On the Wisdom of Crowds (of Economists)

Francis X. Diebold
University of Pennsylvania and NBER

Aarón Mora
University of South Carolina

Minchul Shin
Federal Reserve Bank of Philadelphia

On the Wisdom of Crowds
(of Economists)

Francis X. Diebold

University of Pennsylvania

and NBER

Aarón Mora

University of South Carolina

Minchul Shin

Federal Reserve Bank of Philadelphia

February 6, 2026

Abstract: We study the properties of macroeconomic survey forecast response averages as the number of survey respondents grows. Such averages are "portfolios" of forecasts. We characterize the speed and pattern of the gains from diversification as a function of portfolio size (the number of survey respondents) in both (1) the key real-world data-based environment of the U.S. Survey of Professional Forecasters, and (2) the theoretical model-based environment of equicorrelated forecast errors. We proceed by proposing and comparing various direct and model-based "crowd size signature plots," which summarize the forecasting performance of k-average forecasts as a function of k, where k is the number of forecasts in the average. We then estimate the equicorrelation model for growth and inflation forecast errors by choosing model parameters to minimize the divergence between direct and model-based signature plots. The results indicate near-perfect equicorrelation model fit for both growth and inflation, which we explicate by showing analytically that, under very weak conditions, the direct and fitted equicorrelation model-based signature plots are identical at a particular model parameter configuration. That parameter configuration immediately suggests an analytic closed-form estimator for the direct signature plot, so that equicorrelation ultimately emerges as a device for convenient calculation of direct signature plots, rather than a separate "model" producing separate signature plots. Finally, we find that the gains from survey diversification are greater for inflation forecasts than for growth forecasts, and that they are largely exhausted with inclusion of 5-10 representative forecasters.

Contents

1 Introduction and Basic Framework

The wisdom of crowds, or lack thereof, is traditionally and presently a central issue in psychology, history, and political science; see for example Surowiecki (2005) regarding wisdom, and Aliber et al. (2023) regarding lack thereof. Perhaps most prominently, however, the wisdom of crowds is also—and again, traditionally and presently—a central issue in economics and finance, where heterogeneous information and expectations formation play a prominent role.[1]

In this paper we focus on economics and finance, studying the "wisdom" of "crowds" of professional economists. We focus on the U.S. Survey of Professional Forecasters (SPF), which is important not only in facilitating empirical academic research in macroeconomics and financial economics, but also—and crucially—in guiding real-time policy, business, and investment management decisions.[2,3]

In particular, we study SPF crowd behavior as crowd size grows, asking precisely the same sorts of questions of SPF "forecast portfolios" that one asks of financial asset portfolios:[4] How quickly, and with what patterns, do diversification benefits become operative, and eventually dissipate, as portfolio size (the number of forecasters) grows, and why? Do the results differ across the key SPF variables (growth and inflation), and if so, why? What are the implications for survey size and design?

We answer the above questions using what we call "crowd size signature plots," which summarize the forecasting performance of k-average forecasts as a function of k, where k is the number of forecasts in the average. We examine both "direct signature plots" (estimated directly from the SPF data) and "model-based signature plots" (estimated by fitting a simple model of forecast-error equicorrelation), and we characterize their surprisingly intimate relationship, uncovering en route a remarkably simple analytic closed-form signature plot estimator. We focus throughout on estimating and interpreting signature plots for both growth and inflation, providing individual and comparative assessments of their paths and patterns.

---

[1]Interestingly, moreover, it also features prominently in new disciplines like machine learning and artificial intelligence, via forecast combination methods like ensemble averaging (e.g., Diebold et al., 2023).

[2]On real-time policy and its evaluation, see for example John Tayor's inaugural NBER Feldstein Lecture at https://www.hoover.org/sites/default/files/gmwg-empirically-evaluating-economic-policy-in-real-time.pdf.

[3]For an introduction to the SPF, see the materials at https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/survey-of-professional-forecasters.

[4]Classic early work on which we build includes, for example, Makridakis and Winkler (1983) and Batchelor and Dua (1995). We progress much further, however, particularly regarding analytic characterization.

Related Literature. This paper contributes to the forecast combination literature by examining the performance patterns of simple k-average forecasts as the number of forecasters, k, increases. We note that this paper is not focused on the problem of optimal forecast combination (see Bates and Granger (1969) for a pioneering contribution, and for example Timmermann (2006), Hsiao and Wan (2014) and Wang, Hyndman, Li, and Kang (2022) for recent developments). The motivation for focusing on simple k-averages is twofold. First, it is a well-established empirical fact that simple averages of forecasts produce combined forecasts with surprisingly good forecasting performance despite not being optimal in general (e.g., Clemen (1989), Stock and Watson (2004), and Genre, Kenny, Meyler, and Timmermann (2013)).[5] Second, simple averages can be optimal in specific contexts. For example, Engle and Kelly (2012) show that when forecast errors exhibit equicorrelation, equal weights are optimal; Elliott and Liao (2025) later provide precise necessary and sufficient conditions under which simple averages are optimal.

Our focus on performance patterns of simple k-averages as k increases builds on earlier empirical work. Makridakis and Winkler (1983), using the M-Competition time-series dataset and combinations of forecasts across multiple forecasting methods, document that simple-average combinations become more accurate, with sharply diminishing marginal gains, as the number of component forecasts increases. We find the same qualitative pattern in the SPF when averaging across professional forecasters. We extend these results by providing a theoretical foundation: we derive an exact expression for the direct crowd size signature plot that links its level and slope to two interpretable features of SPF forecast errors (average dispersion and average dependence). This delivers an equivalent equicorrelation representation and explains both the hyperbolic shape of the signature plots and the rapid exhaustion of diversification gains.

In another related paper, Chan and Pauwels (2018) propose a framework to study the properties of a combined forecast, and derive conditions under which the simple average is optimal. In their setting, they point out that the performance of a combined forecast improves as the number of forecasts or models increases, and that a simple average will asymptotically outperform any fixed individual forecast. Although this is consistent with our results, our emphasis is more general, as we characterize the entire trajectory of performance improvements, including the practically relevant setting of a small to moderate number of forecasters.

---

[5]This empirical fact has been called a "forecast combination puzzle"; work focused on studying the causes of this puzzle includes Smith and Wallis (2009), Elliott and Liao (2025), and Claeskens, Magnus, Vasnev, and Wang (2016).

Our results on relating direct crowd size signature plots to an equicorrelation model-based signature plot are connected to the results in Clemen and Winkler (1985). In the case of normally distributed forecast errors, they show that the value of additional information for forecast combination diminishes when the additional information is correlated. This is consistent with our findings documenting diversification gains; however, our results go beyond this by not requiring the normality assumption.

Finally, we contribute to the SPF survey-forecast literature (e.g., Croushore and Stark, 2019; Clements et al., 2022) by providing evidence and theory directly relevant for survey design: how many forecasters are needed to capture most of the benefits of averaging, and why do those benefits differ across variables? Using SPF growth and inflation forecasts, we document a steeply diminishing-returns pattern in crowd-size performance, with most improvements realized by about 5-10 representative forecasters. Moreover, we show that diversification gains are considerably larger for inflation than for growth, and we provide an exact theoretical characterization linking the inflation-growth difference to the underlying variance-covariance structure of individual forecast errors.

We proceed as follows. In section 2 we study the SPF, estimating its crowd size signature plots directly, for both growth and inflation. In section 3 we introduce and estimate the equicorrelation model, beginning in section 3.1 by characterizing its crowd size signature plots analytically for any parameter configuration, and continuing in section 3.2 by estimating its parameters by minimizing divergence between direct and model-based signature plots, again providing empirical results for both growth and inflation. In section 4 we conclude and sketch several directions for future research.

2 Direct SPF Crowd Size Signature Plots

In this section we introduce the idea of crowd size signature plots, and we calculate them directly for SPF point forecasts of real output growth ("growth") and GDP deflator inflation ("inflation"), for forecast horizons $h = 1, 2, 3,$ and 4, corresponding to short-, medium-, and longer-term forecasts.[6] Our sample period is 1968Q4-2023Q2.

The SPF was started in 1968Q4 and is currently conducted and maintained by the Federal

---

[6]The SPF contains quarterly level forecasts of real GDP and the GDP implicit price deflator. We transform the level forecasts into growth and inflation forecasts by computing annualized quarter-on-quarter growth rates, and we compute the corresponding forecast errors using realized values as of December 2023. See Appendix A for details.

Figure 1: SPF Participation
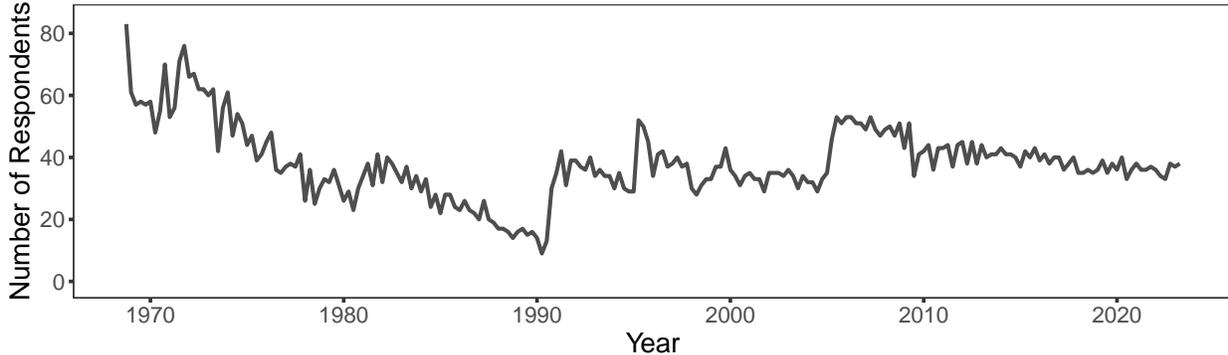
Notes: We show the number of participants in the U.S. Survey of Professional Forecasters, 1968Q4-2023Q2.

Reserve Bank of Philadelphia.[7] In Figure 1 we show the evolution of the number of forecast participants, which declined until 1990Q2, when the Federal Reserve Bank of Philadelphia took control of the survey, after which it has had approximately 40 participants. Participants stayed for 15 quarters on average, with a minimum of 1 quarter and a maximum of 125 quarters.

We us now sketch the basic framework. Let N refer to a set of forecasts with $N \times 1$ zero-mean time-t error vector e_t, $t = 1, ..., T$, and let $k \leq N$ refer to a subset of forecasts. We consider k-forecast averages, and we seek to characterize k-forecast mean-squared forecast error (MSE). For a particular k-forecast average corresponding to group g_k^*, the forecast error is just the average of the individual forecast errors, so we have

$$\widehat{MSE}_T^*(k) = \frac{1}{T} \sum_{t=1}^{T} \left(\frac{1}{k} \sum_{i \in g_k^*} e_{it} \right)^2.$$
(1)

For any choice of k, however, there are $\binom{N}{k}$ possible k-forecast averages. We focus on the k-average $\widehat{MSE}_T^*(k)$ given by equation (1), averaged across all groups of size k,

$$\widehat{MSE}_{NT}^{avg}(k) = \frac{1}{\binom{N}{k}} \sum_{g_k=1}^{\binom{N}{k}} \left(\frac{1}{T} \sum_{t=1}^{T} \left(\frac{1}{k} \sum_{i \in g_k} e_{it} \right)^2 \right),$$
(2)

where g_k is an arbitrary member of the set of groups of size k.

