pooled estimates will likely suffer from heterogeneity bias and do not represent a viable alternative to consensus estimates. At least for the forecasts considered in this paper, if rationality can be tested at all, it should be tested at the individual level.

Bonham and Cohen (1992) performed such tests for GNP deflator forecasts and rejected the null hypothesis of unbiasedness at the 5% significance level for over 50% of the eighty survey respondents in their sample of one-quarter-ahead forecasts. The corresponding fraction of rejections for contemporaneous forecasts was 40%. For forecasts of changes in the GNP deflator, Zarnowitz (1985) rejected the unbiasedness hypothesis (at the 5% level) in 26.6% of the contemporaneous forecasts, 46.8% of the one-quarter-ahead forecasts, and 58.2% of the four-quarter-ahead forecasts. For contemporaneous, one-quarter-ahead, and four-quarter-ahead forecasts of the changes in nominal GNP, the rejection rates were 12.7%, 10.1%, and 11.4% respectively. These results are consistent with the finding of heterogeneity of survey respondents. A significant number of individual respondents to the Survey of Professional Forecasters do not produce unbiased predictions of GNP or the GNP deflator, and their forecasts differ systematically.

4. Conclusions

The literature which conducts direct tests of the REH has relied primarily on consensus measures of expectations from surveys such as the Livingston Survey and the Survey of Professional Forecasters. While a number of researchers have recognized the potential for aggregation bias from the use of consensus forecasts, none has tested whether

[26] See footnote 15 for an exception to the inconsistency of consensus unbiasedness parameters when micro-homogeneity fails.

the sufficient conditions for no PAB are met. Theil (1954) showed that if micro-homogeneity fails, then absent some very restrictive, implausible conditions concerning noncorresponding parameters, OLS estimators on aggregate data are inconsistent. Among the researchers who have explicitly addressed the problems of PAB or PIB in test of the REH, most have chosen to make use of individual forecasts by pooling cross-section and time-series observations from surveys. None of these researchers has recognized that the failure of micro-homogeneity causes heterogeneity bias in pooled regressions, also rendering those estimators inconsistent. Conversely, if the necessary micro-homogeneity condition for pooling survey forecasts is met, then at least for the case of cointegrated predictions and realizations, it is appropriate to use consensus forecasts as well. This paper dispels the myth that pooling is a viable alternative to the use of consensus forecasts in testing rational expectations.

As an example, we show that forecasts of nominal GNP and the GNP deflator from the Survey of Professional Forecasters do not pass tests of micro-homogeneity for unbiasedness regressions. This survey data has been used in a number of papers to perform both consensus and pooled tests of the unbiasedness hypothesis. However, if tests of the unbiasedness hypothesis are possible, they should be conducted at the individual level. Tests of unbiasedness using individual forecasts from the SPF have been reported by Zarnowitz (1985), and Bonham and Cohen (1992). In general, these tests lead to the conclusion that survey respondents do not produce unbiased forecasts. The rejections of micro-homogeneity presented here, and the findings of Zarnowitz (1985) and Bonham and Cohen (1992) constitute evidence of both widespread bias in professional forecasts and considerable forecast heterogeneity. Our results, and the conclusions of numerous studies of survey forecasts described above cast serious doubt on the validity of the homogeneous information assumption necessary for a rigorous derivation of the REH from principles of economic optimization. While this conclusion certainly does not imply that optimizing agents do not form expectations, it does emphasize the importance of collecting direct measures of agents expectations so that the expectations formation process may be studied more carefully using the best data possible.

References

Baltagi, B. H., and B. Raj, "A Survey of Recent Theoretical Developments in the Econometrics of Panel Data," *Empirical Economics*, 1992, 17, 85 - 109.

Bierwag, G. O. and M. A. Grove, "Aggregate Koyck Functions," *Econometrica*, October 1966, 34, No. 4, 828-32.

Bonham, C. S. and R. Cohen, "Testing the Rationality of Price Forecasts: Comment," with Richard Cohen, *The American Economic Review*, March 1995, 85, No. 1, 284-289.

_____, and _____, "The Rationality of Price Level Forecasts: Correct Tests using Micro Data," *University of Hawaii*, September 20, 1992, Working Paper No. 92-4.

_____, and D. Dacy, "In Search of a Strictly Rational Forecast," *Review of Economics and Statistics*, May 1991, 73, no. 2, 245-253.

