

On the Difficulty of Measuring Forecasting Skill in Financial Markets

S.E. Satchell and O.J. Williams

August 2010

CWPE 1039

On the Difficulty of Measuring Forecasting Skill in Financial Markets

S.E. Satchell
Trinity College
University of Cambridge

O.J. Williams*
Scalpel Research Ltd

August 2010

Abstract

The use of correlation between forecasts and actual returns is commonplace in the literature, often used as a measurement of investors' skill. A prominent application of this is the concept of the Information Coefficient (IC). Not only can IC be used as a tool to rate analysts and fund managers but it also represents an important parameter in the asset allocation and portfolio construction process. Nevertheless, theoretical understanding of it has typically been limited to the partial equilibrium context where the investing activities of each agent have no effect on other market participants. In this paper we show that this can be an undesirable oversimplification and we demonstrate plausible circumstances in which conventional empirical measurements of IC can be highly misleading. We suggest that improved understanding of IC in a general equilibrium setting can lead to refined portfolio decision making *ex ante* and more informative analysis of performance *ex post*.

JEL Classification: D53, D82, D84, G11, G17

Keywords: performance measurement, skill, financial forecasting, active management, Information Coefficient, Information Ratio

*The authors thank Robert Kosowski, Hamish Low and Susan Thorp for many useful comments. O.J. Williams acknowledges support from the Economic and Social Research Council (award number PTA-031-2004-00054).

1 Introduction

“...the man who is right always has two forces working in his favor - basic conditions and the men who are wrong” *Reminiscences of a Stock Operator, Edwin Lefèvre (1923, p.124)*

“Some investors are born with high *IC*, some achieve high *IC*, and some have high *IC* thrust upon them” *Anonymous market practitioner during year-end bonus discussions*

The correlation between investors’ *forecasts* of returns and *actual* returns themselves has featured in many contexts, theoretical and empirical, ever since the earliest days of portfolio theory. Among the first appearances of this important quantity was its prominent rôle in Treynor and Black’s (1973) model of active management, one of the earliest theoretical treatments of that topic. Although not referred to as such in that paper, this correlation became known as the Information Coefficient (*IC*) and, under this sobriquet, has featured prominently in the literature ever since.¹

In a related paper, Satchell and Williams (2010) (henceforth, S&W), we applied skill measurements to policy questions concerning relationships between welfare and the profile of skill of agents in the market. We demonstrated the effects of changing the *distribution* of skill in the market, assuming that available skill *levels* remain fixed. We also showed that adjusting skill parameters had generally ambiguous effects on equilibrium price volatility, agents’ returns and wealth volatility.

To explore welfare aspects in that analysis we used the *hit-rate* which is the proportion of future outcomes correctly forecast by an investor, focusing only on market *direction* and ignoring magnitudes. That approach was intended to keep the mathematics simple in order to focus on the economic intuition. In this paper however we take the analysis a stage further by considering the more general case where agents forecast future asset prices over a continuous range. In this context the the hit-rate measure is naturally superseded by the Information Coefficient.

In S&W we assumed that agents had control over the level of their own skill level; we computed marginal effects of skill improvements on the assumption that an agent could adjust their own hit-rate independently of other agents. On the face of it this is an innocuous assumption which is ubiquitous elsewhere in the literature. A typical model is that agents know their own skill levels based on econometric analysis of their past forecasts and actual returns. A rational utility-maximizing investor determines their asset demands given knowledge of their own skill level and the implicit assumption is that this skill level will remain constant over repeated forecasting instances. In other words an agent’s *ex ante* hit-rate (or *IC*) (used in their decision process) is equal, in expectation at least, to their *ex post* hit-rate (or *IC*) (measured from actual performance). In a stylised theory of active management Grinold and Kahn (1999) claim

¹Coggin and Hunter (1983) discuss empirical *IC* measurement issues, Dimson and Marsh (1984) refer to the concept in their empirical study of analysts’ forecasting ability, Khan *et al* (1996) apply *IC* in global asset allocation, Khan (1998) discusses its applicability in bond investing, Grinold and Kahn (1999) cover a range of active management analytics, Lee (2000) uses it as a key parameter in a range of analytics pertaining to tactical asset allocation, Herold (2003) discusses its relevance in qualitative forecasting, Satchell and Williams (2007) consider the question of optimal forecasting horizon, Fishwick (2007) analyses the case of multiple models.

that skill (in their case measured by IC) is what differentiates active managers from passive managers.

This paper questions that crucial assumption. We will show that when an investor has a non-trivial impact on market price then the equality between *ex ante* and *ex post* IC will not hold in general. The basic problem is that as an investor's forecast is priced into equilibrium it inevitably tends to push equilibrium price closer towards the investor's own private forecast. This erodes the amount of alpha available for the investor. In the extreme case of perfect forecasting ($ex\ ante\ IC \rightarrow 1$) the equilibrium price will represent exactly the future price, alpha will be constantly equal to zero and hence measured IC will degenerate to zero.

Some institutional investors may consider their trading sizes to be negligible relative to overall market turnover and therefore turn a blind eye to this issue. However the relevant size measure here is not simply the volume of each institution's own trading in isolation, but the overall proportion of market activity accounted for by that institution aggregated together with others which share substantially the same market views. The extensive literature on style analysis indicates that there is a high degree of clustering between investors who can be sorted into categories within which performance is relatively highly-correlated.² In recent years this phenomenon has become particularly acute in such areas as quantitative strategies (analysed in some detail by Lo (2008) and alluded to by Litterman (2009)). Therefore the issues which we consider in this paper have extremely broad relevance.

Our paper is a distant relative of parts of the market microstructure literature which deal with the process by which traders reveal information about their private forecasts via the equilibrium price, such as the seminal work by Kyle (1985). Similar topics are addressed in the literature on dynamic models of interacting agents, exemplified by Brock and Hommes (1997) and Branch (2002), and by theoretical and empirical studies of order flow and *tâtonnement* (reviewed by Bouchaud *et al* (2009)). However the specific problem which we approach here deliberately abstracts away from the precise mechanics of trade execution; we focus on equilibrium price outcomes irrespective of how exactly these are brought about. This keeps our results at a general level but nevertheless the methodology which we demonstrate is equally applicable in microstructural settings where even more detailed subtleties of investor interaction can be incorporated.

We relate more closely to the empirical performance measurement literature which debates the perennial question of whether results are due to luck or skill. In this research it is common to use estimated alphas (or Information Ratios) as an indicator of skill and compute such metrics as significance levels and test for evidence of persistence. This line of inquiry dates back at least as far as Treynor and Mazuy (1966) and a representative sample of papers includes Dimson and Marsh (1984), Lee and Rahman (1990) and Malkiel (1995). On the whole the evidence in favour of skill dominating luck is inconclusive. However there are many econometric challenges involved in the estimation process, such as highlighted by Ferson and Schadt (1996) and in prominent recent work, Kosowski *et al* (2006) and Kosowski *et al* (2007) deploy robust bootstrap methods which address many of these issues and find that performance of the best mutual funds and

²As a sampling of the relevant literature: Sharpe (1992) and Brown and Goetzmann (1997) approach this for the case of mutual funds while Agarwal and Naik (2000), Fung and Hsieh (2002) and Brown and Goetzmann (2003) address the case of hedge funds.

hedge funds cannot be explained by luck.

In much of the performance measurement literature the maintained hypothesis is that empirical alphas (or their *IC* or Information Ratios counterparts) are a valid representation of skill; something which we will show should not necessarily be taken at face value without further consideration of the market structural context.

A candid understanding of forecasting ability (and its limits) is essential throughout the fund management process. Considerable resources and attention are routinely applied to the business of forecasting returns and their dependency structures as well as optimisation and portfolio construction, however the effectiveness of all this activity is seriously impaired if forecasts bear little resemblance to reality.³ In practice, the lack of certainty in forecasts can be dealt with in numerous alternative ways, e.g. robust optimisation (see Koutsoyannis and Satchell (2007)), Bayesian methods (e.g. Black and Litterman (1992)) or resampling techniques (see Scherer (2007)), however we feel it would be desirable to have a more solid theoretical understanding of structural factors which influence forecast accuracy in the first place.

It is well known among practitioners that long time series are required to obtain tolerable estimates of *IC*; while this may simply be due to erratic forecasting performance by analysts, our results highlight other potential explanatory factors. Without resorting to simulation we also provide a number of analytic results and approximations which may prove useful to practitioners as basic tools to assess the order-of-magnitude of these factors. Furthermore in certain circumstances it may be appropriate to treat *IC* as a stochastic variable in its own right; again our results can be insightful in this regard.

For the sake of clarity throughout this paper we use the term *Realised Information Coefficient* (*RIC*) to refer to *measured ex post IC* in order to distinguish this from the *unobserved IC* which we take to mean the *IC* used by an investor *ex ante* to make their portfolio allocation decision, i.e. the *IC* which they believe represents their personal skill level. At the risk of repetition: the quantities normally measured and discussed elsewhere in the literature are what we term *RIC*.

We will show that the relationship between these two values is a complex function which depends on the profile of skill of all agents in the market and the correlation structure of their forecasts, and indeed the relationship can be non-monotonic. The key message is that while *ex ante IC* may indeed be a valid measure of an investor's innate forecasting *skill*, the *ex post RIC*, necessarily extracted by empirical analysis, depends on a whole slew of factors entirely outside the investor's control (and knowledge). We will furnish various examples which demonstrate that the impact of these factors can vary from entirely negligible (which makes *RIC* a perfectly valid metric) to dramatic (changing the sign and magnitude of *IC* between *ex post* and *ex ante* and rendering *RIC* virtually useless as an indication of skill).

In fact depending on the degree of complexity of the market's skill structure we find that *RIC* can often assume counter-intuitive values in relation to *IC*; for example we will demonstrate circumstances where a neutral investor who never uses forecasts ($IC = 0$) can appear to be a valuable contrarian indicator ($RIC < 0$).⁴

³In particular the hedge fund industry has recently seen a proliferation of so-called 'alpha capture' systems in which systematic quantitative methods are used by funds to rank and filter brokers' recommendations according to their forecast level of accuracy (see Burgess (2006), Mackintosh (2006) and Grene (2008)).

⁴This is somewhat reminiscent of Lewis Carroll's observation that 'even a stopped clock is right twice a day'.

Although a more detailed discussion is beyond the scope of this paper we do believe that an investor has at least partial control over their level of forecasting skill. Plausibly it could be *increased* by more thorough analysis or training, and methods by which it can be *decreased* are trivially obvious. The target of these efforts is clearly *ex ante IC*, untainted as it is by external market effects. However our non-monotonicity result means that (potentially very costly) improvements in personal *IC* may not translate into any measured skill improvements at all, whether in terms of *RIC*, Information Ratios or utility. It also suggests that an optimal level of skill exists for a given type of investor, albeit dependent on the skill of others.