---

[7]For a recent introduction see Croushore and Stark (2019).

Among other things, one may be interested in:

1. Tracking and visualizing $\widehat{MSE}_{NT}^{avg}(k)$ as k grows ("$\widehat{MSE}_{NT}^{avg}(k)$ crowd size signature plots");

2. Tracking and visualizing *functions* of $\widehat{MSE}_{NT}^{avg}(k)$ as k grows, such as:

   (a) the performance from k-averaging ($\widehat{MSE}_{NT}^{avg}(k)$) relative to the performance from no averaging ($\widehat{MSE}_{NT}^{avg}(1)$),

$$\widehat{MSE}_{R,NT}^{avg}(k) = \frac{\widehat{MSE}_{NT}^{avg}(k)}{\widehat{MSE}_{NT}^{avg}(1)} \tag{3}$$

   ("$\widehat{MSE}_{R,NT}^{avg}(k)$ crowd size signature plots," where "R" denotes "relative," or "ratio");

   (b) the *change* (improvement) in $\widehat{MSE}_{NT}^{avg}(k)$ from adding one more forecast to the pool (i.e., moving from k to $k+1$ forecasts), e.g.,

$$\widehat{DMSE}_{NT}^{avg}(k) = \widehat{MSE}_{NT}^{avg}(k) - \widehat{MSE}_{NT}^{avg}(k+1) \tag{4}$$

   ("$\widehat{DMSE}_{NT}^{avg}(k)$ crowd size signature plots," where "D" denotes "delta," or "change"), or, in ratio form,

$$\widehat{DMSE}_{R,NT}^{avg}(k) = \frac{\widehat{DMSE}_{NT}^{avg}(k)}{\widehat{DMSE}_{NT}^{avg}(1)} \tag{5}$$

   ("$\widehat{DMSE}_{R,NT}^{avg}(k)$ crowd size signature plots");

3. Tracking and visualizing not just the *mean* of the squared-error distribution as k grows, as in all of the above signature plots, but rather the *complete* distribution ("$\widehat{F}_{NT}^{avg}(k)$ signature plots", or in ratio form, "$\widehat{F}_{R,NT}^{avg}(k)$ signature plots");

4. Estimating equicorrelation models for SPF forecast errors by minimizing divergence between direct and model-based SPF signature plots, and assessing their fit;

5. Examining the paths and patterns in the above estimated signature plots, and similarities/differences for growth vs. inflation;

6. Drawing practical implications for SPF design.

The "ratio" signature plots, especially the $\widehat{MSE}_{R,NT}^{avg}(k)$ plot defined in equation (3), deserve special mention. Scaling by $\widehat{MSE}_{NT}^{avg}(1)$ (the benchmark MSE corresponding to no averaging) makes $\widehat{MSE}_{R,NT}^{avg}(1) \equiv 1$, which facilitates $\widehat{MSE}_{NT}^{avg}(k)$ comparisons across different variables like growth and inflation. For that reason we use it extensively in the graphical presentations that follow.

$\widehat{MSE}_{NT}^{avg}(k)$ and the related objects above are readily calculated in principle, but complications arise in practice. In particular, calculation is infeasible unless N is very small, due to the huge number of different k-average forecasts. For example, for $N = 40$ and $k = 20$ we obtain $\binom{N}{k} = O(10^{11})$. Hence we proceed by approximating $\widehat{MSE}_{NT}^{avg}(k)$ as follows:

Algorithm 1.

1. *For each period $t = 1, ..., T$, randomly select k forecasts and average their forecast errors; refer to this k-average forecast group as $g_{k,t}^*$;*

2. *Average the forecast errors for $g_{k,t}^*$ across sample periods to calculate $\widehat{MSE}_T^*(k)$;*

3. *Repeat B times, and average the $\widehat{MSE}_T^*(k)$ values across the B draws, where B is large, but not so large as to be computationally intractable.*

As $B \to \infty$, the above approximation will converge to $\widehat{MSE}_{NT}^{avg}(k)$, and in this paper we set $B = 30,000$. The approximation works, moreover, for unbalanced as well as balanced panels, which is important because the SPF panel is unbalanced due to entry and exit of forecasters.

We show direct crowd size signature plots for growth and inflation in Figure 2 $(\widehat{MSE}_{R,NT}^{avg}(k))$, Figure 3 $(\widehat{DMSE}_{R,NT}^{avg}(k))$, and Figure 4 $(\widehat{F}_{R,NT}^{avg}(k))$. Several features are apparent, as distilled in the following remarks:

Remark 1. For both growth and inflation, the $\widehat{MSE}_{R,NT}^{avg}(k)$ signature plot is lowest for $h = 1$, with the signature plots for $h = 2, 3,$ and 4 progressively farther above the $h = 1$ plot in roughly parallel upward shifts, reflecting the fact that the near future is generally easier to predict than the more-distant future.

Remark 2. The growth $\widehat{MSE}_{R,NT}^{avg}(k)$ signature plot does not rise much from $h = 3$ to $h = 4$, in contrast to the inflation $\widehat{MSE}_{R,NT}^{avg}(k)$ signature plot, suggesting that growth predictability drops with horizon more quickly than inflation predictability, effectively vanishing by $h = 3$.

Figure 2: Direct $\widehat{MSE}_{R,NT}^{avg}(k)$ Crowd Size Signature Plots
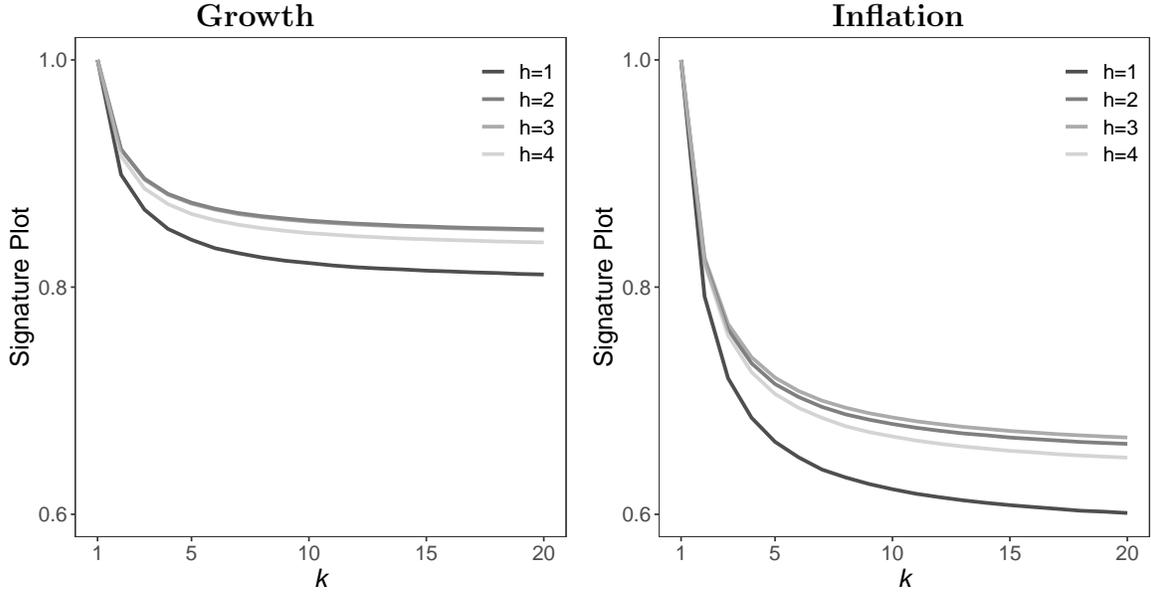
Notes: We show direct $\widehat{MSE}_{R,NT}^{avg}(k)$ crowd size signature plots for SPF growth and inflation forecasts at horizon $h = 1, 2, 3, 4$, for group sizes $k = 1, 2, ..., 20$. For each k, we produce the figure by randomly drawing $B = 30,000$ groups of size k for each $t = 1, ..., T$.

Remark 3. For both growth and inflation and all forecast horizons, the reduction in $\widehat{MSE}_{R,NT}^{avg}(k)$ from $k = 1$ to $k = 5$ dwarfs the improvement from moving from $k = 6$ to $k = 20$, as visually emphasized by the $\widehat{DMSE}_{R,NT}^{avg}(k)$ signature plots in Figure 3. Hence there is little benefit from adding representative forecasters to the pool beyond $k = 5$.

Remark 4. Individual examination of the growth or inflation $\widehat{DMSE}_{R,NT}^{avg}(k)$ signature plots reveals approximately identical behavior at all horizons (i.e., no upward shifts), which is expected because the $\widehat{MSE}_{R,NT}^{avg}(k)$ signature plots shift with horizon in approximately parallel fashion, leaving its "first derivative" plot ($\widehat{DMSE}_{R,NT}^{avg}(k)$) approximately unchanged.

Remark 5. Comparison of the growth and inflation $\widehat{DMSE}_{R,NT}^{avg}(k)$ signature plots reveals identical marginal gains from adding additional representative forecasters. This phenomenon is not accidental, and we shall return to it subsequently in section 3.

Remark 6. Both the growth and inflation $\widehat{F}_{R,NT}^{avg}(k)$ plots show right-skewness in forecast-error distributions for small k, but the distributions become less variable and more symmetric as k grows and the law of large numbers and central limit theorem operate. The small-k skewness is much greater, however, for the inflation forecast-error distribution.