Brown, Bryan W. and Shlomo, Maital, "What Do Economists Know? An Empirical Study of Experts' Expectations," *Econometrica*, March 1981, 49, 491-504.

Carlson, John A., "A Study of Price Forecasts," *Annals of Economic and Social Measurement*, 1977, 6, 27-56.

Clemen, R. T., "Combining Forecasts: A Review and Annotated Bibliography," *International Journal of Forecasting*, 1989, 5, no. 1 (Winter), pp. 27-56.

Croushore, Dean, "Introducing: The Survey of Professional Forecasters," *Philadelphia Federal Reserve Business Review*, 1993, November/December, 3-15.

deLeeuw, Frank and Michael J. McKelvey, "Price Expectations of Business Firms," *Brookings Papers on Economic Activity*, 1981, 1, 299-313.

Dietrich, J. K. and D. H. Joines, "Rational Expectations, Informational Efficiency, and Tests Using Survey Data: A Comment," *Review of Economic and Statistics*, 1983, 65, 525-29.

Ehrbeck, T. "Rejecting Rational Expectations in Panel Data: Some New Evidence," *European University Institute*, May 1992, Working Paper ECO No. 92/91.

Figlewski, S., "Optimal price forecasting using survey data," *Review of Economic and Statistics*, 1983, 65, 13-21.

_____, and T. Urich, "Optimal Aggregation of Money Supply Forecasts: Accuracy, Profitability and Market Efficiency," *The Journal of Finance*, 1983, 28, no. 3, 695-710.

_____, and Paul Wachtel, "The Formation of Inflationary Expectations," *Review of Economic and Statistics*, 1981, 63, 1-10.

_____, and _____, "Rational Expectations, Informational Efficiency, and Tests Using Survey Data: A Reply," *Review of Economic and Statistics*, 1983, 65, 529-531.

Frankel, Jeffrey A. and Kenneth A. Froot, "Using Survey Data to Test Standard Propositions Regarding Exchange Rate Expectations," *American Economic Review*, March 1987, 77, 133-53.

Friedman, Benjamin M., "Survey Evidence on the 'Rationality' of Interest Rate Expectations," *Journal of Monetary Economics*, October 1980, 6, 453-65.

Frydman, Roman, and Edmund S. Phelps (eds.), Individual Forecasting and Aggregate Outcomes, Cambridge: Cambridge University Press, 1983.

Gramlich, Edward M., "Models of Inflation Expectations Formation," *Journal of Money Credit and Banking*, May 1983, 15, 155-73.

Granger, C.W.J., " Developments in the Study of Cointegrated Economic Variables," in Long-Run Economic Relationships, R. F. Engle, and C.W.J. Granger, eds., Oxford, Oxford University Press, 1991, pp. 65-80.

Grossman, S. J., and Robert J. Shiller, "Consumption Correlatedness and Risk Measurement in Economies with Non-Traded Assets and Heterogeneous Information," *Journal of Financial Economics*, 1982, 10, 195-210.

Haavelmo,T., "The role of the econometrician in the advancement of economic theory," *Econometrica*, 1958, 26, 351-357.

Hafer, R. W., S. E. Hein, "On the Accuracy of Time-Series , Interest Rate, and Survey Forecasts of Inflation," *Journal of Business*, 1985, 5, 377-398.

Hansen, L. P. and T. J. Sargent, "Formulating and estimating dynamic linear rational expectations," *Journal of Economic Dynamics and Control*, 1980, 2, 7-46.

Hirsch, Albert A. and Michael C. Lovell, Sales Anticipations and Inventory Behavior, New York: Wiley & Sons, 1969.

Holden K., Peel D. A., and J. L. Thompson, Expectations: Theory and Evidence, Macmillan, London, 1985.

Hsiao, Cheng, Analysis of Panel Data, Cambridge, U. K.: Cambridge University Press, 1986.

Ito, Takatoshi, "Foreign Exchange Rate Expectations: Micro Survey Data," *American Economic Review*, 1990, 80, June, 434-449.

Jonung, Lars, "Perceived and Expected Rates of Inflation in Sweden," *American Economic Review*, 1981, 71, 961-68.

Kane, E. J., and B. G. Malkiel, "The Term Structure of Interest Rates: An Analysis of A Survey of Interest-Rate Expectations," *Review of Economic and Statistics*, 1967, 49, 343-55.

Keane, Michael P. and David E. Runkle, "Testing the Rationality of Price Forecasts: New Evidence from Panel Data," *American Economic Review*, September 1990, 80, 714-735.