It is *ex ante IC* which should be the exogenous primitive of analysis when we consider comparative statics or hypothetical scenarios. Say, for instance, a star stock-picker in US equities (measured in terms of *ex post RIC*) is contemplating investing in a small emerging market where he has no prior experience. What expectations should he (and his investors) have as regards likely *RIC* and Information Ratio for this endeavour? Even if he has a consistent ability to forecast emerging prices just as accurately as US stocks, our results show that naively assuming the portability of his usual *RIC* into the new market is a misleading oversimplification. We suggest that a preferable approach would be to combine the investor's *IC* with knowledge of the skill structure in the new market in order to obtain a new *RIC* value.

A wide array of questions like this arise when we consider the intertemporal stability of *RIC*, effect of herding on *RIC*, policy issues surrounding investor education, and so on. In all these matters we believe knowledge of investors' *ex ante IC* levels is vital to coherent analysis and the temptation to proxy these using measured *RIC* levels should be avoided if at all possible.

Having accurately quantified *RIC* we are able to incorporate this value in an analytic expression for an investor's *ex ante* expected utility and we compare this briefly with the corresponding Information Ratio given by applying Grinold's (1989) Fundamental Law of Active Management.

The paper is organised as follows: Section 2 discusses the sequence of events involved in the forecasting and investing process and explains how various alternative methods of calculating *IC* values fit into this schedule; Section 3 sets up our equilibrium model and derives our central result: the Realised Information Coefficient; Section 4 provides illustrative examples and Section 5 discusses utility and welfare issues. Section 6 concludes.

2 Skill and Price Targets

When publishing buy/sell recommendations it is usual for analysts in the equity market to provide forecasts of future *prices* (known as *price targets*) rather than forecasts of *returns*. To some extent this is for convenience due to delays between analysis and publication during which time the current market price can change, however it is also very consistent with the large body of corporate finance literature which focuses on valuation of companies by methods such as discounted cashflow or earnings multiples (see, for instance, Brealey and Myers (2002)). It is most likely therefore that the elemental forecast is of price level rather than percentage return, with the percentage returns themselves computed as a secondary step given current market prices.

An extensive literature deals with the part which analysts and their forecasts play in financial

markets, encompassing the decision process by which forecasts are devised and the statistical relationship between forecasts and actual prices, as well as discussion of issues concerning incentives and the regulatory environment. A thorough review of recent work in this area is provided by Ramnath *et al* (2008).

A noteworthy paper with relevance to our analysis here is Asquith *et al* (2005) who perform a thorough empirical analysis of the contents of analysts' research reports and their associated market impact. They find that price targets have significant explanatory power over actual price outcomes. Indeed Treynor and Black (1973) also recognise that one of the earliest stages in the analysis process is when 'the analyst begins by appraising the security in question' and subsequently 'compares his appraisal with the current market price of the security'. Nevertheless the literature of financial economics and mathematical finance has been dominated for decades by models in which the evolution of prices is specified exogenously or derived from exogenous primitives such as expected returns, covariances or endowment processes.

From a skill measurement point of view we believe this highlights the need for precise treatment of the sequence and timing of events involved in forecast formulation and subsequent trading and we refer to the specific timeline of events as shown in Table 1. If we measure *IC* over a series of returns computed between I and IV but the transaction in III has market impact then the actual *realised IC* is likely to differ from this since it should be computed based on returns between III and IV. We might think of the *IC* between I and IV as the skill level which an analyst is 'born with', while that between III and IV is endogenous to the general equilibrium which is ground out once all participating investors have interacted based on their own (possibly correlated) research efforts.

We assume that there are J agent types identified by superscripts $i \in [0, J - 1]$, we denote the future price (stage IV) by p_{t+1} , investor type i 's standardised score (stage II) by $S^{(i)}$ and the *known* pre-forecast market price (stage I) by p_0 . We impose joint-normality between p_{t+1} and all scores $S^{(i)}$. Therefore by the standard formula for the conditional expectation of a multivariate normal we have:

$$\mathbb{E}_{t=0} \left[p_{t+1} \mid S^{(i)} \right] - p_0 = \mu + IC^{(i)} \sigma S^{(i)} \quad (1)$$

where μ and σ^2 are, respectively, unconditional mean and variance of future money gains per unit asset. As (1) indicates, $S^{(i)}$ is standardised with $\mathbb{E} [S^{(i)}] = 0$ and $\text{VAR} [S^{(i)}] = 1$. Expectation and variance operators describe temporal distributions and we can think of prices (p_t) and forecasts $(S_t^{(i)})_{i \in [0, J-1]}$ as being a $(J+1)$ -vector of random variables indexed by t . We emphasise, however, that throughout this paper we make no assumptions about the actual dynamics of this process over time. Our analysis relies only on the requirement that scores $S^{(i)}$ (used to determine equilibrium price p_t) be jointly-normal with the actual future price p_{t+1} and we are agnostic about the source of randomness which drives p_{t+1} itself.

We will shortly distinguish between the two alternative formulations of Information Coefficient which we have described and show their relationship. In order to highlight the essence of this without undue mathematical complexity we define *IC*'s here in terms of *per-asset money gains* rather than percentage returns. In other words we assume that investors' money gain

Table 1: Stages in the research and trading process

Stage	Activity		Time
I	Analyst observes current market price		0
II	Analyst performs research	Forecast Price Determined	$[0, t)$
III	Investor executes transaction	Equilibrium Price Determined	t
IV	Future price materialises	Gain or Loss Realised	$t + 1$

forecasts (or scores) and actual money gains follow a bivariate normal distribution.⁵

The normal distribution assumption for future values (as opposed to percentage returns) does appear elsewhere in the finance literature. We list, among many others, the following: Leland and Pyle's (1977) analysis of informational asymmetries between borrowers and lenders in financial markets, the literature on optimal security design (e.g. Allen and Gale (1988), Demange and Laroque (1995)), and Lo and Wang's (2001) intertemporal model of asset-pricing and trading volume where the effects of agents' competition for returns are a key aspect.

3 Model and Distributions

3.1 Equilibrium Price Solution

We assume J investor *types* identified by $i \in (0, J - 1)$. There are a total of K investors in the market and there are $K f^{(i)}$ investors of type i . We assume that each investor has an exponential utility function with differing type-specific coefficients of absolute risk aversion, access to unlimited risk-free borrowing/lending and that all settlements and payoffs occur simultaneously in the same period. Investors of type i all believe the future asset price is distributed $N(p_0 + \mu^{(i)}, \sigma^{(i)2})$ where we emphasise that their subjective mean return per asset (in terms of money gains) $\mu^{(i)}$ is defined as:

$$\mu^{(i)} = \mu + IC^{(i)} \sigma S^{(i)}$$

consistent with the notion of a *price target*. Their subjective variance $\sigma^{(i)2}$ is given by

$$\sigma^{(i)2} = \sigma^2(1 - IC^{(i)2})$$

While the classification of investors into types may seem like a blunt simplification, it very much agrees with the extensive body of literature documenting high degrees of correlation between active managers who share the same style, as well as stylized facts such as herding by momentum investors.

Investor type i has coefficient of absolute risk aversion denoted by $\lambda^{(i)}$ and their asset allocation problem can be written in terms of p as follows:

$$\begin{aligned} \mathbb{E} [u^{(i)}(x^{(i)})] &= -\frac{1}{\lambda^{(i)}} \mathbb{E} \left[\exp(-\lambda^{(i)} x^{(i)} (p_{t+1} - p_t)) \right] \\ &= -\frac{1}{\lambda^{(i)}} \left[\exp \left(-\lambda^{(i)} x^{(i)} \mu^{(i)} + \frac{1}{2} \lambda^{(i)2} x^{(i)2} \sigma^{(i)2} + \lambda^{(i)} x^{(i)} p_t - \lambda^{(i)} x^{(i)} p_0 \right) \right] \end{aligned}$$

⁵In a continuous time dynamic model this simplification could potentially be avoided by modeling asset prices as a lognormal diffusion.

where $x^{(i)}$ indicates *units* of risky asset held by each agent of type i . First order conditions for expected utility maximisation with respect to $x^{(i)}$ therefore provide:

$$x^{(i)} = \frac{p_0 + \mu + IC^{(i)}\sigma S^{(i)} - p_t}{\lambda^{(i)}\sigma^{(i)2}} \quad (2)$$

We now determine the equilibrium price by imposing market clearing, where we denote the aggregate supply of securities by Z (assumed fixed):

$$\begin{aligned} K \sum_{i=0}^{J-1} f^{(i)} \frac{p_0 + \mu^{(i)} - p_t}{\lambda^{(i)}\sigma^{(i)2}} &= Z \\ p_t &= p_0 + \sum_{i=0}^{J-1} \delta^{(i)} \mu^{(i)} - \bar{\sigma}^2 \frac{Z}{K} \end{aligned} \quad (3)$$

We define the weighted harmonic mean of the risk-aversion-variance product as:

$$\bar{\sigma}^2 \equiv \left(\sum_{i=0}^{J-1} f^{(i)} \frac{1}{\lambda^{(i)}\sigma^{(i)2}} \right)^{-1} \quad (4)$$

and a quantity we call *market presence* as:

$$\delta^{(i)} = f^{(i)} \frac{\bar{\sigma}^2}{\lambda^{(i)}\sigma^{(i)2}} \quad (5)$$

Hence we find that the equilibrium price is directly driven by agents' forecasts adjusted by a risk-aversion discount. We have in fact specialised a general result in Lintner (1969) who considers equilibrium with heterogeneous agents; here we parameterise the heterogeneity in a particular fashion but the weighted average form of the expression still holds for subjective means $\mu^{(i)}$ and variances $\sigma^{(i)2}$ derived in other ways.

We have already alluded to the temporal $(J+1)$ -dimensional stochastic process $(p_t, S_t^{(i)})_{i \in [0, J-1]}$ indexed by t . We now introduce a new family of cross-sectional processes indexed by $i \in [0, J-1]$. In this framework, realisations (for fixed t) are $\mu^{(i)}$ and $S^{(i)}$. In fact we could write everything with double indices $(\mu_t^{(i)}, S_t^{(i)})$, etc.) but we avoid this. Possible cross-sectional measures could be $(f^{(0)}, \dots, f^{(J-1)})$ or $(\delta^{(0)}, \dots, \delta^{(J-1)})$. Indeed it is often helpful to think of the vector of market presences $\delta^{(i)}$ as if it were a cross-sectional probability measure which depends on the frequency of agents' types in the market, their skill levels and risk aversion. Under this interpretation we can think of the equilibrium price (3) as being the expected value of the forecast of an agent selected at random in the market, denoted by

$$p_t \mid \mathbf{S} = \mathbb{E}_\delta \left[p_0 + \mu^{(i)} - \bar{\sigma}^2 \frac{Z}{K} \right] \quad (6)$$

where \mathbb{E}_δ denotes expectation with respect to the δ -measure and $\mathbf{S} \equiv [S^{(0)}, \dots, S^{(J-1)}]$.

We will return to this analogy shortly.