7

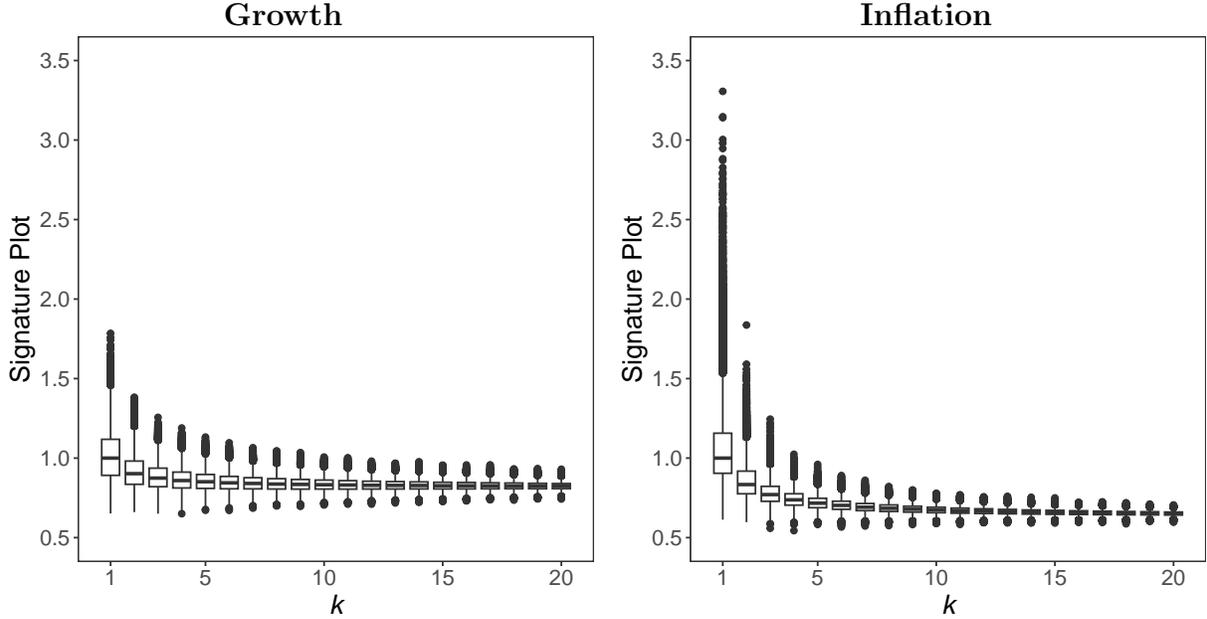Figure 3: Direct $\widehat{DMSE}_{R,NT}^{avg}(k)$ Crowd Size Signature Plots
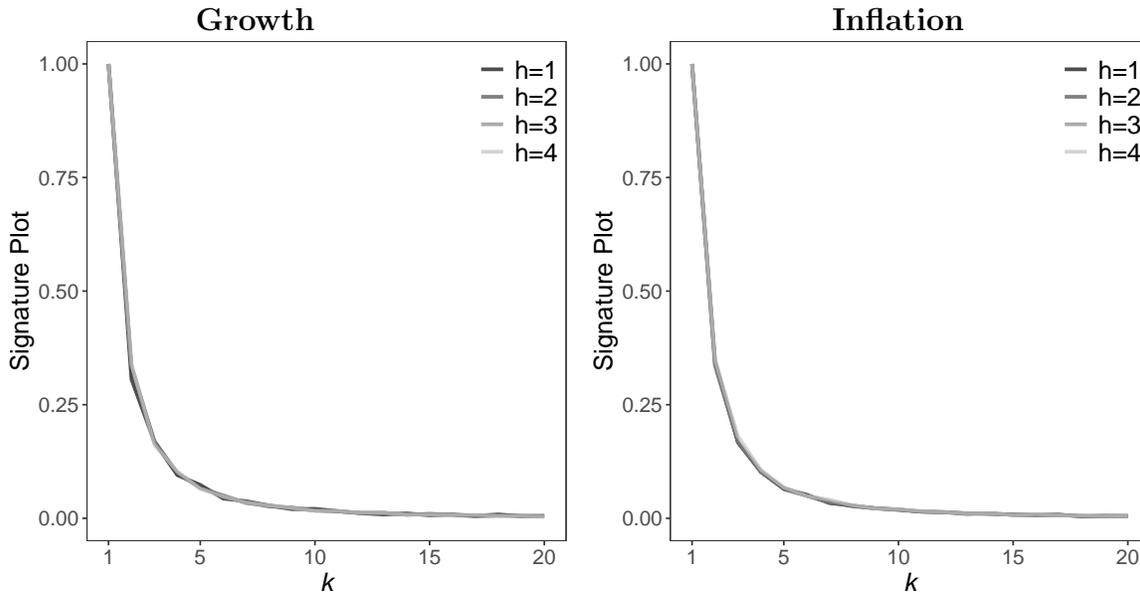
Notes: We show direct $\widehat{DMSE}_{R,NT}^{avg}(k)$ crowd size signature plots for SPF growth and inflation forecasts at horizons $h = 1, 2, 3, 4$, for group sizes $k = 1, 2, ..., 20$. For each k, we produce the figure by randomly drawing $B = 30,000$ groups of size k for each $t = 1, ..., T$.

Remark 7. The growth vs. inflation $\widehat{MSE}_{R,NT}^{avg}(k)$ plots asymptote to very different levels as k increases – approximately 80% for growth and 60% for inflation – which highlights a central result: The benefits of SPF "portfolio diversification" appear to be substantially greater for inflation than for growth.

Figures 2 and 4, together with Remarks 6 and 7, are closely linked and merit additional and joint discussion. Figure 4 and Remark 6 show a markedly stronger consensus in growth forecasts than in inflation forecasts (compare, for example, the growth and inflation forecast distributions in Figure 4 for $k = 1$). This pattern accords with the broader empirical macroeconomics literature, in which low-ordered autoregressions typically provide hard-to-beat descriptions of short-term U.S. growth (e.g., Stock and Watson, 2002; Chauvet and Potter, 2013), whereas inflation dynamics are more complex, warranting consideration of a wider set of models (e.g., Stock and Watson, 2007; Pivetta and Reis, 2007).[8] In this light, Figure 2 and Remark 7 make perfect sense: the greater heterogeneity in inflation forecasts produces

---

[8]A similar situation appears in the agent-based modeling literature, where simulated agents struggle to learn a stable inflation model and often adopt heterogeneous specifications (e.g., Arifovic et al., 2013; LeBaron and Smith, 2025).

Figure 4: Direct $\widehat{F}_{R,NT}^{avg}(k)$ Crowd Size Signature Plots

Notes: We show direct $\widehat{F}_{R,NT}^{avg}(k)$ crowd size signature plots for SPF growth and inflation forecasts. The signature plots are composed of boxplots of squared 1-step-ahead k-average forecast errors, $k = 1, 2, ..., 20$. Each boxplot displays the median, the first and third quartiles, the lower extreme value (first quartile minus 1.5 times interquartile range), the upper extreme value (third quartile plus 1.5 times the interquartile range), and outliers. All growth boxplots and all inflation boxplots are scaled by their respective medians at $k = 1$, so that both the growth and inflation medians equal 1.0 when $k = 1$. For each k, we produce the figure by randomly drawing $B = 30,000$ groups of size k for each $t = 1, ..., T$.

greater gains from inflation forecast diversification, yielding a strikingly lower $\widehat{MSE}_{R,NT}^{avg}(k)$ signature-plot asymptote for inflation (about 60%) than for growth (about 80%).

3 Model-Based SPF Crowd Size Signature Plots

Having empirically characterized crowd size signature plots directly in the SPF data, we now proceed to characterize them analytically in a simple covariance-stationary equicorrelation model, in which $e_t \sim (0, \Sigma)$, where 0 is the $N \times 1$ zero vector and Σ is an $N \times N$ forecast-error covariance matrix displaying equicorrelation, by which we mean that all variances are identical and all correlations are identical.

3.1 Population Results Under Equicorrelation

A trivial equicorrelation example occurs when $\Sigma = \sigma^2 I$, where I denotes the $N \times N$ identity matrix, so that all variances are equal (σ^2), and all correlations are equal (0). Of course the zero-correlation case is unrealistic, because, for example, economic forecast errors are invariably positively correlated due to overlap of information sets, but it will serve as a useful benchmark, so we begin with it.

Simple averaging is the fully optimal forecast combination in the zero-correlation environment, which is obvious since the forecasts are exchangeable. More formally, the optimality of simple averaging (equal combining weights) follows from the multivariate Bates and Granger (1969) formula for MSE-optimal combining weights,

$$\lambda^* = \left(\iota' \Sigma^{-1} \iota \right)^{-1} \Sigma^{-1} \iota, \tag{6}$$

where ι is a k-dimensional column vector of ones. For $\Sigma = \sigma^2 I$ the optimal weights collapse to

$$\lambda^* = (\sigma^{-2} N)^{-1} \sigma^{-2} \iota = \frac{1}{N} \iota.$$

Analytical asymptotic results are straightforward for simple averages in the zero-correlation environment. Let $MSE_{N\infty}^{avg}(k; \sigma) = plim_{T \to \infty} \left(\widehat{MSE}_{NT}^{avg}(k; \sigma) \right)$; then we have

$$MSE_{N\infty}^{avg}(k; \sigma) = plim_{T \to \infty} \left(\frac{1}{\binom{N}{k}} \sum_{g_k=1}^{\binom{N}{k}} \left(\frac{1}{T} \sum_{t=1}^{T} \left(\frac{1}{k} \sum_{i \in g_k} e_{it} \right)^2 \right) \right) \tag{7}$$

$$= \frac{1}{\binom{N}{k}} \sum_{g_k=1}^{\binom{N}{k}} \left(E \left[\left(\frac{1}{k} \sum_{i \in g_k} e_{it} \right)^2 \right] \right)$$

$$= \frac{1}{\binom{N}{k}} \sum_{g_k=1}^{\binom{N}{k}} \frac{1}{k^2} E \left[\left(\sum_{i \in g_k} e_{it} \right)^2 \right]$$

$$= \frac{1}{\binom{N}{k}} \sum_{g_k=1}^{\binom{N}{k}} \frac{1}{k^2} \sum_{i \in g_k} E \left[e_{it}^2 \right]$$

$$= \frac{\sigma^2}{k}.$$

Moreover,

$$DMSE_{N\infty}^{avg}(k;\sigma) = \frac{\sigma^2}{k(k+1)} \tag{8}$$

and

$$MSE_{R,N\infty}^{avg}(k) = \frac{1}{k}. \tag{9}$$

Notice that σ cancels in the $R_{N\infty}^{avg}(k;\sigma)$ calculation, so we simply write $R_{N\infty}^{avg}(k)$.

We now move to a richer equicorrelation case with equal but nonzero correlations, so that instead of $\Sigma = \sigma^2 I$ we have

$$\Sigma = \sigma^2 R, \tag{10}$$

where

$$R = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}, \tag{11}$$

and $\rho \in \left]\frac{-1}{N-1}, 1\right[$.[9] Recent work, in particular Engle and Kelly (2012), has made use of equicorrelation in the context of modeling multivariate financial asset return volatility.

Importantly, the optimality of simple averaging under zero correlation is preserved under equicorrelation. That is, equicorrelation is sufficient for the optimality of simple averaging – an immediate implication of the results of Elliott and Liao (2025), who shows that a necessary and sufficient condition for optimality of simple averaging is that row sums of Σ be equal. Equicorrelation is one such case, although there are of course others, obtained by manipulating correlations in their relation to variances to keep row sums equal, but none are nearly so compelling and readily interpretable as equicorrelation.

To understand the optimality of simple averaging under equicorrelation in the context of Bates and Granger (1969), just as we did earlier under zero correlation, consider the inverse covariance matrix in the Bates-Granger expression for the optimal combining weight vector, (6). In the equicorrelation case we have

$$\Sigma^{-1} = \frac{1}{\sigma^2} R^{-1}, \tag{12}$$

---

[9]R is positive definite if and only if $\rho \in \left]\frac{-1}{N-1}, 1\right[$. See Lemma 2.1 of Engle and Kelly (2012).

where[10]

$$R^{-1} = \frac{1}{1-\rho}I - \frac{\rho}{(1-\rho)(1+(N-1)\rho)}\iota\iota'. \tag{13}$$

Then, using equation (13), the first part of the optimal combining weight (6) is

$$\iota'\Sigma^{-1}\iota = \frac{N}{\sigma^2}\frac{(1+(N-1)\rho)-\rho N}{(1-\rho)(1+(N-1)\rho)}, \tag{14}$$

and the second part is

$$\Sigma^{-1}\iota = \frac{1}{\sigma^2}\frac{(1+(N-1)\rho)-\rho N}{(1-\rho)(1+(N-1)\rho)}\iota. \tag{15}$$

Inserting equations (14) and (15) into equation (6) again yields

$$\lambda^* = \frac{1}{N}\iota, \tag{16}$$

establishing the optimality of equal weights.