Kmenta, Jan, Elements of Econometrics, 2nd ed., New York: Macmillan Publishing Co., Inc., 1986.

Keynes, John Maynard, The General Theory of Employment, Interest, and Money, 1936.

Klein, Lawrence R., "Remarks on the Theory of Aggregation," *Econometrica*, 1946, 14, 303-312.

_____, "Applications of survey methods and data to the analysis of economic fluctuations," in Klein, L. R., ed., Contributions of Survey Methods to Economic Fluctuations, New

York: Columbia University Press, 1954.

Lahiri, Kajal, and T. S. Chun, "Some Tests for Unbiasedness in the Long Run Using Survey Data," *International Economic Review*, Summer 1989, 3, 27-42.

Lee, K. C., Pesaran, M. H., and R. G. Pierse, "Testing for Aggregation Bias in Linear Models," *The Economic Journal*, Conference 1990, 100, 137-150.

Leonard, Jonathan S., "Wage Expectations in the Labor Market: Survey Evidence on Rationality," *Review of Economic and Statistics*, February 1982, 64, 57-61.

Lovell, M. C., "Tests of the Rational Expectations Hypothesis," *American Economic Review*, 1986, 76, 110-24.

Mullineaux, D. J., "On testing for rationality: another look at the Livingston price expectations data," *Journal of Political Economy*, 1978, 86, 329-36.

Mullineaux, D. J., "Inflation expectations and money growth in the United States," *American Economic Review*, 1980, 70, 149-61.

Muth, J. F., "Rational expectations and the theory of price movements," *Econometrica*, 1961, 29, 315-335.

Nelson, Charles R., and Charles I. Plosser, "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics*, 1982, 10, 139-162.

Pearce, Douglas K., "Comparing Survey and Rational Measures of Inflation," *Journal of Money Credit and Banking*, November 1979, 11, 447-56.

Pesando, James E., "A Note on the Rationality of the Livingston Price Expectations," *Journal of Political Economy*, August 1975, 83, 849-58.

Pesaran, Hashem. The Limits To Rational Expectations. New York: Basil Blackwell, 1987.

Runkle, David E., "Liquidity constraints and the permanent-income hypothesis," *Journal of Monetary Economics*, 1991, 27, 73 -98.

Stock, J. H., "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors, *Econometrica*, 1987, 55, 1035-1056.

Su, Vincent, and Josephine Su, "An Evaluation of ASA/NBER Business Outlook Survey Forecasts," *Explorations in Economic Research*, 1975, 2, 588-618.

Theil, Henry, Linear Aggregation of Economic Relations, Amsterdam: North-Holland Publishing Company, 1954.

Tinbergen, J., Statistical Testing of Business Cycle Theories; vol. 1: A Method and its Applications to Investment Activity, Geneva 1939.

Turnovsky, S. J., "Empirical evidence on the formation of price expectations," *Journal of the American Statistical Association*, 1970, 65, 1441-1454.

Urich, Thomas, and Paul Wachtel, "The Structure of Expectations of the Weekly Money Supply announcement," *Journal of Monetary Economics*, March 1984, 13, 183-94.

Zarnowitz, Victor, "The Accuracy of Individual and Group Forecasts From Business Outlook Surveys," *Journal of Forecasting* , 1984, 3, 11-26.

_____, "Rational Expectations and Macroeconomic Forecasts," *Journal of Business and Economic Statistics*, October 1985, 3, 293-311.

_____, and Phillip Braun, "Twenty-Two Years Of The NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisons of Forecasting Performance," *National Bureau of Economic Research Working Paper Series*, January 1992, No. 3965, Cambridge.

Zeldes, S., "Consumption and Liquidity Constraints: An Empirical Investigation," *Journal of Political Economy*, April 1989, 97, 305-46.

Zellner, Arnold, "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias," *Journal of the American Statistical Association*, June 1962a , 348-68.

_____, "Estimation of Cross-Section Relations: Analysis of a Common Specification Error," *Metronomica*, 1962b, 14, 111-117.