Proposition 3.1. *If $0 \leq IC^{(i)} < 1 \forall i$ and $\sum_{i=0}^{J-1} \delta^{(i)} = \sum_{i=0}^{J-1} f^{(i)} = 1$ then $\delta^{(i)} \geq 0 \forall i$ iff $f^{(i)} \geq 0 \forall i$.*

Proof. See A.3.1. □

3.2 Distribution of Price over Repeated Forecasting Instances

We now consider the *ex ante* distribution of equilibrium prices p_t over repeated realisations of $S^{(i)}$ given perfect knowledge of $\delta^{(i)}$ and $IC^{(i)}$. Note that in the course of setting up our investor's optimal portfolio problem we implicitly assumed that she has no knowledge of the structure of skill in the market (i.e. the levels of other types' IC and the inter-forecaster correlations, none of which featured in her optimal portfolio decision). However here we put ourselves in the position of an omniscient social planner. Alternatively one can think of this distribution as that which any investor would gradually learn over time by empirical observation of the behaviour of prices.

We have already established that the equilibrium price p_t will be a δ -weighted linear combination of individual agents' forecasts minus a risk premium. Conditional on knowledge of all agents' forecasts (represented by $S^{(i)}$) we know that p_t is deterministic therefore. From an unconditional point of view, however, we know that this linear combination will itself be normal since we have assumed joint normality of forecasts and actual prices. Therefore we focus our attention here on establishing the unconditional (over i and t) expectation and variance of p_t as follows:

$$\mathbb{E}[p_t | \mathbf{S}] = \mathbb{E} \left[\mathbb{E}_\delta \left[p_0 + \mu^{(i)} - \bar{\sigma}^2 \frac{Z}{K} \right] \right] = p_0 + \mu - \frac{\bar{\sigma}^2 Z}{K} \quad (7)$$

and

$$\mathbb{E}(p_t) = \mathbb{E}[\mathbb{E}[p_t | \mathbf{S}]] = p_0 + \mu - \frac{\bar{\sigma}^2 Z}{K} \quad (8)$$

$$\begin{aligned} \text{VAR}(p_t) &= \mathbb{E}[\text{VAR}(p_t | \mathbf{S})] + \text{VAR}[\mathbb{E}(p_t | \mathbf{S})] \\ &= \mathbb{E} \left(\mathbb{E}[p_t | \mathbf{S} - \mathbb{E}[p_t | \mathbf{S}]]^2 \right) + 0 \\ &= \mathbb{E} \left\{ \left[\sum_{i=0}^{J-1} \delta^{(i)} IC^{(i)} \sigma S^{(i)} \right] \left[\sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)} \sigma S^{(j)} \right] \right\} \\ &= \sum_{i=0}^{J-1} \sum_{j=0}^{J-1} \delta^{(i)} \delta^{(j)} IC^{(i)} IC^{(j)} \phi_{ij} \sigma^2 \\ &\equiv V \sigma^2 \end{aligned} \quad (9)$$

where $0 \leq V < 1$ is a scalar multiplier of the unconditional future price variance which represents the degree to which uncertainty over future prices is 'transmitted' into current equilibrium price p_t . We refer to V as the Variance Factor and with some abuse of terminology we interpret this as a measure of volatility. ϕ_{ij} represents correlation between the forecasts of type i and type j , (i.e. $S^{(i)}$ and $S^{(j)}$).

Although it is perhaps most immediate to think of the repeated forecasting instances as taking place sequentially over successive time periods (thereby making V loosely analogous to time-series volatility) we might alternatively consider forecasts as relating to multiple different assets but which take place at the same instant in time. The latter interpretation is analogous

to cross-sectional volatility (CSV) as discussed by S  n  chal (2004).⁶

If investors all have zero skill, i.e. $IC = 0$, then an uninformed investor in fact has no uncertainty over p_t : it will be that which clears the market on the basis of unconditional expectations and p_t is determined unambiguously. However if all investors in the market are known to have high skill levels with $IC \rightarrow 1$ and correlation $\phi_{jk} = 1$, then uncertainty over p_t is at its highest and tends towards the unconditional variance of future prices.

If we apply the probability interpretation of $\delta^{(i)}$ introduced earlier then (10) can be understood in the following fashion. Suppose that we select two investors independently at random (with replacement) according to the $\delta^{(i)}$ measure; the probability of choosing a given pair of agent types (i, j) is clearly the product $\delta^{(i)}\delta^{(j)}$ (we represent this measure by $\delta \times \delta$). We denote this pair of types by $\mathbf{B} = (B_1, B_2)$ and their respective forecasts by $\mu^{(B_1)}$ and $\mu^{(B_2)}$. Now the covariance between the agents' forecasts is a random variable in its own right, denoted by $\text{COV}[\mu^{(B_1)}, \mu^{(B_2)} | \mathbf{B}]$. We then pose the question: what is the unconditional covariance between two agents' forecasts? We can show that this is also the expression in (10) since we can re-write it as

$$\begin{aligned} \text{COV}(\mu^{(B_1)}, \mu^{(B_2)}) &= \mathbb{E}_{\delta \times \delta} \left(\text{COV}[\mu^{(B_1)}, \mu^{(B_2)} | \mathbf{B}] \right) + \text{COV}_{\delta \times \delta} \left(\mathbb{E}[\mu^{(B_1)} | \mathbf{B}], \mathbb{E}[\mu^{(B_2)} | \mathbf{B}] \right) \\ &= \mathbb{E}_{\delta \times \delta} \left(\sum_{i=0}^{J-1} \sum_{j=0}^{J-1} I_{\{B_1\}}(i) I_{\{B_2\}}(j) IC^{(B_1)} IC^{(B_2)} \phi_{B_1, B_2} \sigma^2 \right) + \text{COV}_{\delta \times \delta}(\mu, \mu) \\ &= \sum_{i=0}^{J-1} \sum_{j=0}^{J-1} \delta^{(i)} \delta^{(j)} IC^{(i)} IC^{(j)} \phi_{ij} \sigma^2 + 0 \\ &= V \sigma^2 \end{aligned}$$

where $I_{\{B_k\}}(x)$ represents the indicator function which takes the value 1 iff $x = B_k$ and 0 otherwise.

We therefore see a close relationship between the degree of *agreement* between agents and the level of *volatility* in the market.

On an informal level we can extend this analogy further: in the event where the two agents are of the same type there will be perfect agreement on forecasts and therefore zero trade; however whenever the agent types differ there will almost surely be trade since one forecast will be strictly higher than the other. When skill is low then forecasts will tend to cluster relatively close to the unconditional mean future price (generating relatively low V) and when skill is higher the forecasts will be more dispersed (higher V).

When correlation is high then the equilibrium price at which trades take place will more closely track individual forecasts (rather than pairwise averages) and therefore mirror the degree of dispersion. When correlation is low, however, an extreme forecast on the part of type i is more likely to be matched with a more modest forecast by type j , thus dampening V . In the extreme case where correlation is significantly negative, for instance, then a particularly high forecast from i is more likely to be matched by a correspondingly low forecast from j which will

⁶Whether we consider the time-series or cross-sectional interpretation, the obvious expeditious assumption here is that the skill structure of the market and unconditional moments of p_{t+1} remain unchanged over repeated instances.

tend to neutralize it completely.

3.3 Return and Wealth Distributions

We now turn our attention to the distribution of investors' future wealth. We are interested in the unconditional distributions of two particular random variables: the deviation of the future (time $t + 1$) price from the (time t) equilibrium price (we denote this by Y) and the type i agent's *anticipated* money gain on one unit of the primary securities (denoted $A^{(i)}$). Note that *actual* money gain Y is common to all types whereas *anticipated* gain is type-specific. From $A^{(i)}$ it is straightforward to obtain the agent's holding of the asset (which we denote $H^{(i)}$).

First we have (with derivations in A.1.1):⁷

$$Y \equiv (p_{t+1} - p_t) \sim N \left(\bar{\sigma}^2 \frac{Z}{K}, \sigma^2 \left[1 + V - 2 \sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)2} \right] \right) \quad (11)$$

and

$$\begin{aligned} A^{(i)} &\equiv (p_0 + \mu^{(i)} - p_t) \\ &\sim N \left(\bar{\sigma}^2 \frac{Z}{K}, \sigma^2 \left[IC^{(i)2} + V - 2 \sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)} IC^{(i)} \phi_{ij} \right] \right) \end{aligned} \quad (12)$$

hence type i agent's *holding* of the asset is given by:

$$\begin{aligned} H^{(i)} &\equiv \frac{A^{(i)}}{\lambda \sigma^{(i)2}} \\ &= \frac{p_0 + \mu^{(i)} - p_t}{\lambda \sigma^{(i)2}} \\ &\sim N \left(\frac{\bar{\sigma}^2 Z}{K \sigma^{(i)2}}, \frac{\sigma^2}{(\lambda \sigma^{(i)2})^2} \left[IC^{(i)2} + V - 2 \sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)} IC^{(i)} \phi_{ij} \right] \right) \end{aligned} \quad (13)$$

3.4 Realised Information Coefficient

An investor's gains over time depend on the relationship between their forecast and the forecasts of others and also on the general structure of skill in the market. *RIC* is the correlation between an agent's *forecast* money gain (per unit asset) $A^{(i)}$ and *actual* money gain Y . For a given type of agent i , the difference between the two measurements depends on both how faithfully the equilibrium price represents future prices (i.e. the level of skill of others) as well as the degree of correlation between forecast of agent type i and others.

Proposition 3.2. *The Realised Information Coefficient achieved by type i is defined as the correlation between **actual** gains (11) $Y \equiv p_{t+1} - p_t$ and type i 's **anticipated** gains (12)*

⁷The mean is a non-stochastic risk premium which follows from (8) above.

$A^{(i)} \equiv p_0 + \mu^{(i)} - p_t$. It is given by the formula:

$$RIC^{(i)} = \frac{IC^{(i)2} - \sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)2} - \sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)} IC^{(i)} \phi_{ij} + V}{\sqrt{\left[1 + V - 2 \sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)2}\right] \left[IC^{(i)2} + V - 2 \sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)} IC^{(i)} \phi_{ij}\right]}} \quad (14)$$

Proof. See A.3.2. □

Proposition 3.3. *For all equi-frequent types with equal constant absolute risk-aversion $RIC = 0$ whenever all types share the same IC irrespective of the degree of correlation or shared level of skill.*

Proof. See A.3.3. □

This proposition provides us with the important result that if all agent types have equal constant absolute risk aversion and equal headcount in the market and increase their skill in tandem, then their ability to capture alpha remains at zero. This situation is somewhat redolent of an ‘arms race’ and is analogous to the case of the two-agent discrete state model outlined in S&W where skill was measured using *hit-rate*. We note also that if Information Ratio is naively computed using IC instead of RIC then it is liable to be extremely misleading in this situation. Indeed if we were to derive a social welfare function based on IC (which, we emphasise once again, is our *ex ante* measure) then it would show the benefit of increasing skill when the reality would be that skill improvement is futile.