Having now introduced equicorrelation and shown that it implies optimality of simple average forecast combinations, it is of interest to motivate it – despite its stark simplicity – in terms of economic considerations. First, obviously but importantly, the information sets of economic forecasters are quite highly overlapping, so it is not necessarily unreasonable to suppose that the forecast error variances are roughly equal, and that pairs of forecast errors are roughly equally correlated.

Second, less obviously but very importantly, equicorrelation is closely linked to factor structure, which is a great workhorse of modern macroeconomic and business-cycle analysis (e.g., Stock and Watson, 2016). In particular, equicorrelation arises when forecast errors have single-factor structure with equal factor loadings and equal idiosyncratic shock variances, as for example in:

$$e_{it} = \delta z_t + w_{it} \tag{17}$$
$$z_t = \phi z_{t-1} + v_t,$$

where $w_{it} \sim iid(0, \sigma_w^2)$, $v_t \sim iid(0, \sigma_v^2)$, and $w_{it} \perp v_t$, $\forall i, t$, $i = 1, ..., N$, $t = 1, ..., T$.

Finally, a large literature from the 1980s onward documents and exploits the routine outstanding empirical performance of simple average forecast combinations, despite the fact that simple averages are not optimal in general (e.g., Clemen, 1989; Genre et al., 2013; Elliott and Timmermann, 2016; Diebold and Shin, 2019). As we have seen, however, equicorrelation is sufficient (and almost necessary) for optimality of simple averages, so that if simple

---

[10]See Lemma 2.1 of Engle and Kelly (2012).

averages routinely perform well, then the equicorrelation model is routinely reasonable – and a natural model to pair with the simple averages embodied in the SPF.

Analytical results for $MSE_{N\infty}^{avg}(\cdot)$, $DMSE_{N\infty}^{avg}(\cdot)$ and $R_{N\infty}^{avg}(\cdot)$ are easy to obtain under equicorrelation, just as they were under zero correlation. Immediately,

$$MSE_{N\infty}^{avg}(k; \rho, \sigma) = plim_{T \to \infty} \left(\frac{1}{\binom{N}{k}} \sum_{g_k=1}^{\binom{N}{k}} \left(\frac{1}{T} \sum_{t=1}^{T} \left(\frac{1}{k} \sum_{i \in g_k} e_{it} \right)^2 \right) \right) \quad (18)$$

$$= \frac{1}{\binom{N}{k}} \sum_{g_k=1}^{\binom{N}{k}} \left(E \left[\left(\frac{1}{k} \sum_{i \in g_k} e_{it} \right)^2 \right] \right)$$

$$= \frac{1}{\binom{N}{k}} \sum_{g_k=1}^{\binom{N}{k}} \frac{1}{k^2} E \left[\left(\sum_{i \in g_k} e_{it} \right)^2 \right]$$

$$= \frac{1}{\binom{N}{k}} \sum_{g_k=1}^{\binom{N}{k}} \frac{1}{k^2} \left[k\sigma^2 + k(k-1)Cov(e_{it}, e_{jt}) \right] \quad \text{for} \ i \neq j$$

$$= \frac{\sigma^2}{k} \left[1 + (k-1)\rho \right].$$

In addition, for MSE changes and ratios we have

$$DMSE_{N\infty}^{avg}(k; \rho, \sigma) = \frac{\sigma^2}{k(k+1)}(1-\rho) \quad (19)$$

and

$$MSE_{R,N\infty}^{avg}(k; \rho, \sigma) = MSE_{R,N\infty}^{avg}(k; \rho) = \frac{1}{k} \left[1 + (k-1)\rho \right]. \quad (20)$$

One can immediately verify that if $\rho = 0$ then the equicorrelation signature plot results (18)-(20) collapse to the corresponding earlier zero-correlation results (7)-(9).

In Figure 5 we show $MSE_{R,N\infty}^{avg}(k; \rho)$ as a function of k, for various equicorrelations, ρ. Several remarks are in order:

Remark 8. From equation (20), the $MSE_{R,N\infty}^{avg}(k; \rho)$ curves decrease to a limiting value (ρ) as the combining pool grows ($k \to \infty$). Indeed the gains from increasing k are initially large but decrease hyperbolically quickly. The MSE improvement, for example, in moving from $k = 1$ to $k = 5$ consistently dwarfs that of moving from $k = 5$ to $k = 20$.

13

Figure 5: Theoretical Equicorrelation $MSE_{R,N\infty}^{avg}(k;\rho)$ Crowd Size Signature Plots
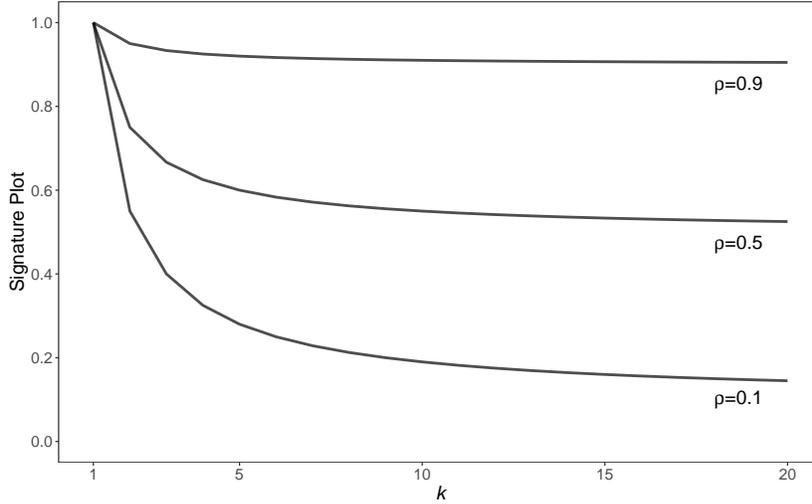
Notes: We show theoretical equicorrelation $MSE_{R,N\infty}^{avg}(k;\rho)$ crowd size signature plots for equicorrelations $\rho = 0.1, 0.5, 0.9$ and group sizes $k = 1, ..., 20$.

Remark 9. Overall, the value of averaging across forecasters is determined by the interaction of k and ρ. Other things equal, it is highest when k is small (small pool) or when ρ is low (weakly correlated forecast errors). In particular, for realistic values of ρ, around 0.5, say, most gains from increasing k are obtained by $k = 5$.

Remark 10. The fact that, for realistic values of ρ, most $MSE_{R,N\infty}^{avg}(k;\rho)$ gains from increasing k are obtained by $k = 5$ does *not* necessarily indicate that typical surveys use too many forecasters. $MSE_{R,N\infty}^{avg}(k;\rho)$ is an *average* across all k-forecast combinations, and the best and worst k-average combinations, for example, will have very different *MSE*s. Figure 6 speaks to this; it shows $MSE_{R,N\infty}^{min}(k;\rho)$, $MSE_{R,N\infty}^{avg}(k;\rho)$, and $MSE_{R,N\infty}^{max}(k;\rho))$ under equicorrelation with $\rho = 0.5$, for $k = 1, ..., 20$.[11]

Remark 11. The equicorrelation case is the only one for which analytic results are readily obtainable. For example, even if we maintain the assumption of equal correlations but simply allow different forecast-error variances ("weak equicorrelation"), the MSE of the k-person average forecast becomes a function of $(k, \rho, \sigma_1^2, ..., \sigma_N^2)$, and little more can be said.[12]

---

[11]We use $N = 40$ as an approximation to the average number of forecasters participating in the SPF in any given quarter, and we use $T = 160$ to mimic the total sample size when working with 40 years of quarterly data, as in the SPF.

[12]See Appendix B for derivation of optimal combining weights in the weak equicorrelation case.

Figure 6: Theoretical Equicorrelation $MSE_{R,N\infty}^{min}(k;.5)$, $MSE_{R,N\infty}^{avg}(k;.5)$, and $MSE_{R,N\infty}^{max}(k;.5)$ Crowd Size Signature Plots
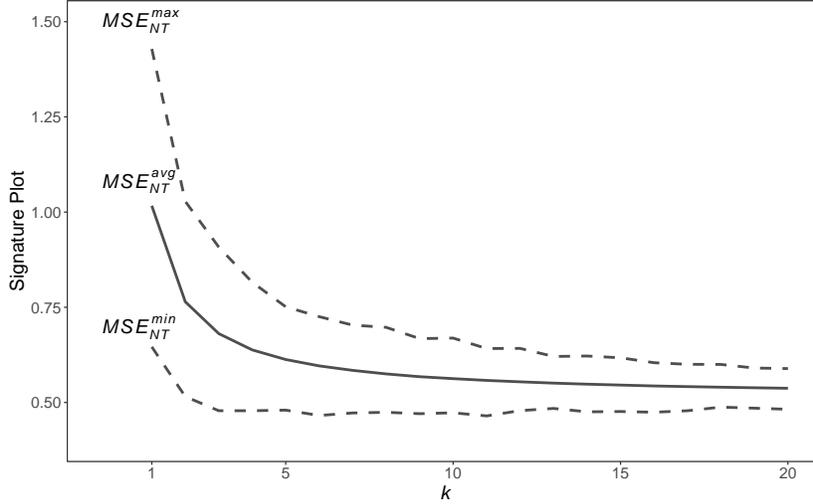
Notes: The data generating process is the equicorrelation model given by equations (10)-(11), with $\rho = .5$, $\sigma = 1$, $N = 40$, and $T = 160$. For each group size k, 30,000 groups of size k were drawn at random for each $t = 1, ..., T$ to produce the figure.

Remark 12. As mentioned earlier, the equicorrelation case naturally matches the provision of survey averages, because in that case simple averages are optimal. Hence, as we now proceed to a model-based empirical analysis of real forecasters, we work with the equicorrelation model, asking what values of ρ and σ make the equicorrelation model-based $MSE_{N\infty}^{avg}(k;\rho,\sigma)$ signature plot as close as possible to the direct $\widehat{MSE}_{NT}^{avg}(k)$ signature plot.

3.2   Estimating the Equicorrelation Model

Here we estimate the equicorrelation model by choosing its parameters ρ and σ to make the equicorrelation $MSE_{N\infty}^{avg}(k;\rho,\sigma)$ as close as possible to the SPF $\widehat{MSE}_{NT}^{avg}(k)$. This estimation strategy is closely related to, but different from, GMM estimation. Rather than matching model and data moments, it matches more interesting and interpretable functions of those moments, namely model-based and direct crowd size signature plots – as per the "indirect inference" of Smith Jr (1993) and Gourieroux et al. (1993). Henceforth we refer to it simply as the "matching estimator."

Specifically, we solve for $(\hat{\rho}, \hat{\sigma})$ such that

$$(\hat{\rho}, \hat{\sigma}) = \arg\min_{(\rho,\sigma)} Q(\rho, \sigma), \tag{21}$$

where

$$Q(\rho, \sigma) = \frac{1}{N} \sum_{k=1}^{N} \left(\widehat{MSE}_{NT}^{avg}(k) - MSE_{N\infty}^{avg}(k; \rho, \sigma) \right)^2,$$

and the minimization is constrained such that $\sigma > 0$ and $\rho \in \left] \frac{-1}{N-1}, 1 \right[$.