Appendices

I. A Consensus of Rational Individuals With Private Information Sets

In this appendix we describe a model of individual, rational forecasters similar to FW (1983). We pay special attention to the stochastic properties of the public and private information used in forming expectations. Suppose that the realized time series, A_{t+k}, is generated by the following data generating process (DGP):

$$(1.1)\quad A_t = X_t\gamma + \mu_t$$
$$= X_{0t}\gamma_0 + X_{1t}\gamma_1 + X_{2t}\gamma_2 + \mu_t,$$

where $\mu_t \sim \text{iid}[0,\sigma_\mu^2]$ and $E[\mu_t\, \mathbf{X}_t] = 0$. Since many economic time series are well described by I(1) processes (see Nelson and Plosser (1982)), we assume that $X_{0t} \sim I(1)$, while X_{1t} and X_{2t} are both serially uncorrelated, I(0) processes with zero means. It follows that $A_t \sim I(1)$, A_t and X_{0t} cointegrate with cointegration coefficient γ_0, and X_{0t} is uncorrelated with either X_{1t} or X_{2t}.

Following FW (1983), assume that there are two types of individuals. Both individuals form predictions of A_t rationally in the sense that they know the structure and parameters of the DGP in equation (1.1). Yet each individual lacks the others' private information. To make our example more concrete, assume that A_t is the realized price level in some economy divided into two regions. X_{0t} is the publicly known aggregate money supply, while X_{1t} and X_{2t} are the privately observed excess demands for goods in regions 1 and 2 respectively. To form their predictions, agents in region 1 use their knowledge of the true DGP and the information set $\phi_{1,t}=\{X_{0t}, X_{1t}\}$, while agents in region 2 use knowledge of equation (1.1) and the information set $\phi_{2,t} = \{X_{0t}, X_{2t}\}$. Unlike FW's (1983) model, the additional assumption that private information variables are mutually independent is not needed here due to the dominance of the I(1) variable X_{0t} in both agents' predictions and the realization.

Rational predictions formed by type 1 and 2 agents are given by:

$$(1.2)\quad \begin{aligned} P_{1,t} &= E[A_t|\phi_{1,t}] &&= X_{0t}\gamma_0 + X_{1t}\gamma_1 \\ P_{2,t} &= E[A_t|\phi_{2,t}] &&= X_{0t}\gamma_0 + X_{2t}\gamma_2 \end{aligned}$$

For each agent, a test of the unbiasedness principle of the REH could be conducted by estimating the following regression:

$$(1.3)\quad A_t = \alpha_i + \beta_i P_{i,t} + \varepsilon_{i,t}, \qquad i = 1,2.$$

Grouping terms in equation (1.1), the DGP for A_t may be written as:

$$(1.4)\quad \begin{aligned} A_t &= X_{0t}\gamma_0 + X_{1t}\gamma_1 + X_{2t}\gamma_2 + \mu_t, \qquad t=1,\ldots,T, \\ &= (X_{0t}\gamma_0 + X_{1t}\gamma_1) + (X_{2t}\gamma_2 + \mu_t), \\ &= \alpha_1 + \beta_1 P_{1,t} + \varepsilon_{1,t}, \end{aligned}$$

where $\alpha_1 = 0$, $\beta_1 = 1$, and $\varepsilon_{1,t} = (X_{2t}\gamma_2 + \mu_t)$ is the forecast error for type 1 individuals. Given our assumptions, regarding the DGP for A_t, the forecast errors, $\varepsilon_{1,t}$, are stationary, mean zero, and uncorrelated with the forecast, $P_{1,t}$. Furthermore, the residuals from estimation of equation (1.4) are equal to the individual's rational forecast error, and hence plim $\hat{\alpha}_1 = \alpha = 0$, and plim $\hat{\beta}_1 = \beta = 1$. A similar construction holds for type 2 agents, so that plim $\hat{\alpha}_1 =$ plim $\hat{\alpha}_2 = 0$, and plim $\hat{\beta}_1 =$ plim $\hat{\beta}_2 = 1$. Therefore, for individual rational forecasts, OLS estimates of the slope and intercept in individual unbiasedness regressions are consistent even in the presence of omitted stationary private information. Below we show that this conclusion also holds for the consensus forecast, i.e., PIB is not a problem under our assumptions. Furthermore, since each individual has the same slope and intercept, Theil's (1954) sufficient condition of micro-homogeneity holds and there will be no PAB in forming a consensus forecast.[1]

[1] The assumption that all private information variables have zero unconditional means is necessary for plim $\hat{\alpha}_i = 0$. For micro-homogeneity, all that is necessary is that individuals' private information variables have the same unconditional mean.