4 Example Skill Profiles

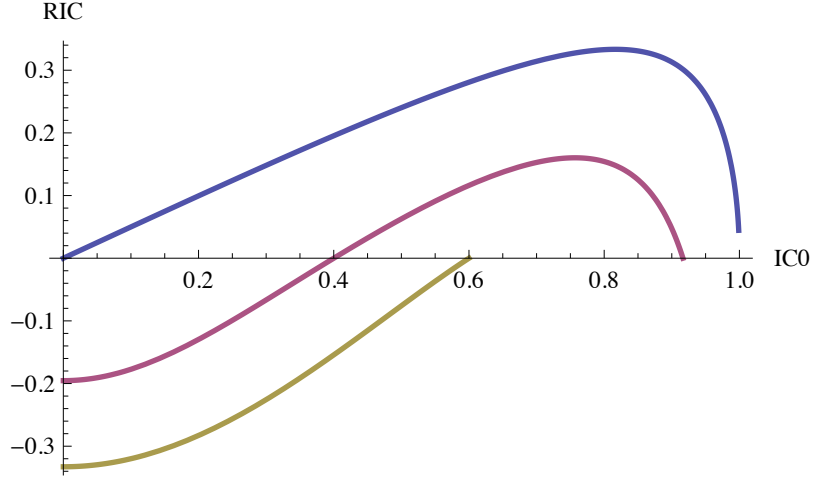
We now illustrate the RIC concept and the calculation of variances for two example structures: (1) two equi-frequent types with differential skill and uncorrelated forecasts, and (2) a continuum of an infinite number of correlated types where skill is distributed unevenly across the investor population. In both cases we assume equal constant absolute risk aversion across all agent types. These represent two extreme cases and we would suggest that many real-world markets may lie somewhere in between these two. For each case we are also able to present various relevant propositions.

4.1 Two equi-frequent types with equal constant absolute risk-aversion, differential skill and uncorrelated forecasts

Suppose that $IC^{(0)} > IC^{(1)}$ with $f^{(0)} = f^{(1)} = 0.5$ and $\phi_{01} = 0$, i.e. one investor type has a clear forecasting advantage over the other. We now have $\delta^{(0)} > 0.5 > \delta^{(1)}$ and therefore σ_A^2 will differ between the two types.

$$\begin{aligned} \sigma_{A^{(i)}}^2 &= \sigma^2 \left[IC^{(i)2} + V - 2\delta^{(i)} IC^{(i)2} \right] \\ &= \sigma^2 \left[V + IC^2(1 - 2\delta^{(i)}) \right] \end{aligned} \quad (15)$$

Figure 1: Realised Information Coefficient (RIC) for agent type 0 in model of two equi-frequent uncorrelated types; $IC^{(1)}$ fixed at 0, 0.4 and 0.8 (from highest to lowest line)



Specifically, by virtue of different signs in the $(1 - 2\delta^{(i)})$ component in (15), $\sigma_{A(0)}^2 < V < \sigma_{A(1)}^2$. This means that the superior forecaster sees money returns which are less variable than spot prices while the opposite will be true for the inferior forecaster. Note that this will be the case irrespective of the exact magnitudes of $IC^{(0)}$ and $IC^{(1)}$. This effect arises since the superior forecaster's larger position-taking (reflected in $\delta^{(0)} > 0.5$) tends to work against her when making investments: she will always tend to push the market towards her private valuation thereby reducing the range of returns accessible per unit asset. However it is important to realise that this only describes price on a *per asset* basis: her superior forecasting ability and larger position-taking somewhat compensate for these effects when considering overall portfolio returns.

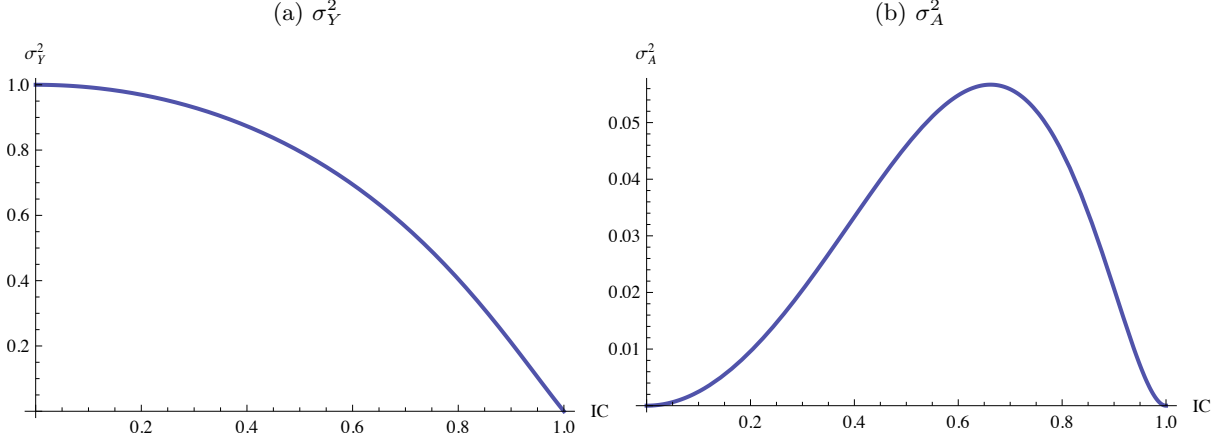
Figure 1 illustrates RIC for this skill structure for selected ranges of IC pairs consistent with positive-definiteness in the overall matrix of forecasts and returns.⁸ Fortunately we are also able to obtain useful analytic results in certain particular cases, such as the case when we have one skilled and one unskilled type which we consider below.

Firstly we adopt the perspective of the **skilled** type and compute his variances and RIC below, with illustrative plots in Figure 2.

$$\begin{aligned}\sigma_Y^2 &= \sigma^2 \left[1 + V - 2\delta^{(skilled)} IC^{(skilled)2} \right] \\ &= \sigma^2 \left[1 + \left(\frac{IC^{(skilled)}}{2 - IC^{(skilled)2}} \right)^2 - \left(\frac{2IC^{(skilled)2}}{2 - IC^{(skilled)2}} \right) \right]\end{aligned}$$

⁸Ensuring positive definiteness is an important consideration which we return to (and discuss further) in Proposition 4.4.

Figure 2: Variances of actual return (Y) and **skilled** type's anticipated return (A) for equi-frequent skilled/unskilled case; $\sigma^2 = 1$



$$\begin{aligned}
 \sigma_{A^{(skilled)}}^2 &= \sigma^2 \left[IC^{(skilled)2} + V - 2\delta^{(skilled)} IC^{(skilled)2} \right] \\
 &= \sigma^2 \left[IC^{(skilled)2} + \delta^{(skilled)2} IC^{(i)2} - 2\delta^{(skilled)} IC^{(skilled)2} \right] \\
 &= \sigma^2 IC^{(skilled)2} \left(1 - \delta^{(skilled)} \right)^2
 \end{aligned}$$

$$\sigma_Y \mid IC=0 = 1 \quad (16)$$

$$\sigma_Y \mid IC=1 = 0 \quad (17)$$

$$\sigma_A \mid IC=0 = 0 \quad (18)$$

$$\sigma_A \mid IC=1 = 0 \quad (19)$$

For the **skilled** type, RIC reduces to:

$$\begin{aligned}
 &= \frac{\sqrt{IC^{(skilled)2} - 2\delta^{(skilled)} IC^{(skilled)2} + V}}{\sqrt{1 + V - 2\delta^{(skilled)} IC^{(skilled)2}}} \\
 &= \frac{\sigma_{A^{(skilled)}}}{\sigma_Y}
 \end{aligned}$$

hence we have

$$RIC^{(skilled)} \mid IC^{(skilled)}=0 = RIC^{(skilled)} \mid IC^{(skilled)}=1 = 0$$

Proposition 4.1. *For the case of one skilled and one unskilled type with equal constant absolute risk-aversion the following approximation holds when $f^{(skilled)} \approx 0$*

$$RIC^{(skilled)} \approx IC(1 - f^{(skilled)})$$

Proof. See A.4.1. □

Proposition 4.2. *For two uncorrelated types with only one skilled and equal constant absolute risk-aversion, the maximum RIC is obtained when $IC = \frac{1}{\sqrt{1+f}}$ and the maximum RIC value at this point is $\frac{1-f}{1+f}$.*

Proof. See A.4.2. □

We turn now to the perspective of the **unskilled** type. Here we find that σ_Y^2 is identical to the case above, however we have

$$\begin{aligned}\sigma_{A(UNskilled)}^2 &= \sigma^2 \left[IC^{(UNskilled)2} + V - 2\delta^{(UNskilled)} IC^{(UNskilled)2} \right] \\ &= \sigma^2 V\end{aligned}$$

this indicates that the unskilled type will tend to take larger and larger positions as the available returns contract due to the increased skill of the skilled type.

Proposition 4.3. *For two uncorrelated types with only one skilled*

$$RIC^{(UNskilled)} = -RIC^{(skilled)} \tag{20}$$

irrespective of the frequency of the types.

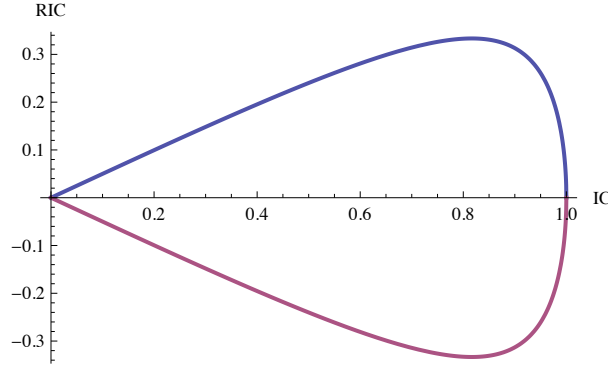
Proof. See A.4.3. □

We illustrate the above proposition in Figure 3 where we show how a completely unskilled investor ($IC = 0$) can nevertheless appear to be a useful contrarian indicator when placed in a market alongside an equally-frequent skilled type. This arises because the unskilled investor's forecast will always be the unconditional mean which, on average, will be *below* the skilled type's forecast in a bull market and *above* it in a bear market. Since the equilibrium price at which both investors trade will be the weighted average of these two forecasts, the unskilled investor will tend to short the market when it's more likely to rally and go long when it's more likely to fall.

To summarise this highly simplistic example: we have shown how RIC can become a zero sum game, with specific level highly dependent on the degree to which forecasts are distributed in the population. The lesson we draw from this is that it is vital to appreciate the competitive dynamic which drives a wedge between investor's innate skill (measured by IC) and the returns which they are actually able to capture (measured by RIC). While forecasting *is not* a competitive activity (the quality of one investor's forecast is not influenced by the quality of other investors' forecasts), trading clearly *is* a competitive activity. The assumption that $RIC = IC$ ignores this reality.

Once again we draw attention to the evident dangers of drawing conclusions about forecasting skill *per se* by exclusively focusing on RIC . In the example above it is clear that the skilled type can increase RIC by increasing their personal IC , but these gains come at the expense of the unskilled type. The unskilled type unwittingly suffers despite having no change in their own level of skill whatsoever. Although we do not examine it here, it is easy to imagine that the arrival of a new third agent type into this market, with $IC > 0$, might dramatically change

Figure 3: RIC for skilled (positive) and unskilled (negative) agents



measured RIC levels of all agents. Given the rapid flow of entry and exit in today's financial markets it seems concerning to make inferences based on measurements which are not robust to such structural changes.

4.2 Infinite Correlated Types with Pareto-distributed Precision

A drawback of the simple two-type model is that analysis of comparative statics and scenarios requires all agents of the same type to change their skill in unison. If, for instance, a single atomistic individual wished to change their personal IC then it would constitute the creation of a new third type and require a re-specification of the model. Such effects are, however, of considerable interest to us and we can address these successfully in a model with a continuum of available skill levels.