3.2.1 Calculating the Solution

Calculating the minimum in equation (21) is very simple, because the bivariate minimization can be reduced to a univariate minimization in $\rho \in \left] \frac{-1}{N-1}, 1 \right[$. To see this, recall that $MSE_{N\infty}^{avg}(k; \rho, \sigma) = \frac{\sigma^2}{k}[1 + (k-1)\rho]$, so that the first-order condition for σ^2 is

$$\frac{1}{N} \sum_{k=1}^{N} \left(\widehat{MSE}_{NT}^{avg}(k) - MSE_{N\infty}^{avg}(k; \rho, \sigma) \right) \left(\frac{1}{k} + \frac{k-1}{k}\rho \right) = 0, \tag{22}$$

and the first-order condition for ρ is

$$\frac{1}{N} \sum_{k=1}^{N} \left(\widehat{MSE}_{NT}^{avg}(k) - MSE_{N\infty}^{avg}(k; \rho, \sigma) \right) \left(1 - \frac{1}{k} \right) \sigma^2 = 0. \tag{23}$$

Combining equations (22) and (23) yields

$$\sigma^2 = \frac{c_1}{c_2 + c_3 \rho}, \tag{24}$$

where $c_1 = \sum_{k=1}^{N} \frac{\widehat{MSE}_{NT}^{avg}(k)}{k}$, $c_2 = \sum_{k=1}^{N} \frac{1}{k^2}$ and $c_3 = \sum_{k=1}^{N} \frac{k-1}{k^2}$. Hence, at the optimum and conditional on the data, there is a deterministic inverse relationship between ρ and σ (see Figure 7), enabling one to restrict the parameter search to the small open interval $\rho \in \left] \frac{-1}{N-1}, 1 \right[$, as well as to explore the objective function visually as a function of ρ alone (see Figure 8).

In Table 1 we show the complete set of estimates (for σ^2 and ρ, for growth and inflation, for $h = 1, ...4$). The $\hat{\rho}$ estimates accord with the earlier-discussed asymptotes of the direct signature plots, and the $\hat{\sigma}^2$ estimates increase with forecast horizon, reflecting the fact that the distant future is harder to forecast than the near future, and implying that

16

Figure 7: Illustration of the Relationship Between ρ and σ^2
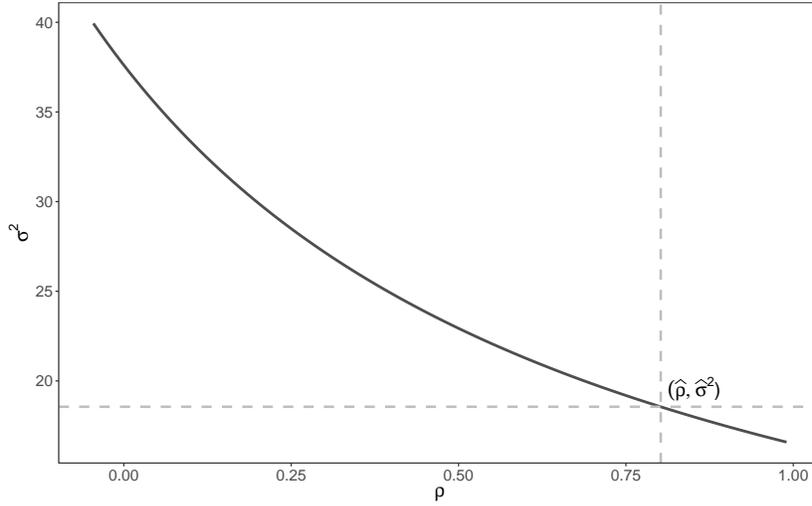
Notes: We show the relationship between ρ and σ^2 given by equation (24), for 1-step-ahead growth forecast errors. The highlighted values of $\hat{\rho} = 0.8$ and $\hat{\sigma}^2 = 18.6$ are the estimated values for growth forecasts at horizon $h = 1$. See Table 1.

Figure 8: Illustration of the Objective Function and its Minimum
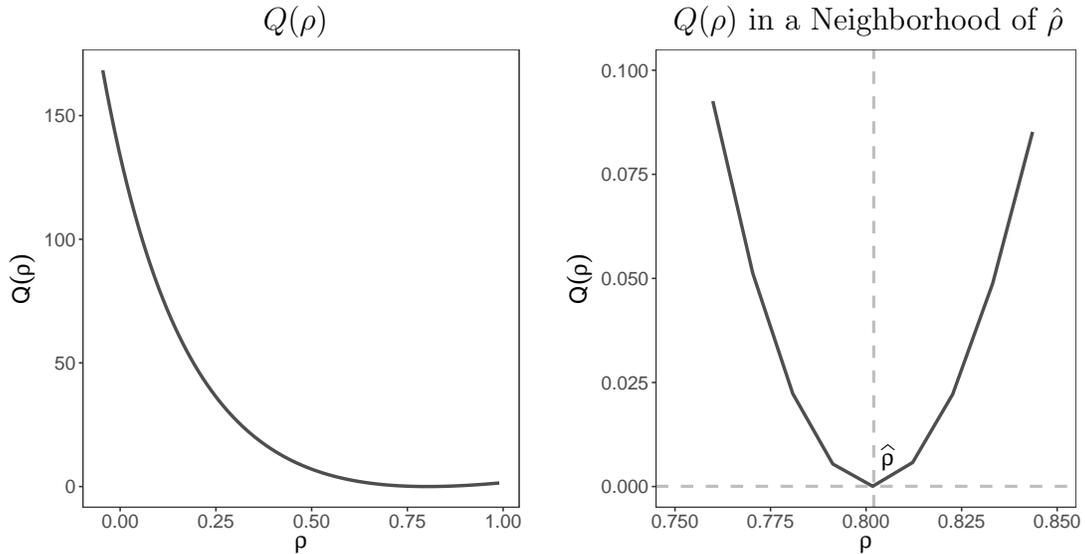
Notes: We show the objective function $Q(\rho)$ of the matching estimator expressed as a function of ρ, for 1-step-ahead growth forecast errors. The highlighted value of $\hat{\rho} = 0.8$ is the estimated value for growth forecasts at horizon $h = 1$. See Table 1.

Table 1: Equicorrelation Model Estimates

Growth

|	$h = 1$	$h = 2$	$h = 3$	$h = 4$
$\hat{\sigma}^2$	18.562	21.170	22.713	23.275
	(0.606)	(0.546)	(0.589)	(0.646)
$\hat{\rho}$	0.801	0.843	0.842	0.831
	(0.036)	(0.028)	(0.028)	(0.030)
$MSE^{avg}_{R,N\infty}(1;\hat{\rho})$	1.000	1.000	1.000	1.000
$MSE^{avg}_{R,N\infty}(5;\hat{\rho})$	0.841	0.874	0.874	0.865
$MSE^{avg}_{R,N\infty}(15;\hat{\rho})$	0.815	0.853	0.853	0.842
$Q(\hat{\sigma}^2, \hat{\rho})$	5.99E-05	3.95E-06	7.95E-06	1.01E-05

Inflation

	$h = 1$	$h = 2$	$h = 3$	$h = 4$
$\hat{\sigma}^2$	3.662	4.343	5.123	6.094
	(0.253)	(0.255)	(0.295)	(0.370)
$\hat{\rho}$	0.580	0.644	0.650	0.630
	(0.082)	(0.068)	(0.066)	(0.071)
$MSE^{avg}_{R,N\infty}(1;\hat{\rho})$	1.000	1.000	1.000	1.000
$MSE^{avg}_{R,N\infty}(5;\hat{\rho})$	0.664	0.715	0.720	0.704
$MSE^{avg}_{R,N\infty}(15;\hat{\rho})$	0.608	0.668	0.673	0.655
$Q(\hat{\sigma}^2, \hat{\rho})$	2.21E-06	9.56E-07	1.82E-06	4.36E-05

Notes: We show equicorrelation model parameter estimates for SPF growth and inflation forecast errors at various horizons, with standard errors computed via 1000 bootstrap samples. We also show estimated relative MSE with respect to no averaging, $MSE^{avg}_{R,N\infty}(k;\hat{\rho})$ for $k = 1, 5, 15$. In the final line of each panel we show the value of the objective function evaluated at the estimated parameters, $Q(\hat{\sigma}^2, \hat{\rho})$.

Figure 9: Direct $\widehat{MSE}_{R,NT}^{avg}(k)$ and Estimated Equicorrelation $MSE_{R,N\infty}^{avg}(k; \hat{\rho})$ Crowd Size Signature Plots
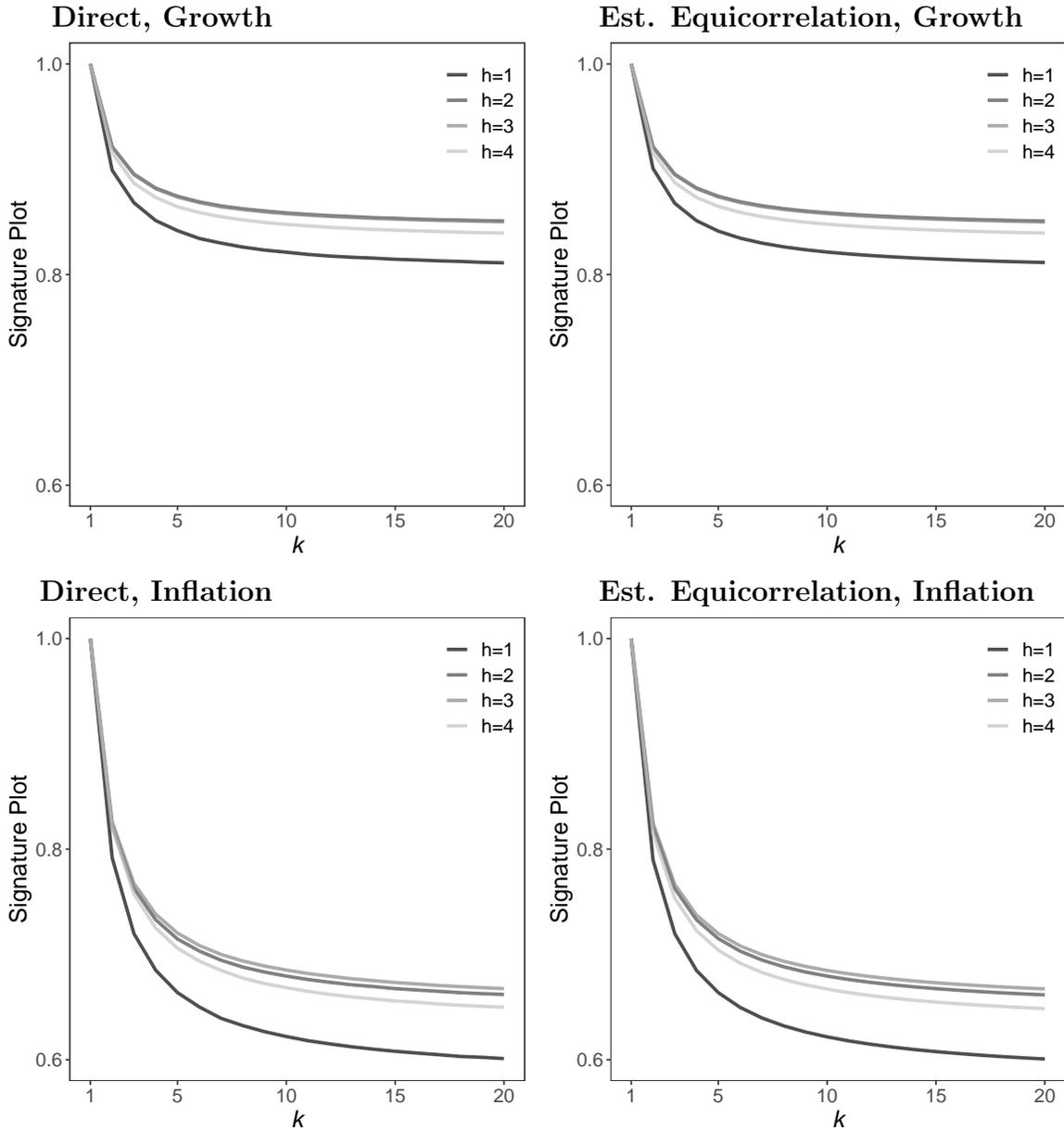
Notes: We show direct and estimated equicorrelation ratio crowd size signature plots $(\widehat{MSE}_{R,NT}^{avg}(k)$ and $MSE_{R,N\infty}^{avg}(k; \hat{\rho})$, respectively) for growth and inflation forecasts at horizons $h = 1, 2, 3, 4$, for group sizes $k = 1, 2, ..., 20$.