Consider the following relationship between the consensus forecast and the realization:

$$(1.5) \quad A_t \quad = \alpha + \beta P_t + \varepsilon_t,$$

where $P_t = \frac{1}{N}\sum_{i=1}^{N} P_{i,t}$ is the consensus forecast. The consensus prediction may be written as:

$$(1.6) \quad P_t \quad = X_{0t}\gamma_0 + \frac{n}{N}X_{1t}\gamma_1 + \frac{N-n}{N}X_{2t}\gamma_2,$$

where the number of type 1 and type 2 individual forecasters are n and N-n respectively.

Following the same method used to derive (1.4), the DGP for A_t may be written as:

$$\begin{aligned}(1.7) \quad A_t \quad &= X_{0t}\gamma_0 + X_{1t}\gamma_1 + X_{2t}\gamma_2 + \mu_t \\ &= (X_{0t}\gamma_0 + \frac{n}{N}X_{1t}\gamma_1 + \frac{N-n}{N}X_{2t}\gamma_2) + (\frac{N-n}{N}X_{1t}\gamma_1 + \frac{n}{N}X_{2t}\gamma_2 + \mu_t) \\ &= \alpha + \beta P_t + \varepsilon_t,\end{aligned}$$

where $\alpha = 0$, $\beta = 1$, and $\varepsilon_t = (\frac{N-n}{N}X_{1t}\gamma_1 + \frac{n}{N}X_{2t}\gamma_2 + \mu_t)$ is a mean zero, serially independent, stationary consensus forecast error. As in the case of individual forecasts, the consensus forecast error is uncorrelated with the consensus prediction, and $\hat{\varepsilon}_t$, the residuals from estimation of (1.7) are identical to the forecast errors so that plim $\hat{\alpha} = \alpha = 0$, and plim $\hat{\beta} = \beta = 1$.

In the case of rational forecasts of integrated series, OLS estimates of the slope and intercept in consensus unbiasedness regressions are consistent even in the presence of stationary private information. Unlike the results of FW (1983), there will be no bias present from the aggregation of individual, rational predictions formed on the basis of private as well as public information. Furthermore, micro-homogeneity holds, soTheil's (1954) sufficient condition for perfect linear aggregation (no PAB), is met.[2] Allowing for omitted (stationary) public information variables does not change our conclusion about the lack of PAB or PIB, or the consistency of individual or consensus estimates in

[2] See Appendix III for a proof of the sufficiency of micro-homogeneity for no PAB in unbiasedness tests on consensus forecasts.

unbiasedness regressions.[3] However, dropping the assumption of individual rationality will.

It is important to note that the assumptions regarding the stationarity, mean zero, and time independence of the private information variables and the random error in (1.2) are requirements for a forecast to possess the optimality properties associated with the strong version of the REH due to Muth. For example, if type 1 agents did not know the excess demand in region 2, and that excess demand variable was I(1), then these individuals' prediction errors, $\varepsilon_{1,t}$, would be integrated, correlated with their prediction, and $E[\varepsilon_{1,t}|\phi_{1,t}] \neq 0$, all violations of the REH. However, so long as $\phi_{i,t}$ is a proper information set in the sense that it includes lagged realizations, A_{t-j}, rational individuals could produce autoregressive forecasts with stationary errors.

[3] Omitted stationary public information does however violate the informational-efficiency property of a rational expectation. In other words, even though $E[\varepsilon_{t+k}|P_t] = 0$, $E[\varepsilon_{t+k}|\Omega_t] \neq 0$, because the omitted public I(0) variable is in both the consensus forecast error and the public information set, Ω_t, available at time t.

II. Consensus Forecasts When the DGP is Unknown

FW (1983, p. 531) argue that their conclusion—that individual unbiasedness parameter estimates are consistent, while consensus parameter estimates are inconsistent—holds even when they relax the assumption of knowledge of the true structure of the economy or the mutual independence of private information variables. The orthogonality of private information variables with the individual or consensus forecast is guaranteed by our model as long these private variables are stationary. As shown in equations (1.4) and (1.7), omitting private stationary variables does not affect the consistency of parameter estimates in unbiasedness regressions or the conclusion that individual and consensus forecasts are unbiased. Therefore, lack of knowledge of these variables or their true parameters does not affect our conclusions regarding unbiasedness tests. On the other hand, omitting either public or private integrated variables, or assuming incorrect parameters for included integrated variables will generally result in biased and informationally-inefficient individual and consensus forecasts, inconsistent parameter estimates in individual and consensus regressions, and a failure of micro-homogeneity.