In this case we devise a means of parameterising a population-wide skill structure such that we can specify a *density* for the type frequency f rather than a discrete vector. Although there are infinitely-many ways of achieving this we aim here for parsimony combined with clarity in algebra and so we limit consideration to structures which lead to tractable calculations of various necessary moments and RIC . Again for clarity we assume equal constant absolute risk aversion.

A key consideration in this process regards inter-forecaster correlations ϕ_{ij} . When the number of types increases significantly it is no longer convenient to specify all pairwise correlations individually. Specifying a single pairwise constant correlation is one approach but this does not simplify the RIC calculation very much. The proposition below provides sufficient conditions for parameterisation of correlation such that (a) positive definiteness is maintained in a population of arbitrarily-many types and (b) the RIC calculation is somewhat simplified.

Proposition 4.4. *Suppose the correlation between forecasts of type i and type j is denoted by ϕ_{ij} , then in a population of multiple types with differential skill the overall matrix of forecasts and returns will be positive definite if*

$$\begin{aligned}\phi_{ij} &= kIC^{(i)}IC^{(j)} \text{ where } i \neq j \\ \phi_{ij} &= 1 \text{ where } i = j\end{aligned}$$

$$1 < k < \frac{1}{IC_{max}^2}$$

where IC_{max}^2 denotes the largest IC^2 value in the population.

Proof. See A.4.4. □

This indicates that the higher the maximum IC , the less freedom we have to influence levels of correlation, e.g. as $IC \rightarrow 1$ we require $k \rightarrow 1$. However if we set the maximum skill level even modestly high (e.g. $IC_{max} = 0.7$) then we can vary k between 1 and 2 which gives some useful flexibility.

How should we interpret k ? Consider the covariance between two agents' forecasts (types i and j): it will be $kIC^{(i)2}IC^{(j)2}\sigma^2$. (We use here the fact that their forecasts each have the form $\mu + IC^{(i)}\sigma S^{(i)}$ which have standard deviation $IC^{(i)}\sigma$ so we obtain the covariance by multiplying these standard deviations together and then multiply by the correlation ϕ_{ij} .) The variance of type i 's own forecast is $IC^{(i)2}\sigma^2$. Hence we can compute the beta between the two agents' forecasts as $kIC^{(j)2}$, i.e. type i is more influenced by type j when type j 's skill is high.

If, for instance, we take the case of $IC^{(j)} = 0.5$ and $k = 2$ then the beta between type i and type j will be 0.5 irrespective of the skill level of type i . The skill level of type i depends on how much of their own 'noise' they add to this common component (and the common components they share with all the other types).

This captures the idea that there is some degree of imitation between different agent types with the more skilled exerting a relatively high influence on the less skilled. This is consistent with the real-world situation where particularly successful forecasters develop a higher public profile over time than others, perhaps due to a chronic survivorship bias. It also captures the hypothetical (yet plausible) situation where prominent investment firms establish trading positions for the 'house' before publicising influential research recommendations afterwards; such recommendations then rapidly become self-fulfilling prophecies supported by the weight of client money.

Of course our model does not explicitly embody such dynamic, and potentially strategic, interactions. Nevertheless it is clear that in general the correlation structure of forecasts, combined with appropriate type classifications, can represent quite a high degree of richness in investor behaviour, albeit in reduced form. The imitation strategies of investors are discussed by Hirshleifer and Teoh (2009) together with a broad review of related topics in herding and contagion.

We now consider the distribution of skill in the population. For convenience we choose to work initially with each type's *precision*, which we define as $\pi = \frac{1}{1-IC^2}$ rather than IC directly. This is convenient because the range of precision is $[1, \pi_{max})$ which makes it more amenable to fitting with a standard continuous distribution. Here we use the upper-truncated Pareto distribution for this purpose, which has exactly this support. This also seems a reasonable choice given the popularity of Pareto distributions in modeling quantities like income and wealth; at first glance there seems fair justification to believe that the distribution of precision in the population might be similar.

Hence the density function for precision is:

$$f(\pi) = \frac{s}{1 - \pi_{max}^{-s}} \pi^{-(s+1)} ; 1 \leq \pi < \pi_{max} \quad (21)$$

where $s > 1$ is a shape parameter. From standard features of the truncated Pareto distribution (e.g. Zaninetti and Ferraro (2008)), the mean precision is given by

$$\mathbb{E}_f [\pi] = \frac{s}{s-1} \frac{1 - \pi_{max}^{1-s}}{1 - \pi_{max}^{-s}} \quad (22)$$

We now proceed to calculate the corresponding precision-weighted (or ‘market presence’) density which as usual we denote δ .⁹

Proposition 4.5. *When the precision of investors is distributed in the population according to the density $f(\pi)$ which is Pareto-distributed with parameter s , truncated at π_{max} then the δ density is*

$$\delta(\pi) = \frac{(s-1)}{1 - \pi_{max}^{1-s}} \pi^{-s} ; 1 \leq \pi < \pi_{max} \quad (23)$$

i.e. the δ distribution is the original unweighted f distribution but with the shape parameter reduced by one.

Proof. See A.4.5. □

As an aside, it is apparent here that we require $s > 2$ if we wish to ensure that π has a finite mean under the δ -density.

Having defined both correlation structure and skill distribution we are now able to compute the variances and RIC . Thanks to the particular correlation structure which we imposed we are able to write all these quantities in terms of $\mathbb{E}_\delta [IC^2]$, which is the expected value of IC^2 computed under the ‘market presence’ density δ . For brevity we denote this by $\overline{IC^2}$ and begin with the following helpful result:

Proposition 4.6. *When investor precision is Pareto-distributed according to $f(\pi)$ truncated at π_{max} , as in (21), we have*

$$\begin{aligned} \overline{IC^2} &= \frac{1}{1 - \pi_{max}^{1-s}} \left[\frac{1}{s} (1 - \pi_{max}^{-s}) - \pi_{max}^{-s} (\pi_{max} - 1) \right] \\ &\rightarrow \frac{1}{s} as \pi_{max} \rightarrow \infty \end{aligned}$$

Proof. See A.4.6 □

⁹The derivation of δ and associated moments which follows is somewhat informal in that we make the extension to a continuum of states by replacing mass functions with densities and summations with integrals. In the appendix (A.1.2) we sketch some formalities which clarify our method.

Proposition 4.7. *As $\pi_{max} \rightarrow \infty$ the mean of IC^2 under the δ -density and the mean IC^2 under the f -density have the following properties*

$$\begin{aligned} \mathbb{E}_\delta [IC^2] &> \mathbb{E}_f [IC^2] \\ \frac{\partial \mathbb{E}_\delta [IC^2]}{\partial \mathbb{E}_f [IC^2]} &> 1 \end{aligned}$$

Proof. The mean under the f -density is given by the expression in Proposition 4.6 with s replaced by $s + 1$ and the results above follow directly. \square

The above proposition will be of central importance later in our discussion of skill improvement. The intuition is that an adjustment in skill distribution which raises or lowers average IC^2 under the f density (representing unweighted headcount) will lead to a proportionally *larger* change in average IC^2 under the δ density because the weighting in the latter calculation is itself determined by (the increased) precision levels.

We now calculate the key variance components which appear in (11) and (12). Since we are now dealing with a continuum of types the type-specific σ_A^2 is now expressed as a function of agent's IC . Derivations are provided in A.4.7.

$$V = k \left(\overline{IC^2} \right)^2 \quad (24)$$

$$\sigma_Y^2 = 1 - \overline{IC^2} \left(2 - k \overline{IC^2} \right) \quad (25)$$

$$\sigma_A^2(IC) = IC^2 \left(1 - 2k \overline{IC^2} \right) + k \left(\overline{IC^2} \right)^2 \quad (26)$$

(Recall that for positive definiteness we have $1 < k < \frac{1}{\overline{IC^2}_{max}}$.)

By simple differentiation we see that σ_Y^2 decreases monotonically as $\overline{IC^2}$ increases, i.e. as average skill increases in the market we find equilibrium price more accurately represents the future actual price and hence available returns per unit asset reduce. This is similar to the effect noted in the previous example of two types illustrated in Figure 2(a).

Similarly we find that σ_A^2 is increasing in *own* skill (IC^2) only if $k \overline{IC^2} < \frac{1}{2}$, i.e. when average skill and correlation are relatively low. In other words an agent will only find that improved skill translates into increased *ex ante* alpha per unit asset if there is relatively little skilled competition and/or he is relatively uncorrelated with others. Again this is qualitatively the same as the case with only two types illustrated in Figure 2(b).

Proposition 4.8. *When investor precision is Pareto-distributed according to $f(\pi)$ truncated at π_{max} , as in (21), the Realised Information Coefficient (RIC) for an investor with personal Information Coefficient denoted by IC is given as follows:*

$$RIC(IC) = \frac{\left[IC^2 - \overline{IC^2} \right] \left[1 - k \overline{IC^2} \right]}{\sqrt{\left[1 - \overline{IC^2} \left(2 - k \overline{IC^2} \right) \right] \left[IC^2 \left(1 - 2k \overline{IC^2} \right) + k \overline{IC^2} IC^2 \right]}} \quad (27)$$

where we denote the expected value of IC^2 under the δ -density by $\overline{IC^2}$ and k is the correlation

parameter as defined in Proposition 4.4 with $1 < k < \frac{1}{IC_{max}^2}$.¹⁰

Proof. This expression follows from substituting continuum versions of (24), (25) and (26) into Proposition 3.2. \square

Various observations arise from this expression:

(1) Since the denominator will always be positive the sign of RIC is given unambiguously by the numerator. From the correlation structure we know that $k < \frac{1}{IC_{max}^2}$ and hence we can be sure that the second factor is positive since $\overline{IC^2} < IC_{max}^2$. Therefore the sign is uniquely determined by whether $IC^{(i)2}$ is greater or lesser than the δ -weighted mean $\overline{IC^2}$. This is the analog (with a continuum of types) of the result in the two type model where the skilled and unskilled agents' RIC 's sum to zero. Once again we see that RIC is very much determined by competitive forces.

(2) A type with $IC^2 = \overline{IC^2}$ will have RIC of zero irrespective of correlation; in other words a hypothetical 'average investor' has zero *ex post* skill. This is somewhat analogous to the result in the finite types model where increasing all types' skill in tandem fails to lead to any utility improvement. Here we would find that any changes in the distribution of skill nevertheless leave the 'average' type no better off in terms of RIC .

(3) According to the Fundamental Law of Active Management, type-specific Information Ratios will be approximately proportional to RIC . We might therefore obtain a crude indication of overall social welfare by computing the expected value of RIC across the population according to the unweighted f -density. The types with IC below the 'threshold' $\overline{IC^2}$ will contribute negatively to this average and those above $\overline{IC^2}$ will contribute positively. Loosely speaking we would therefore expect the f -weighted mean RIC to lie 'close to' zero.