fitted model-based equicorrelation signature plots should shift upward with horizon. In Figure 9 we show side-by-side direct $\widehat{MSE}_{NT}^{avg}(k)$ (left column) and equicorrelation model-based $MSE_{R,N\infty}^{avg}(k;\hat{\rho})$ (right column) signature plots, which reveal a remarkably good equicorrelation model fit – the direct and model-based plots are effectively identical.

3.2.2 Understanding the Near-Perfect Equicorrelation Fit

Here we present a closed-form solution for the direct crowd size signature plot. The result is significant in its own right and reveals why our numerical matching estimates for the equicorrelation model produce fitted signature plots that align so closely with direct signature plots. To maintain precision it will prove useful to state it as a formal theorem.

Theorem 1. *Let e_t be any covariance stationary $N \times 1$ vector with mean zero and covariance matrix Σ, given by*

$$\Sigma = \begin{pmatrix} \sigma_1^2 & c_{12} & c_{13} & \cdots & c_{1N} \\ c_{21} & \sigma_2^2 & c_{23} & \cdots & c_{2N} \\ c_{31} & c_{32} & \sigma_3^2 & \cdots & c_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{N1} & c_{N2} & c_{N3} & \cdots & \sigma_N^2 \end{pmatrix},$$

and define the population k-average MSE,

$$MSE_{N\infty}^{avg}(k) = \frac{1}{\binom{N}{k}} \sum_{g_k=1}^{\binom{N}{k}} E\left[\left(\frac{1}{k} \sum_{i \in g_k} e_{it} \right)^2 \right],$$

where g_k represents any subset of e_t of size k ($k \in [1, N]$). Then

$$MSE_{N\infty}^{avg}(k) = \frac{\overline{\sigma^2}}{k} \left(1 + (k-1)\overline{\rho} \right) \tag{25}$$

and

$$MSE_{R,N\infty}^{avg}(k) = \frac{MSE_{N\infty}^{avg}(k)}{MSE^{avg}(1)} = \frac{1}{k} \left(1 + (k-1)\overline{\rho} \right), \tag{26}$$

where

$$\overline{\sigma^2} = \frac{1}{N} \sum_{i=1}^{N} \sigma_i^2$$

$$\overline{c} = \frac{1}{\binom{N}{2}} \sum_{1 \leq i < j \leq N} c_{ij}$$

$$\overline{\rho} = \frac{\overline{c}}{\overline{\sigma^2}}.$$

Proof. We have:

$$MSE_{N\infty}^{avg}(k) = \frac{1}{\binom{N}{k}} \sum_{g_k=1}^{\binom{N}{k}} E\left[\left(\frac{1}{k}\sum_{i\in g_k} e_{it}\right)^2\right]$$

$$= \frac{1}{k^2}\frac{1}{\binom{N}{k}} \sum_{g_k=1}^{\binom{N}{k}} \left[\sum_{i\in g_k} \sigma_i^2 + \sum_{i,j\in g_k, i\neq j} c_{ij}\right]$$

$$= \frac{1}{k^2}\frac{1}{\binom{N}{k}} \left[\underbrace{\sum_{g_k=1}^{\binom{N}{k}}\sum_{i\in g_k} \sigma_i^2}_{\text{grand sum of variances}} + \underbrace{\sum_{g_k=1}^{\binom{N}{k}}\sum_{i,j\in g_k, i\neq j} c_{ij}}_{\text{grand sum of covariances}}\right].$$

First consider the term related to variance. Note that, of the $\binom{N}{k}$ groups, there are $\binom{N-1}{k-1}$ groups that include e_i^2. The sum of $\sum_{i\in g_k}\sigma_i^2$ over all groups g_k is therefore

$$\sum_{g_k=1}^{\binom{N}{k}}\sum_{i\in g_k}\sigma_i^2 = \binom{N-1}{k-1}\sum_{i=1}^{N}\sigma_i^2. \tag{27}$$

Now consider the term related to covariance. When summing $\sum_{i,j\in g_k, i\neq j} c_{ij}$ across all groups g_k, covariances between all possible pairs of e_{it} and e_{jt} are accounted for. Because we are summing over all arbitrary groups g_k, each pair (i,j) appears the same number of times in the grand summation. To compute this number we observe that each group g_k contains $k(k-1)$ pairwise covariances and that there are $\binom{N}{k}$ possible groups. Hence the total number of individual covariance terms in the grand sum is $k(k-1)\binom{N}{k}$. The number of times that each individual covariance term c_{ij} appears in the grand sum is $\frac{k(k-1)\binom{N}{k}}{\binom{N}{2}}$, where $\binom{N}{2}$ is the total number of distinct pairs (i,j). The grand sum of covariances is therefore

$$\frac{1}{\binom{N}{k}} \sum_{g_k=1}^{\binom{N}{k}}\sum_{i,j\in g_k, i\neq j} c_{ij} = \frac{k(k-1)\binom{N}{k}}{\binom{N}{2}} \sum_{i\leq i<j\leq N} c_{ij}. \tag{28}$$

Combining equations (27) and (28), we have:

$$MSE_{N\infty}^{avg}(k) = \frac{1}{k^2} \frac{1}{\binom{N}{k}} \binom{N-1}{k-1} \sum_{i=1}^{N} \sigma_i^2 + \frac{1}{k^2} \frac{1}{\binom{N}{k}} \frac{k(k-1)\binom{N}{k}}{\binom{N}{2}} \sum_{i \leq i < j \leq N} c_{ij}$$

$$= \frac{1}{k} \left[\left(\frac{1}{N} \sum_{i=1}^{N} \sigma_i^2 \right) + (k-1) \left(\frac{1}{\binom{N}{2}} \sum_{i \leq i < j \leq N} c_{ij} \right) \right]$$

$$= \frac{\overline{\sigma^2}}{k} \left[1 + (k-1)\overline{\rho} \right].$$

This completes the proof. $\square$

Several remarks are in order:

Remark 13. Theorem 1 reveals that in large samples the direct crowd size signature plot is simply the equicorrelation model-based signature plot evaluated at particular values of the equicorrelation model parameters ρ and σ, *regardless of whether the forecast errors are truly equicorrelated*. Hence in large samples the two plots will always match perfectly. The same identity holds exactly in finite samples upon replacing expectations by time averages, using $\widehat{\sigma}_i^2 = T^{-1} \sum_{t=1}^{T} e_{it}^2$ and $\widehat{c}_{ij} = T^{-1} \sum_{t=1}^{T} e_{it} e_{jt}$ when the panel is balanced.

Remark 14. Theorem 1 also immediately suggests an alternative, closed-form, matching estimator for the equicorrelation model:

$$\widehat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^{N} \widehat{\sigma}_i^2 \tag{29}$$

$$\widehat{\rho} = \frac{\frac{1}{\binom{N}{2}} \sum_{1 \leq i < j \leq N} \widehat{c}_{ij}}{\frac{1}{N} \sum_{i=1}^{N} \widehat{\sigma}_i^2}. \tag{30}$$

Remark 15. Combining Remarks 13 and 14, it emerges that, regardless of whether one chooses to approach the signature plot from the "direct side" or from the "equicorrelation side," there is no need for simulation (as earlier from the direct side, when implementing the direct estimator via Algorithm 1) or numerical optimization (as earlier from the equicorrelation side, when implementing the matching estimator via solution of the first-order conditions (22)-(23)). Instead, the trivially simple closed-form estimator (29)-(30) is always applicable.

Let us now provide a characterization of $DMSE_{N\infty}^{avg}(k)$ (and $DMSE_{R,N\infty}^{avg}(k)$), which of course follows immediately from our earlier characterization of $MSE_{N\infty}^{avg}(k)$ (and $MSE_{R,N\infty}^{avg}(k)$). Immediately,

Corollary 1. *Under the conditions of Theorem 1, we have that*

$$DMSE_{N\infty}^{avg}(k) = \frac{\overline{\sigma^2}(1-\overline{\rho})}{k(k+1)},$$

so that

$$DMSE_{R,N\infty}^{avg}(k) = \frac{DMSE^{avg}(k)}{DMSE^{avg}(1)} = \frac{2}{k(k+1)}.$$

In particular, because $DMSE_{R,N\infty}^{avg}(k)$ depends only on k and not on the data, it will be identical for all variables forecast.

Remark 16. We earlier observed that the empirical $DMSE_{R,N\infty}^{avg}(k)$ signature plots for growth and inflation appeared identical. The corollary shows this is not a data-specific coincidence. Theoretically, the two curves *must* coincide for every k.

Remark 17. Similarly, for both growth and inflation we earlier observed negligible diversification gains from expanding the forecast pool beyond $k = 10$ (say), and the corollary shows why. By the time we reach $k = 10$, the marginal gain from expanding the pool drops to just 1.8% of the initial marginal gain $(2/[10(10 + 1)] = 0.018)$.

3.3   On the Assessment of Equicorrelation Model Fit

As we have seen, equicorrelation ultimately emerges as a device for convenient calculation and understanding of direct signature plots, rather than as a separate "model" producing separate signature plots. In particular, nothing we have done requires that the equicorrelation be "true." Nevertheless, it may be of interest in some contexts to assess whether forecast errors are truly equicorrelated, which can be done (under stronger assumptions than those invoked thus far) by maximum-likelihood estimation of a dynamic-factor model, followed by likelihood-ratio tests of the equicorrelation restrictions.