Assume that type 1 and type 2 agents are unaware of the true structural parameters of the included I(1) public information. We replace equation (1.1) above with the following irrational prediction equations:

$$(2.1)\quad \begin{aligned} P_{1,t} &= X_{0t}\gamma_{1,0} + X_{1t}\gamma_1, \qquad \gamma_{1,0} = \gamma_0 - \Delta_1; \\ P_{2,t} &= X_{0t}\gamma_{2,0} + X_{2t}\gamma_2, \qquad \gamma_{2,0} = \gamma_0 - \Delta_2, \end{aligned}$$

where Δ_1, Δ_2 may be positive, negative, or zero. Following the same procedure used to derive equation (1.4), for type 1 agents we arrive at:

$$(2.2)\quad \begin{aligned} A_t &= X_{0t}\gamma_0 + X_{1t}\gamma_1 + X_{2t}\gamma_2 + \mu_t \\ &= (X_{0t}(\gamma_{1,0} + \Delta_1), + X_{1t}\gamma_1 + (X_{2t}\gamma_2 + \mu_t) \\ &= (X_{0t}\gamma_{1,0} + X_{1t}\gamma_1) + (X_{0t}\Delta_1 + X_{2t}\gamma_2 + \mu_t) \\ &= \alpha_1 + \beta_1 P_{1,t} + \varepsilon_{1,t}, \end{aligned}$$

where the forecast error, $\varepsilon_{1,t} = (X_{0t}\Delta_1 + X_{2t}\gamma_2 + \mu_t)$, is no longer mean zero, or serially uncorrelated. In fact, individuals' forecast errors are integrated and cointegrated with their predictions. Hence, $\alpha_i \neq 0$, $\beta_i \neq 1$, and OLS parameter estimates are inconsistent, with plim $\hat{\alpha}_i > 0$ and plim $\hat{\beta}_i < 1$ for $\Delta_i > 0$, and vice versa. Furthermore, unless we assume Δ_1

$= \Delta_2$, micro-homogeneity does not hold since plim $\hat{\alpha}_1 \neq$ plim $\hat{\alpha}_2$, and plim $\hat{\beta}_1 \neq$ plim $\hat{\beta}_2$. Therefore, Theil's (1954) sufficient condition for perfect aggregation is violated and PAB may occur if consensus estimation is used.

Consider the relationship between the consensus forecast and the realization given by equation (1.5), where the consensus forecast, P_t, may be written as:

$$\text{(2.3)}\quad \begin{aligned} P_t &= \frac{n}{N}P_{1,t} + \frac{N-n}{N}P_{2,t}, \\ &= \frac{n}{N}X_{0t}(\gamma_0-\Delta_1) + \frac{N-n}{N}X_{0t}(\gamma_0-\Delta_2) + \frac{n}{N}X_{1t}\gamma_1 + \frac{N-n}{N}X_{2t}\gamma_2, \\ &= \frac{N}{N}X_{0t}\gamma_0 - X_{0t}\left(\frac{n}{N}\Delta_1 + \frac{N-n}{N}\Delta_2\right) + \frac{n}{N}X_{1t}\gamma_1 + \frac{N-n}{N}X_{2t}\gamma_2. \end{aligned}$$

In (2.3), the number of type 1 and type 2 individual forecasters is n and N-n respectively.

Following the same method used to derive (1.4), the DGP for A_t may be written as:

$$\text{(2.4)}\quad \begin{aligned} A_t &= X_{0t}\gamma_0 + X_{1t}\gamma_1 + X_{2t}\gamma_2 + \mu_t \\ &= \left[X_{0t}\gamma_0 - X_{0t}\left(\frac{n}{N}\Delta_1 + \frac{N-n}{N}\Delta_2\right) + \frac{n}{N}X_{1t}\gamma_1 + \frac{N-n}{N}X_{2t}\gamma_2\right] \\ &\quad + \left[X_{0t}\left(\frac{n}{N}\Delta_1 + \frac{N-n}{N}\Delta_2\right) + \frac{N-n}{N}X_{1t}\gamma_1 + \frac{n}{N}X_{2t}\gamma_2 + \mu_t\right] \\ &= \alpha + \beta P_t + \varepsilon_t, \end{aligned}$$

where the consensus forecast error, $\varepsilon_t = X_{0t}\left(\frac{n}{N}\Delta_1 + \frac{N-n}{N}\Delta_2\right) + \frac{N-n}{N}X_{1t}\gamma_1 + \frac{n}{N}X_{2t}\gamma_2 + \mu_t$ is no longer a mean zero, serially uncorrelated, rational forecast error. In fact, the consensus errors are integrated and cointegrated with the consensus predictions. Hence, $\alpha \neq 0$, $\beta \neq 1$, and OLS estimates are inconsistent. In general, when some individuals are unaware of the true structure or parameters of the DGP, micro-homogeneity does not hold, and parameter estimates from consensus-unbiasedness regressions are inconsistent.