(4) The average $\overline{IC^2}$ which we focus on in Proposition 4.8 is calculated under the δ -density. However Proposition 4.7 reminds us that this is not the same as the unweighted population average (calculated under f). Indeed raising average skill of the whole population under the f density will raise $\overline{IC^2}$ even more. Very informally, the effect of this is to raise the 'threshold' $\overline{IC^2}$ level disproportionately high relative to the new f -weighted average and intuitively this causes us to suspect that the social welfare impact of such skill improvement might not necessarily be positive. We investigate this phenomenon more exactly via numerical methods later in the paper.

(5) By (rather lengthy) differentiation of (27) it can be shown that

$$\frac{\partial RIC(IC)}{\partial \overline{IC^2}} < 0$$

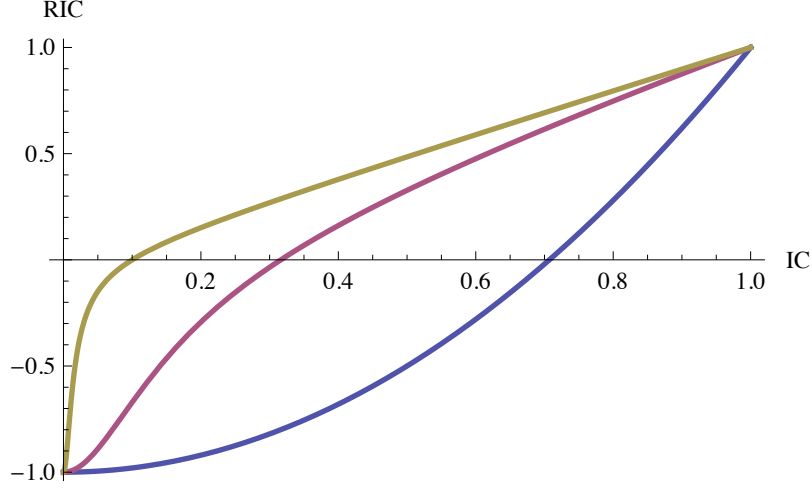
and

$$\frac{\partial RIC(IC)}{\partial IC^2} > 0$$

i.e. RIC is increasing in an agent's own skill IC^2 but decreasing in the δ -weighted average skill of the population $\overline{IC^2}$.

¹⁰We reiterate that the mean quantity $\overline{IC^2}$ is *not* the headcount-weighted mean which would be given by the f distribution, rather it is the precision-weighted mean. However given that the distributional assumptions which we have made there is a simple relationship between f and δ and it is straightforward to convert between the two means as required.

Figure 4: RIC against IC for three values of s : 2, 10 and 100 (from lowest to highest line); in each case $k = 1$



Plotting RIC against IC helps us gain further insight. Figure 4 shows example plots for $k = 1$ and s either 2, 10 or 100. When $s = 2$ the mean value of skill is relatively high ($\overline{IC^2} = 0.5$) and the distribution of skill decays fairly slowly across the population. In this case the intensity of competition for returns is very high and hence only the highest levels of IC translate into positive RIC . This case is the lowest line in the figure. By contrast when $s = 100$ the skill distribution is heavily concentrated around low-skilled types. These are much better conditions for the relatively high skilled since they face less competition for returns and the opportunity of trading with a multitude of less well-informed counterparties. Hence we see that positive RIC is obtained at lower levels of IC . This is shown in the highest line on the plot.

This plot also sheds light on approximate comparative statics of skill improvement. Suppose a single atomistic investor were to change his personal IC , then the new level of his RIC can be approximated by reference to the appropriate curve. (We assume here that the magnitude of the individual investor's demand in the market is negligible so that the effect of this change on the f and δ distributions is immaterial.) Such deductions cannot be made from the model of two types in the previous section unless *all* investors of the same type change skill together.

It is also helpful to state the inverse form of Proposition 4.8 as follows:

Corollary 4.9. *The personal Information Coefficient (IC) of an investor with Realised Information Coefficient RIC is given by:*

$$IC(RIC) = \left[\overline{IC^2} + \frac{1}{2}\theta^2 RIC^2 + \theta RIC \sqrt{\frac{1}{4}\theta^2 RIC^2 + \overline{IC^2} \left(\frac{1 - k\overline{IC^2}}{1 - 2k\overline{IC^2}} \right)} \right]^{\frac{1}{2}}$$

where

$$\theta^2 = \frac{\left[1 - \overline{IC^2} (2 - k\overline{IC^2}) \right] (1 - 2k\overline{IC^2})}{\left[1 - k\overline{IC^2} \right]^2}$$

Proof. See A.4.8 □

Having established the analytic distinction between IC and RIC we now consider the shapes of the unweighted distributions of these quantities across the population, given the assumption of Pareto-distributed precision which we have already presented. We tackle this in stages, beginning with computation of the distribution of IC^2 which we denote f_{IC^2} .

Proposition 4.10. *When $\pi = \frac{1}{1-IC^2}$ is Pareto-distributed with shape parameter s , truncated at $\pi_{max} \equiv \frac{1}{1-IC_{max}^2}$, the unweighted density of IC^2 is*

$$f_{IC^2}(IC^2) = \frac{s}{1 - \pi_{max}^{-s}} \left(\frac{1}{1 - IC^2} \right)^{-s+1} ; 0 \leq IC^2 < IC_{max}^2$$

Proof. See A.4.9 □

This density is illustrated in Figure 5. We now consider the unweighted distribution of RIC .

Proposition 4.11. *When $\pi = \frac{1}{1-IC^2}$ is Pareto-distributed with shape parameter s , truncated at $\pi_{max} \equiv \frac{1}{1-IC_{max}^2}$, the density of RIC is given by*

$$f_{RIC}(RIC) = \frac{s}{1 - \pi_{max}^{-s}} \left(\frac{1}{1 - \overline{IC^2} - \frac{1}{2}\theta^2 RIC^2 - \theta RIC \sqrt{\frac{1}{4}\theta^2 RIC^2 + \overline{IC^2} \left(\frac{1-k\overline{IC^2}}{1-2k\overline{IC^2}} \right)}} \right)^{-s+1} \times$$

$$\left[\theta^2 RIC + \theta \sqrt{\frac{1}{4}\theta^2 RIC^2 + \overline{IC^2} \left(\frac{1-k\overline{IC^2}}{1-2k\overline{IC^2}} \right)} + \frac{\frac{1}{4}\theta^3 RIC^2}{\sqrt{\frac{1}{4}\theta^2 RIC^2 + \overline{IC^2} \left(\frac{1-k\overline{IC^2}}{1-2k\overline{IC^2}} \right)}} \right]$$

where θ is defined in Corollary 4.9.

Proof. See A.4.10 □

Finally we turn our attention to the distribution of IC (our *ex ante*) measurement.

Proposition 4.12. *When $\pi = \frac{1}{1-IC^2}$ is Pareto-distributed with shape parameter s , truncated at $\pi_{max} \equiv \frac{1}{1-IC_{max}^2}$, the density of IC is given by*

$$f_{IC}(IC) = \frac{2s}{1 - \pi_{max}^{-s}} \left(\frac{1}{1 - IC^2} \right)^{-s+1} IC ; 0 \leq IC < IC_{max}$$

Proof. See A.4.11 □

Proposition 4.13. *The mode of IC under the unweighted f distribution (described in Proposition 4.12) is given by*

$$IC_{mode} = \frac{1}{\sqrt{2s-1}} \tag{28}$$

Proof. See A.4.12 □

Figure 5: Unweighted Pareto distributions of IC^2 for alternative shape parameters ($k = 1$)

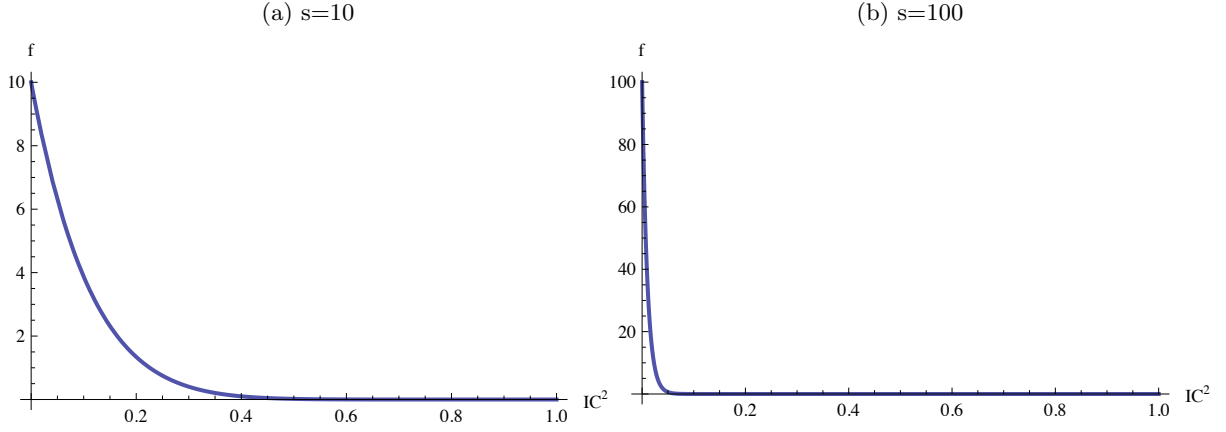
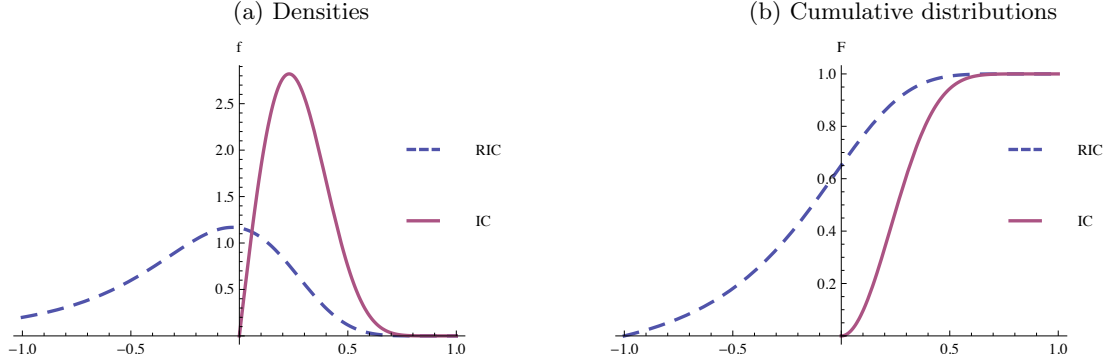


Figure 6: Unweighted IC distribution (solid line) and RIC distribution (dashed line) where $k=1$ and $s = 10$; $IC_{mode} \approx 0.229$

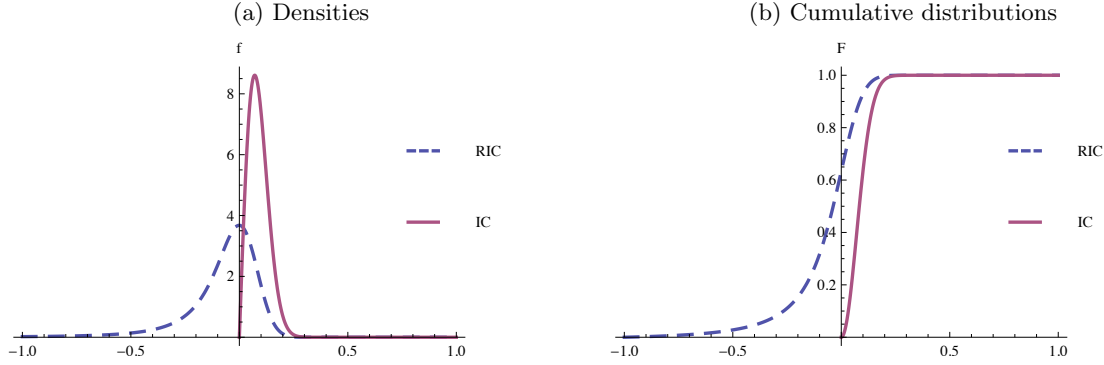


We plot examples of these densities in Figures 6 and 7. These plots are a vivid summary of our analysis: we show that the initial distribution of IC is uniformly positive, relatively evenly distributed around the mode and with an upper tail extending to the perfect foresight case of $IC = 1$. In comparison, however, the distribution of RIC is shifted to the left, extends across the full range of correlations from -1 to $+1$ and is less peaked. (In this case with $k = 1$ we have $IC_{max} = +1$ so we observe a full range of correlations.)