Consider in particular a standard dynamic single-factor model (DFM),

$$e_{it} = \delta_i z_t + w_{it} \tag{31}$$

$$z_t = \phi z_{t-1} + v_t,$$

where $w_{it} \sim iid\, N(0, \sigma_{wi}^2)$, $v_t \sim iid\, N(0, \sigma_v^2)$, and $w_{it} \perp v_t$, $\forall i, t$. The implied forecast-error covariance matrix Σ fails to satisfy equicorrelation; that is,

$$\Sigma \neq \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix},$$

because the forecast-error variances generally vary with i, and their correlations generally vary with i and j. In particular, simple calculations reveal that

$$\sigma_i^2 \equiv var(e_{i,t}) = \delta_i^2 var(z_t) + \sigma_{wi}^2$$
$$= \delta_i^2 \left(var(z_t) + \frac{\sigma_{wi}^2}{\delta_i^2} \right), \ \forall i,$$

where $var(z_t) = \frac{\sigma_v^2}{1-\phi^2}$, and

$$\rho_{ij} \equiv corr(e_{i,t}, e_{j,t}) = \frac{\delta_i \delta_j var(z_t)}{\sqrt{\delta_i^2 var(z_t) + \sigma_{wi}^2} \sqrt{\delta_j^2 var(z_t) + \sigma_{wj}^2}}$$
$$= \frac{1}{\sqrt{1 + \frac{\sigma_{wi}^2}{\delta_i^2 var(z_t)}} \sqrt{1 + \frac{\sigma_{wj}^2}{\delta_j^2 var(z_t)}}}, \ \forall i, j. \quad (32)$$

Nevertheless, certain simple restrictions on the DFM (31) produce certain forms of equicorrelation. First, it is apparent from equation (32) that $\rho_{ij} = \rho$, $\forall i \neq j$, if and only if

$$\frac{\sigma_{wi}^2}{\delta_i^2} = \frac{\sigma_{wj}^2}{\delta_j^2}, \ \forall i \neq j, \quad (33)$$

so that imposition of the constraint (33) on equation (31) produces a "weak" form of equicorrelation with identical correlations (ρ) but allowing for potentially different idiosyncratic

shock variances $(\sigma_1^2, ..., \sigma_N^2)$. That is,

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho & \cdots & \rho \\ \rho & \sigma_2^2 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & \sigma_N^2 \end{pmatrix}.$$

Second, it is also apparent from equation (32) that if we impose the stronger restriction,

$$\sigma_{wi}^2 = \sigma_{wj}^2 \quad \text{and} \quad \delta_i = \delta_j, \ \forall i, j, \tag{34}$$

which of course implies the weaker restriction (33), then we obtain ("strong") equicorrelation as we have defined it throughout this paper, with identical correlations (ρ) *and* identical idiosyncratic shock variances (σ^2). That is,

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}.$$

The DFM (31) is already in state-space form, and one pass of the Kalman filter yields the innovations needed for exact Gaussian likelihood evaluation and optimization, and it also accounts for missing observations associated with survey entry and exit. One may also impose the weak or strong equicorrelation restrictions (33) or (34), respectively, and assess them using likelihood-ratio tests. Unsurprisingly, such formal tests (not reported) reject equicorrelation – a stylized "toy" data-generating process if ever there was one – for both growth and inflation, so we will not pursue formal testing.

More interestingly, we will provide a preliminary assessment of the *size* of deviations from equicorrelation, for both growth and inflation, quite apart from assessing whether they happen to be exactly zero. We construct a balanced panel with no entry or exit, 2010Q1-2019Q4 (40 quarters), containing ten forecasts for growth and inflation, and we examine their covariance matrices, as shown in Figure 10.[13]

First consider growth forecast errors, in the upper panel of Figure 10. No shading means that the absolute deviation of the object in the cell (a variance or covariance) from the median

---

[13]The forecasts are ordered from lowest to highest error variance for each variable.

Figure 10: Growth and Inflation Forecast Error Variances and Covariances, Percent Deviations from Median

Growth

Inflation
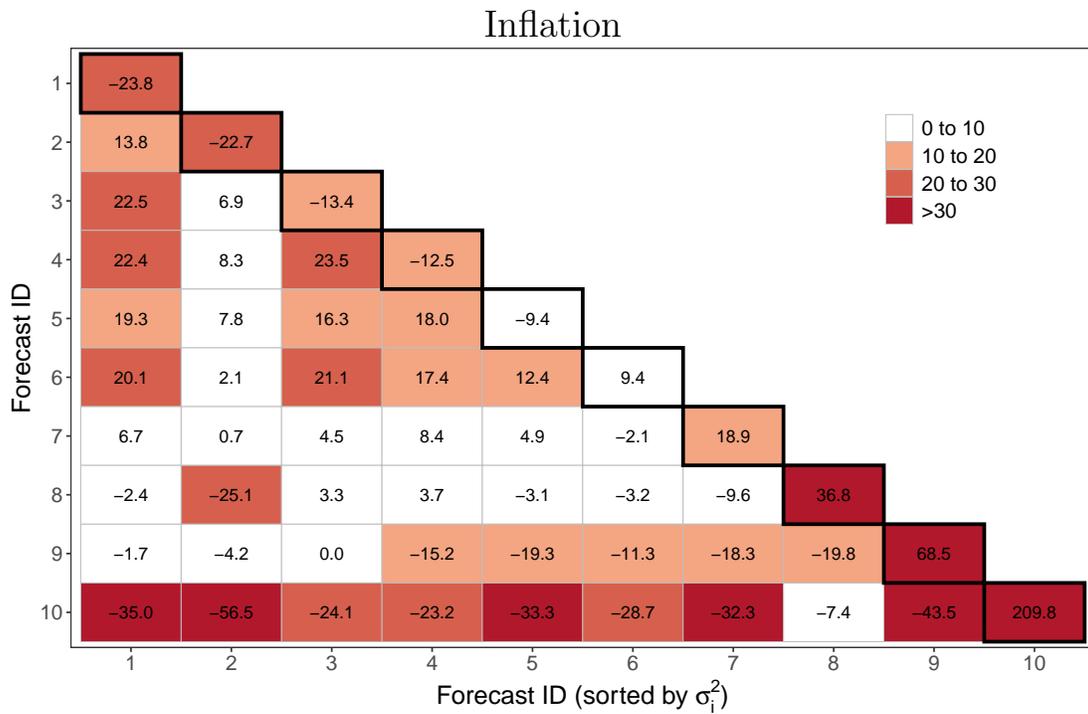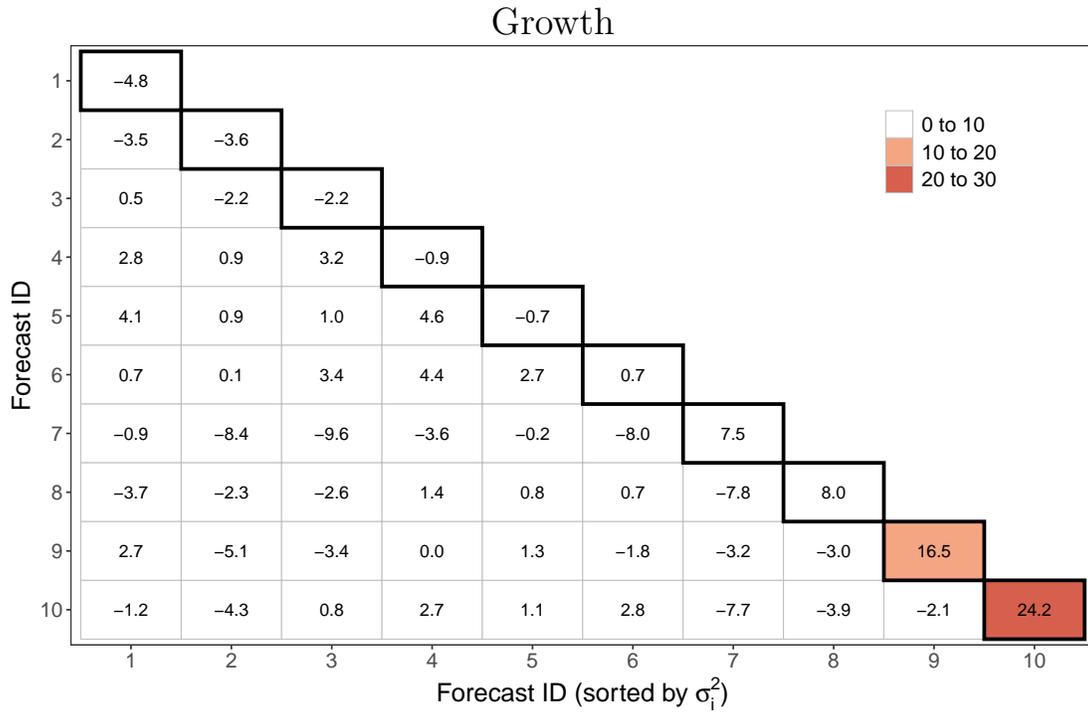
Notes: We work with a balanced panel of ten forecasters, 2010Q1-2019Q4 (40 quarters). See text for details.

across all cells (all variances or all covariances) is less than 10%. For example, the forecast with ID 3 has a variance of growth forecast error that is -2.2% below the median variance, so its cell is unshaded. Similarly, light red shading means that the absolute deviation from the median is 10-20%, and darker red shading means that the absolute deviation from the median is 20-30%. All growth covariance cells are white, as are all but two variance cells. Hence, at least by our intentionally rough metric, equicorrelation appears to be not too bad an approximation for growth forecast errors.

Now consider inflation forecast errors, in the lower panel of Figure 10. The shading convention is the same as earlier, but now there is a fourth, dark red, category corresponding to >30%. The situation is sharply different, with many red cells, for both variances and covariances. This echoes our earlier discussion of the relative lack of consensus in inflation forecasts due to the difficulty of determining an appropriate inflation model, as per Figures 2 and 4, together with Remarks 6 and 7.

4   Summary, Conclusions, and Directions for Future Research

We have studied the properties of macroeconomic survey forecast response averages as the number of survey respondents grows, characterizing the speed and pattern of the "gains from diversification" and their eventual decrease with "portfolio size" (the number of survey respondents) in both (1) the key real-world data-based environment of the U.S. Survey of Professional Forecasters (SPF), and (2) the theoretical model-based environment of equicorrelated forecast errors. We proceeded by proposing and comparing various direct and model-based "crowd size signature plots," which summarize the forecasting performance of k-average forecasts as a function of k, where k is the number of forecasts in the average. We then estimated the equicorrelation model for growth and inflation forecast errors by choosing model parameters to minimize the divergence between direct and model-based signature plots.

The results indicate near-perfect equicorrelation model fit for both growth and inflation, which we explicated by showing analytically that, under conditions, the direct and fitted equicorrelation model-based signature plots are identical at a particular model parameter configuration, which we characterize. We find that the gains from diversification are greater for inflation forecasts than for growth forecasts, but that both the inflation and growth diversification gains nevertheless decrease quite quickly (which we also explain analytically),

so that fewer SPF respondents than currently used may be adequate.

Several directions for future research appear promising, including, in no particular order:

1. Instead of considering MSEs across $\binom{N}{k}$ possible k-average forecasts and averaging to obtain a "representative k-average" forecast MSE as a function of k, one may want to consider "*best k-average*" forecast MSE as a function of k, where the unique best k-average forecast is obtained in each period as the k-average that performed best historically.