There are a number of possible cases which can be analyzed with the help of equations (2.2) and (2.4). The most general case is presented first, followed by three special cases:

1. $\Delta_1 \neq \Delta_2$, with both Δ_1 and Δ_2 nonzero.
   In this case, micro-homogeneity fails: $\alpha_i \neq \alpha_j \neq 0$, $\beta_i \neq \beta_j \neq 1$, and both individual and consensus coefficient estimates are inconsistent.

2. $\Delta_1 = \Delta_2$, with both Δ_1 and Δ_2 nonzero.
   In this case, micro-homogeneity holds: $\alpha_i = \alpha_j \neq 0$, $\beta_i = \beta_j \neq 1$, there is no PAB, however, both individual and consensus coefficients estimates are inconsistent.

3. $\Delta_1 = -\Delta_2$, with both Δ_1 and Δ_2 nonzero.
   In this case, micro-homogeneity fails: $\alpha_i \neq \alpha_j \neq 0$, $\beta_i \neq \beta_j \neq 1$, and both individual and consensus coefficient estimates are inconsistent.

4. $\Delta_1 = -\Delta_2$, with both Δ_1 and Δ_2 nonzero, and there are an equal number of agents of type 1 and 2.
   In this case, micro-homogeneity fails: $\alpha_i \neq \alpha_j \neq 0$, $\beta_i \neq \beta_j \neq 1$, individual coefficient estimates are inconsistent, but $\alpha=0$, $\beta=1$, and consensus estimates are consistent.[4]

Except in the unlikely event that all forecasters share exactly the same bias (case 2), micro-homogeneity fails when individuals are unaware of the true structure or parameters of the DGP. In fact, micro-homogeneity is an implication of the REH. Only in the unlikely event that individual biases average out will consensus parameters be estimated consistently even though micro-homogeneity fails.[5] In this case, the consensus and pooled regressions lead to a false acceptance of the unbiasedness hypothesis (see Figure 1 in the text).

---

[4] More generally, for k types of forecasters, if $\sum_{i=1}^{k} w_i \cdot \Delta_i = 0$, where w_i is the proportion of the sample made up of forecasters of type i, then consensus estimates are consistent and lead to false acceptance of the unbiasedness hypothesis.

[5] The conclusion that micro-homogeneity fails, but consensus parameter estimates are consistent also holds if individual's forecasts differ from unbiased predictions by only a constant. Many individual forecasts could be biased, yet individual and consensus estimates may be consistent. However, false acceptance could still result because of partial averaging out of individual biases. Furthermore, it is unclear why one would choose to form a consensus forecast (or a pooled regression) and test for unbiasedness once micro-homogeneity has been rejected.

III. Micro-Homogeneity: A Sufficient Condition For No Pure Aggregation Bias

The classic problem of aggregation bias is to define the relationship between the micro coefficients in

$$(3.1) \quad A_t = \alpha_i + \beta_i P_{i,t} + \varepsilon_{i,t}, \qquad i=1, \ldots N.$$

and the macro coefficient in

$$(3.2) \quad A_t = \alpha + \beta P_t + \varepsilon_t, \qquad t=1, \ldots, T.$$

In (3.2), $P_t = \frac{1}{N}\sum_{i=1}^{N} P_{i,t}$ is the average or "consensus" prediction.

Theil defined aggregation bias as the difference between the macro coefficient and the average of the "corresponding" micro coefficients and showed that a sufficient condition for zero (pure) aggregation bias is the equality of all individual coefficient vectors (micro-homogeneity), i.e. $\alpha_i = \alpha_j$, and $\beta_i = \beta_j$, for all i, j. To show this result for the special case of an unbiasedness regression, we follow Theil (1954) and introduce the following auxiliary regressions:

$$(3.3) \quad P_{i,t} = \omega_{0i} + \omega_{1i} P_t + \eta_{i,t}, \qquad i = 1, \cdots, N,$$

where ω_{0i} and ω_{1i} are regression coefficients, and $\eta_{i,t}$ is a residual series with the following properties: $E(\eta_{i,t}) = 0$ and $E[\eta_{i,t} P_t] = 0$.[6] While no formal economic interpretation can be provided for (3.3), it simplifies the calculation of the bias from the lack of micro-homogeneity substantially.