We do not provide analytic expressions for the mean, median or mode of RIC in this paper however it is clear from inspection of the plot that these are likely to be closer to zero than in the IC distribution. We provide some preliminary numerical evidence of this in the next section. This is entirely consistent with Proposition 4.8 which shows that a key driver of the RIC value is the *difference* between IC^2 and the mean $\overline{IC^2}$ rather than absolute values of IC .

The cumulative distributions plotted in panel (b) of Figures 6 and 7 demonstrate how higher quartiles of IC map to lower quartiles of RIC .

Figure 7: Unweighted IC distribution (solid line) and RIC distribution (dashed) where $k=1$ and $s = 100$; $IC_{mode} \approx 0.071$



5 Utility and Social Welfare

5.1 Future Wealth Density Function

Having now determined RIC it is feasible for us to go a step further and derive the shape of the distribution of random future wealth W . From this we can obtain an expression for unconditional expected utility as a function of the skill parameters in the market. This gives us an exact picture of the marginal utility of skill and associated welfare issues without resorting to the Information Ratio proxy.

The question of determining the distribution of an *active* investor's future wealth has been addressed in the literature only to a limited extent, for instance Lee (2000), one of the few treatments of this topic, computes the first two moments but not the full distribution.

In the case of a *passive* investor in a CAPM-type setting, future wealth is simply the product of portfolio weights and actual asset returns. Since portfolio weights are a deterministic function of exogenous means and covariances this is straightforward to derive. Passive investors are rewarded for taking systematic market risk with whatever risk premium prevails in equilibrium.

For active managers, however, the portfolio weights are stochastic and - assuming non-zero forecasting skill - there is a dependency between these and actual returns (which are of course also stochastic).¹¹ This means that the distribution of future wealth is more complex. However some useful (and fairly innocuous) simplification can be obtained if we pay attention to separating market risk premium from alpha due to active management. In the analysis below we are interested in the distribution of the product of the random variables in (11) and (13) which respectively represent money gains per unit asset and holding size (in units). The mean of each of these components is a non-stochastic quantity proportional to risk premium.

In the analysis below it can be shown that the general functional form of the wealth distribution does not depend on the magnitude of this risk premium, however it does of course influence the values of various parameters of the distribution. Non-zero risk premium increases the complexity of the parameters and thereby diverts attention from the pure skill effects. Inspection of (11) and (13) reveals that the absolute magnitude of the risk premium depends on exogenous

¹¹Satchell and Hwang (2001) consider the implications of this for the calculation of tracking error in passive portfolios.

factors such as unconditional price variance σ^2 , risk-aversion levels and overall headcount in the market (K). Since our focus in this paper is very much on the actively-managed component of returns (i.e. the alpha), for clarity of exposition we assume that the risk premium equals zero, but note that incorporating a non-zero risk premium is perfectly tractable at the cost of additional algebra.

Proposition 5.1. *We denote by W a type i investor's random future wealth which is the product of per-unit money gain Y and number of units held $H^{(i)}$. (See (11) and (13) for precise definitions of these random variables).*

The density of W is given by

$$f_W(w) = \frac{e^{\frac{RICw}{(1-RIC^2)\sigma_Y\sigma_{H^{(i)}}}} K_0 \left[\frac{Abs(w)}{(1-RIC^2)\sigma_Y\sigma_{H^{(i)}}} \right]}{\pi\sqrt{1-RIC^2}\sigma_Y\sigma_{H^{(i)}}} \quad (29)$$

where K_0 denotes the modified Bessel function of the third kind (with imaginary argument) and the variances σ_Y^2 and $\sigma_{H^{(i)}}^2$ are given by (11) and (13).

Proof. See A.2. □

As an aside, this distribution appears elsewhere in the (unrelated) finance literature in the form of the variance-gamma model of Madan and Seneta (1990) and Madan, Carr and Chang (MCC) (1998). Perhaps the earliest reference (in a non-financial context) is Craig (1936).

The first two cumulants are given by:

$$\mu_W = RIC\sigma_Y\sigma_H \quad (30)$$

$$Var(W) = (1 + RIC^2)\sigma_Y^2\sigma_H^2 \quad (31)$$

However since the wealth distribution is manifestly non-normal it also important to consider higher moments. These are computed by Craig (1936) and MCC and we can show that

$$Skewness = \frac{2RIC(3 + RIC^2)}{(1 + RIC^2)^{\frac{3}{2}}}$$

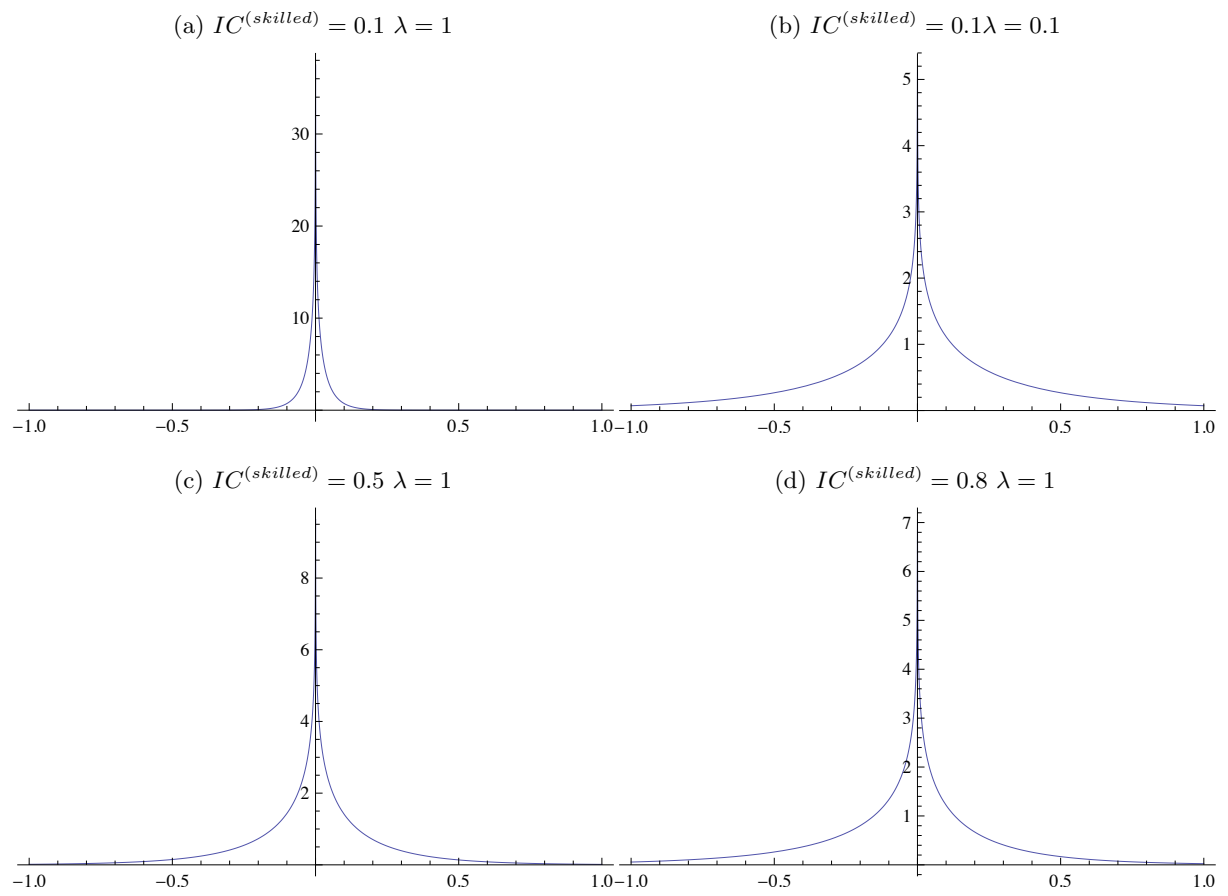
which will be approximately zero for small RIC . Also

$$Kurtosis = \frac{9 + 42RIC^2 + 9RIC^4}{(1 + RIC^2)^2}$$

Hence excess kurtosis will be approximately 6 for small RIC . We note also that skewness is increasing in RIC while kurtosis is increasing in the absolute value of RIC .

A selection of plots of the future wealth density appears in Figure 8. These relate to the equiprequent skilled/unskilled structure from the perspective of the skilled type.

Figure 8: Probability Density of Skilled Agent's Future Wealth (two equi-frequent types, $\sigma^2 = 1$)



5.2 Portfolio Information Ratio

We define the per-period Information Ratio of the investor's portfolio to be¹²

$$IR = \frac{RIC}{\sqrt{1 + RIC^2}}$$

If we consider the first example skill structure in Section 4 (two equifrequent types: one skilled and one unskilled) then we have shown in Proposition 4.1 that we may approximate RIC by $IC(1 - f)$ for an uncorrelated forecaster when f is relatively small. We can therefore write

$$IR \approx \frac{IC(1 - f)}{\sqrt{1 + IC^2(1 - f)^2}}$$

As per Grinold and Kahn (1999), this per-period Information Ratio should be multiplied by $\sqrt{breadth}$ as appropriate, where *breadth* indicates the number of times per year that investments are made based on forecasts.

Hence we can generalise Grinold's (1989) Fundamental Law of Active Management (in this endogenous returns model of equifrequent skilled/unskilled types) to

$$IR \approx \frac{IC(1 - f)}{\sqrt{1 + IC^2(1 - f)^2}} \sqrt{breadth}$$

where it is clear that for an uncorrelated atomistic investor with negligible market impact ($f \rightarrow 0$) we have Grinold's original formulation.