2. One may want to allow for time-varying equicorrelation parameters, as σ^2 might, for example, move downward with the Great Moderation, while ρ might move counter-cyclically. The strong equicorrelation model in dynamic-factor form becomes

$$e_{it} = \delta_t z_t + w_{it}$$

$$z_t = \phi z_{t-1} + v_t,$$

where $w_{it} \sim iid(0, \sigma_{wt}^2)$, $v_t \sim iid(0, \sigma_v^2)$, and $w_{it} \perp v_t, \forall i, t$. Immediately,

$$e_t \sim iid\left(0,\ \Sigma_t(\rho_t)\right),$$

where

$$\Sigma_t(\rho_t) \;=\; \sigma_t^2 \begin{pmatrix} 1 & \rho_t & \cdots & \rho_t \\ \rho_t & 1 & \cdots & \rho_t \\ \vdots & \vdots & \ddots & \vdots \\ \rho_t & \rho_t & \cdots & 1 \end{pmatrix}$$

$$\sigma_t^2 = \delta_t^2 var(z_t) + \sigma_{wt}^2$$

$$\rho_t = \frac{\delta_t^2 var(z_t)}{\delta_t^2 var(z_t) + \sigma_{wt}^2}.$$

3. One may want to complement our exploration of the U.S. SPF with a comparative exploration of the European SPF.[14] Doing so appears feasible but non-trivial, due to cross-survey differences in sample periods, economic indicator concepts (e.g., inflation), and timing conventions.

---

[14]For an introduction to the European SPF, see the materials at `https://data.ecb.europa.eu/methodology/survey-professional-forecasters-spf`.

Appendices

A   Data Definitions and Sources

We obtain U.S. quarterly level forecasts of real output and the GDP deflator from the Federal Reserve Bank of Philadelphia's *Individual Forecasts: Survey of Professional Forecasters* (variables $RGDP$ and $PGDP$, respectively). We transform the level forecasts into annualized growth rate forecasts using:

$$g_{t+h|t-1} = 100 \left(\left(\frac{f_{t+h|t-1}}{f_{t+h-1|t-1}} \right)^4 - 1 \right), \tag{A1}$$

where $f_{t+h|t-1}$ is a quarterly level forecast (either $RGDP$ or $PGDP$) for quarter $t+h$ made using information available in quarter $t-1$. For additional information, see `https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/individual-forecasts`.

We obtain the corresponding realizations from the Federal Reserve Bank of Philadelphia's *Forecast Error Statistics for the Survey of Professional Forecasters* (December 2023 vintage). The realizations are reported as annualized growth rates, as in equation (A1) above, so there is no need for additional transformation. For additional information, see `https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/error-statistics`.

B   Optimal Combining Weights Under Weak Equicorrelation

Here we briefly consider the "weak equicorrelation" case, with correlation ρ and variances σ_1^2, σ_2^2, ..., σ_N^2; that is,

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho & \cdots & \rho \\ \rho & \sigma_2^2 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & \sigma_N^2 \end{pmatrix},$$

where $\rho \in \left] \frac{-1}{N-1}, 1 \right[$. We can decompose the covariance matrix Σ as

$$\Sigma = DRD = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N \end{pmatrix} \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N \end{pmatrix},$$

where R is positive definite if and only if $\rho \in \left] \frac{-1}{N-1}, 1 \right[$. The inverse of the covariance matrix is

$$\Sigma^{-1} = D^{-1} R^{-1} D^{-1} = \begin{pmatrix} \sigma_1^{-1} & 0 & \cdots & 0 \\ 0 & \sigma_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N^{-1} \end{pmatrix} \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix}^{-1} \begin{pmatrix} \sigma_1^{-1} & 0 & \cdots & 0 \\ 0 & \sigma_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N^{-1} \end{pmatrix},$$

where

$$R^{-1} = \frac{1}{1-\rho} I - \frac{\rho}{(1-\rho)(1+(N-1)\rho)} \iota \iota',$$

I stands for an $N \times N$ identity matrix, and ι is a N-vector of ones.

Recall that, as noted in the text, the optimal combining weight is

$$\lambda^* = \left(\iota' \Sigma^{-1} \iota \right)^{-1} \Sigma^{-1} \iota, \tag{C1}$$

The first part of the optimal combining weight (C1) is

$$\iota' \Sigma^{-1} \iota = \frac{(1+(N-1)\rho) \left(\sum_{i=1}^{N} \sigma_i^{-2} \right) - \rho \left(\sum_{i=1}^{N} \sigma_i^{-1} \right) \left(\sum_{i=1}^{N} \sigma_i^{-1} \right)}{(1-\rho)(1+(N-1)\rho)}, \tag{C2}$$

and the second part is

$$\Sigma^{-1} \iota = \frac{(1+(N-1)\rho)\boldsymbol{\sigma}^{-2} - \rho \left(\sum_{i=1}^{N} \sigma_i^{-1} \right) \boldsymbol{\sigma}^{-1}}{(1-\rho)(1+(N-1)\rho)}, \tag{C3}$$

where

$$\boldsymbol{\sigma}^{-2} = \begin{pmatrix} \sigma_1^{-2} \\ \sigma_2^{-2} \\ \vdots \\ \sigma_N^{-2} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\sigma}^{-1} = \begin{pmatrix} \sigma_1^{-1} \\ \sigma_2^{-1} \\ \vdots \\ \sigma_N^{-1} \end{pmatrix}.$$

Inserting equations (C2) and (C3) into equation (C1), we get the optimal weight for the ith forecast as

$$\lambda_i^* = \frac{\sigma_i^{-2} + \rho(N-2)\sigma_i^{-2} - \rho\left(\sum_{j\neq i}\sigma_i^{-1}\sigma_j^{-1}\right)}{\sum_{i=1}^{N}\left(\sigma_i^{-2} + \rho(N-2)\sigma_i^{-2} - \rho\left(\sum_{j\neq i}\sigma_i^{-1}\sigma_j^{-1}\right)\right)}. \tag{C4}$$

To check the formula, note that for $N = 2$ we obtain the standard Bates and Granger (1969) optimal bivariate combining weight,

$$\lambda_1^* = \frac{\sigma_1^{-2} - \rho\sigma_1^{-1}\sigma_2^{-1}}{\sigma_1^{-2} + \sigma_2^{-2} - 2\rho\sigma_1^{-1}\sigma_2^{-1}} = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2},$$

and for any N, but with $\sigma_j^2 = \sigma^2 \ \forall j$ (equicorrelation), we obtain weights,

$$\lambda_i^* = \frac{(1-\rho)\sigma^{-2}}{N(1-\rho)\sigma^{-2}} = \frac{1}{N}, \quad \forall i.$$

References

Aliber, R.Z., C.P. Kindleberger, and R.N. McCauley (2023), *Manias, Panics, and Crashes: A History of Financial Crises*, 8th Edition, Palgrave MacMillan.

Arifovic, J., J. Bullard, and O. Kostyshyna (2013), "Social Learning and Monetary Policy Rules," *The Economic Journal*, 123, 38–76.

Batchelor, R. and P. Dua (1995), "Forecaster Diversity and the Benefits of Combining Forecasts," *Management Science*, 41, 68–75.

Bates, J.M. and C.W.J Granger (1969), "The Combination of Forecasts," *Operations Research Quarterly*, 20, 451–468.

Chan, F. and L. L. Pauwels (2018), "Some Theoretical Results on Forecast Combinations," *International Journal of Forecasting*, 34, 64–74.

Chauvet, M. and S. Potter (2013), "Forecasting Output," *Handbook of Economic Forecasting*, 2, 141–194.

Claeskens, G., J. R. Magnus, A. L. Vasnev, and W Wang (2016), "A Simple Theoretical Explanation of the Forecast Combination Puzzle," *International Journal of Forecasting*, 32, 754–62.

Clemen, R. T. and R. L. Winkler (1985), "Limits for the Precision and Value of Information from Dependent Sources," *Operations Research*, 33, 427–442.

Clemen, R.T. (1989), "Combining Forecasts: A Review and Annotated Bibliography (With Discussion)," *International Journal of Forecasting*, 5, 559–583.

Clements, M.P., R.W. Rich, and J.S. Tracy (2022), "Surveys of Professionals," Federal Reserve Bank of Cleveland Working Paper.

Croushore, D. and T. Stark (2019), "Fifty Years of the Survey of Professional Forecasters," *Economic Insights*, 4, 1–11.

Diebold, F.X. and M. Shin (2019), "Machine Learning for Regularized Survey Forecast Combination: Partially-Egalitarian Lasso and its Derivatives," *International Journal of Forecasting*, 35, 1679–1691.

Diebold, F.X., M. Shin, and B. Zhang (2023), "On the Aggregation of Probability Assessments: Regularized Mixtures of Predictive Densities for Eurozone Inflation and Real Interest Rates," *Journal of Econometrics*, 237, 105321.

Elliott, G. and A. Timmermann (2016), *Economic Forecasting*, Princeton University Press.

Elliott, Graham and Jie Liao (2025), "Combining Forecasts – On Why Averaging Beats Optimal Linear Weights," Working paper, University of California, San Diego. Version: May 5, 2025.

Engle, R.F. and B.T. Kelly (2012), "Dynamic Equicorrelation," *Journal of Business and Economic Statistics*, 30, 212–228.

Genre, V., G. Kenny, A. Meyler, and A. Timmermann (2013), "Combining Expert Forecasts: Can Anything Beat the Simple Average?" *International Journal of Forecasting*, 29, 108–121.

Gourieroux, C., A. Monfort, and E. Renault (1993), "Indirect Inference," *Journal of Applied Econometrics*, 8, S85–S118.

Hsiao, C. and S. K. Wan (2014), "Is there an Optimal Forecast Combination?" *Journal of Econometrics*, 178, 294–309.

LeBaron, B. and K. Smith (2025), "Learning Integrated Inflation Forecasts in a Simple Multi-Agent Macroeconomic Model," *Journal of Economic Dynamics and Control*, 172, 104979.

Makridakis, S. and R.L. Winkler (1983), "Averages of Forecasts: Some Empirical Results," *Management Science*, 29, 987–996.

Pivetta, F. and R. Reis (2007), "The Persistence of Inflation in the United States," *Journal of Economic Dynamics and Control*, 31, 1326–1358.

Smith, J. and K. F. Wallis (2009), "A Simple Explanation of the Forecast Combination Puzzle," *Oxford Bulletin of Economics and Statistics*, 71, 331–355.

Smith Jr, A. A. (1993), "Estimating Nonlinear Time-Series Models using Simulated Vector Autoregressions," *Journal of Applied Econometrics*, 8, S63–S84.

Stock, J. H. and M. W. Watson (2002), "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business & Economic Statistics*, 20, 147–162.

Stock, J. H. and M. W. Watson (2004), "Combination Forecasts of Output Growth in a Seven-Country Data Set," *Journal of Forecasting*, 23, 405–430.

Stock, J. H. and M. W. Watson (2007), "Why Has US Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking*, 39, 3–33.

Stock, J.H. and M.W. Watson (2016), "Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics," In J.B. Taylor and H. Uhlig (eds.), *Handbook of Macroeconomics, vol. 2A*, Elsevier, 415-526.

Surowiecki, J. (2005), *The Wisdom of Crowds*, Vintage Books.

Timmermann, A. (2006), "Forecast Combinations," *Handbook of Economic Forecasting*, 1, 135–196.

Wang, X., R.J. Hyndman, F. Li, and Y. Kang (2022), "Forecast Combinations: An Over 50-year Review," ArXiv:2205.04216 (https://arxiv.org/abs/2205.04216).