[6] For example, type 1 agents, whose rational prediction is given by equation (2.1), have auxiliary regressions: $X_{0t}\cdot\gamma_0 + X_{1t}\cdot\gamma_1 = \omega_{01} + \omega_{11}(X_{0t}\cdot\gamma_0 + \frac{n}{N}X_{1t}\cdot\gamma_1 + \frac{N-n}{N}X_{2t}\cdot\gamma_2) + \eta_{1,t}$. For rational individual forecasters, $E[\eta_{1,t} P_t] = 0$ as required. Note that for the case where predictions and realizations are stationary series, the auxiliary regression residual and the consensus forecast will be correlated.

Theil (1954) shows for the general case that:

$$(3.4)\quad \begin{aligned} \sum_{i=1}^{N}\omega_{0i} &= 0, \\ \sum_{i=1}^{N}\omega_{1i} &= N, \text{ and} \\ \sum_{i=1}^{N}\eta_{i,t} &= 0. \end{aligned}$$

(The second summation will equal unity where the aggregation is achieved by summation rather than averaging.) Averaging the unbiasedness regression equation (3.1) over all N individuals, we get:

$$(3.5)\quad A_t = \frac{1}{N}\sum_{i=1}^{N}\alpha_i + \frac{1}{N}\sum_{i=1}^{N}\beta_i P_{i,t} + \frac{1}{N}\sum_{i=1}^{N}\varepsilon_{i,t}.$$

Now substitute the auxiliary regression (3.3) in the above equation and collect terms.

$$(3.6)\quad A_t = \frac{1}{N}\sum_{i=1}^{N}\alpha_i + \frac{1}{N}\sum_{i=1}^{N}\beta_i\omega_{0i} + \frac{1}{N}\sum_{i=1}^{N}\beta_i\omega_{1i}P_t + \frac{1}{N}\sum_{i=1}^{N}\beta_i\eta_{i,t} + \frac{1}{N}\sum_{i=1}^{N}\varepsilon_{i,t}.$$

Comparing (3.6) with the consensus regression equation (3.2), we find:

$$(3.7)\quad \begin{aligned} \alpha &= \frac{1}{N}\sum_{i=1}^{N}[\alpha_i + \beta_i\omega_{0i}] \\ \beta &= \frac{1}{N}\sum_{i=1}^{N}\beta_i\omega_{1i} \end{aligned}$$

$$\varepsilon_t = \frac{1}{N}\sum_{i=1}^{N}[\beta_i\eta_{i,t} + \varepsilon_{i,t}]$$

Thus, in the general case, when micro-homogeneity does not hold ($\alpha_i \neq \alpha_j$, and $\beta_i \neq \beta_j$), the macro intercept is a weighted average of both the corresponding (α_i) and non-corresponding (β_i) micro parameters. The macro-slope is a weighted average of the corresponding micro-slope coefficients, β_i.

If micro-homogeneity holds, and using the results in equation (3.4), the macro parameters may be written as:

$$\begin{aligned}
\alpha &= \frac{1}{N}\sum_{i=1}^{N}\bar{\alpha} + \bar{\beta}\sum_{i=1}^{N}\omega_{0i} &&= \frac{N}{N}\bar{\alpha} = \alpha_i, \forall\, i; \\
(3.8) \quad \beta &= \bar{\beta}\frac{1}{N}\sum_{i=1}^{N}\omega_{1i} &&= \frac{N}{N}\bar{\beta} = \beta_i, \forall\, i; \\
\varepsilon_t &= \frac{1}{N}\sum_{i=1}^{N}\varepsilon_{i,t} + \bar{\beta}\sum_{i=1}^{N}\eta_{it} &&= \frac{1}{N}\sum_{i=1}^{N}\varepsilon_{i,t},
\end{aligned}$$

and there will be no aggregation bias (PAB) in the consensus regression. Micro-homogeneity is a sufficient condition for no PAB. Theil also shows that micro-homogeneity is not a necessary condition.