5.3 Unconditional Expected Utility

Having determined future wealth distribution it is straightforward to compute unconditional expected utility. This entails taking the expectation of

$$u(W) = -\frac{1}{\lambda} \exp(-\lambda W) \quad (33)$$

with respect to the probability measure given by (29). The resulting expression is a product of $-\frac{1}{\lambda}$ and the moment generating function of the product of two correlated normals. Craig

¹²We define the Information Ratio as:

$$IR = \frac{r_P - r_B}{\sigma_\epsilon} \quad (32)$$

where r_P is the investor's portfolio return, r_B is the return of a benchmark portfolio and σ_ϵ is the standard deviation of the so-called *tracking-error* (the periodic difference between the portfolio and benchmark returns). This measurement has its roots in Sharpe (1981). Grinold (1989) shows that (in certain conditions) for investors with mean-variance utility the Information Ratio defines a preference ordering between portfolios equivalent to expected utility maximisation. Here we consider the investor's benchmark portfolio to be a strategy of doing nothing, i.e. having no holdings in risky assets whatsoever ($r_B = 0$), and hence *ex ante* the expression reduces to:

$$IR = \frac{RIC}{\sqrt{1 + RIC^2}}$$

since we replace r_P with *expected* wealth from (30) and we use standard deviation in wealth obtained from (31) in place of σ_ϵ . Effectively we make this step by multiplying both numerator and denominator of (32) by the investor's arbitrary initial wealth; note that the investor's portfolio choice has no dependence on initial wealth due to the exponential form of the utility function so its magnitude here is irrelevant.

(1936) provides a closed form expression for this: once again if we make the assumption of zero risk premium then

$$\begin{aligned}
\mathbb{E}[\exp(t\zeta_1\zeta_2)] &= [1 - (1 + \rho)t]^{-\frac{1}{2}} [1 + (1 - \rho)t]^{-\frac{1}{2}} \\
&= [1 - (1 + \rho)t + (1 - \rho)t - (1 + \rho)(1 - \rho)t^2]^{-\frac{1}{2}} \\
&= [1 - 2\rho t - (1 - \rho^2)t^2]^{-\frac{1}{2}}
\end{aligned} \tag{34}$$

where we denote two *standard* correlated normal random variables by ζ_1 and ζ_2 and their correlation by ρ . Hence from (33) expected utility here will be:

$$\mathbb{E}[u] = -\frac{1}{\lambda} \mathbb{E}[\exp(t\zeta_1\zeta_2)]$$

with $\rho = RIC$ and $t = -\frac{\sigma_A\sigma_Y}{(1-IC^2)\sigma^2}$ since $W = YH^{(i)}$.

For illustration we plot in Figure 9 certainty equivalent wealth levels for type-0 investors in the case of two equi-frequent uncorrelated investor types. We compute this by the following transformation:

$$w_{ce} = -\frac{1}{\lambda} \log(-\lambda \mathbb{E}[u(W)]) \tag{35}$$

This is directly comparable with the *RIC* plot in Figure 1 and some qualitative similarity is immediately apparent.

By differentiation of (34) we can obtain a general expression for unconditional expected marginal utility of skill:

$$\begin{aligned}
\frac{\partial \mathbb{E}[u]}{\partial IC} &= \frac{1}{2\lambda} \frac{\left[-2\frac{\partial \rho}{\partial IC}t - 2\rho\frac{\partial t}{\partial IC} - 2t\frac{\partial t}{\partial IC}(1 - \rho^2) + 2\rho\frac{\partial \rho}{\partial IC}t^2\right]}{[1 - 2\rho t - (1 - \rho^2)t^2]^{\frac{3}{2}}} \\
&= \lambda^2 [\mathbb{E}[u]]^3 \left[\frac{\partial \rho}{\partial IC}t + \rho\frac{\partial t}{\partial IC} + t\frac{\partial t}{\partial IC}(1 - \rho^2) - \frac{\partial \rho}{\partial IC}\rho t^2\right] \\
&= \lambda^2 [\mathbb{E}[u]]^3 \left[\frac{\partial \rho}{\partial IC}(t - \rho t^2) + \frac{\partial t}{\partial IC}(\rho + t - t\rho^2)\right]
\end{aligned} \tag{36}$$

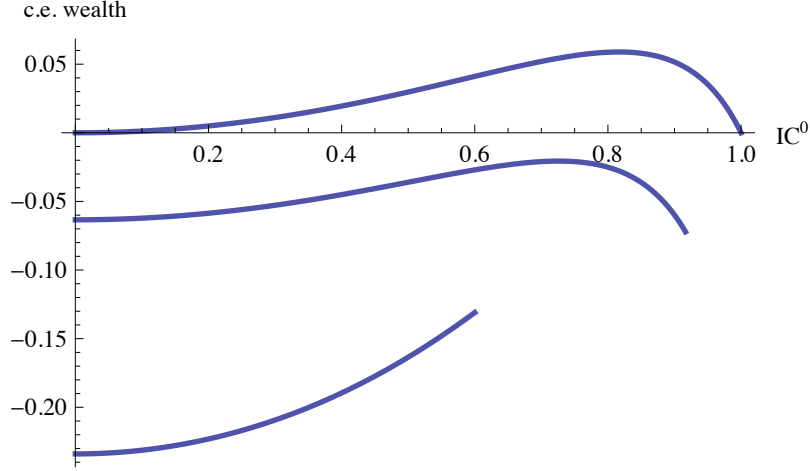
and we now use this to revisit the skill structures examined in Section 4 from a marginal utility perspective.

5.3.1 Two equi-frequent types: one skilled and one unskilled (skilled perspective)

In this case we have

$$\begin{aligned}
t &= -\frac{\sigma_A\sigma_Y}{1 - IC^2} \\
\rho t^2 &= \frac{\sigma_A}{\sigma_Y} \frac{\sigma_A^2\sigma_Y^2}{(1 - IC^2)^2} = \frac{\sigma_A^3\sigma_Y}{(1 - IC^2)^2} \\
\rho &= \frac{\sigma_A}{\sigma_Y} \\
t\rho^2 &= -\frac{\sigma_A^2}{\sigma_Y^2} \frac{\sigma_A\sigma_Y}{(1 - IC^2)} = -\frac{\sigma_A^3}{\sigma_Y(1 - IC^2)}
\end{aligned}$$

Figure 9: Certainty-Equivalent Wealth of investor type 0 in equi-frequent two independent agent model; $IC^{(1)}$ fixed at 0, 0.4 and 0.8 (from highest to lowest line)



Hence from the variances defined in (11) and (12) we have

$$\begin{aligned}
 & \rho + t - t\rho^2 \\
 = & \frac{\sigma_A}{\sigma_Y} - \frac{\sigma_A \sigma_Y}{(1 - IC^2)} + \frac{\sigma_A^3}{\sigma_Y(1 - IC^2)} \\
 = & \frac{\sigma_A}{\sigma_Y(1 - IC^2)} [(1 - IC^2) - \sigma_Y^2 + \sigma_A^2] \\
 = & 0
 \end{aligned}$$

and therefore (36) simplifies to

$$\begin{aligned}
 \frac{\partial \mathbb{E}[u]}{\partial IC} &= \lambda^2 [\mathbb{E}[u]]^3 \frac{\partial \rho}{\partial IC} (t - \rho t^2) \\
 &= \lambda^2 [\mathbb{E}[u]]^3 \frac{\partial \rho}{\partial IC} t(1 - \rho t)
 \end{aligned}$$

with

$$\text{sign} \left[\frac{\partial \mathbb{E}[u]}{\partial IC} \right] = \text{sign} \left[\frac{\partial \rho}{\partial IC} \right]$$

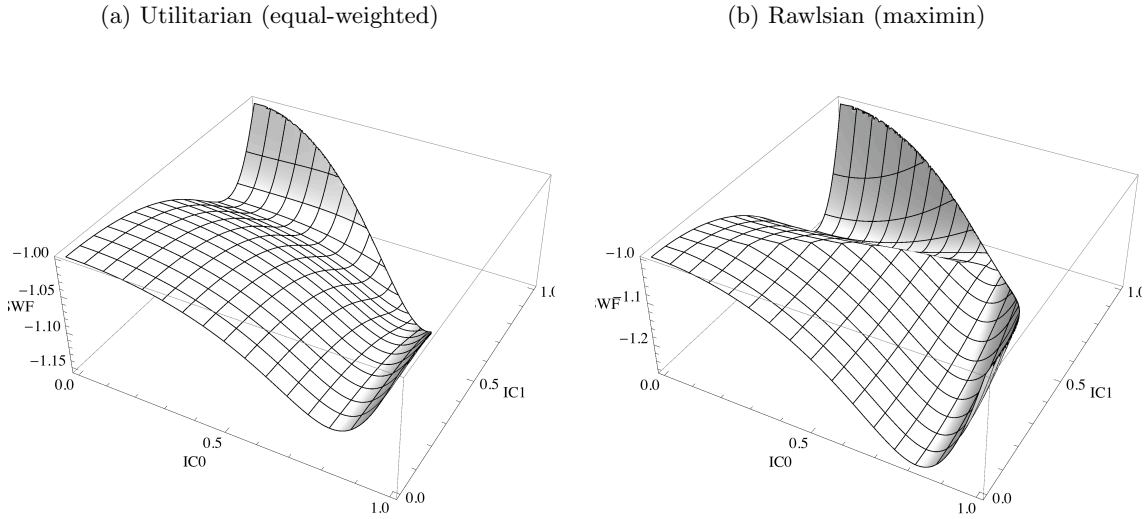
Hence we find that when RIC reaches a maximum, so too does expected utility. This indicates that for this particular skill structure it is not unreasonable to proxy expected utility by Information Ratio or RIC . This will not be true in general for all skill structures however.

We also consider the limiting cases: $IC = 0$ and $IC = 1$. In this case it is clear from (16), (17), (18) and (19) that

$$\begin{aligned}
 \mathbb{E}[u] &= -\frac{1}{\lambda} [1 + (1 + \rho)\sigma_Y\sigma_{H(i)}]^{-\frac{1}{2}} [1 - (1 - \rho)\sigma_Y\sigma_{H(i)}]^{-\frac{1}{2}} \\
 &= -\frac{1}{\lambda}
 \end{aligned}$$

An investor skill Laffer curve? Consider the case of one skilled and one unskilled type. The non-monotonic utility/skill trade-off depicted in Figure 9 is a little reminiscent of the Laffer

Figure 10: Social Welfare Functions for two equi-frequent uncorrelated agents with equal constant absolute risk-aversion



curve.¹³ As investors improve their forecasting skill there can exist distributions of skill in the market which cause them to become ‘victims of their own success’. This arises because although their increased precision leads to larger position-taking, this in turn causes the variance of market prices to increase. On an *ex ante* basis, beyond a certain point of IC , this extra variance in market prices leads to variance in returns which more than offsets the benefits of larger-position taking.

Once more our results here echo those of the simpler discrete model in S&W. Improvements from relatively low levels of skill do improve the RIC (and expected utility) of the smarter agent but at the direct expense of the other. Beyond a certain critical point, however, the RIC of the ‘smarter’ agent begins to decline.

Finally we follow the approach in S&W of computing social welfare functions where we combine the expected utilities of both agent types into a single overall measurement. These are illustrated in Figure 10 and are qualitatively very similar in shape to those obtained in the model of two discrete states detailed in S&W. If we ignore relatively implausible extreme values of skill (close to perfect certainty $IC \approx 1$) we see that social welfare is in fact maximised when both agents are equally unskilled ($IC = 0$) and any improvement in skill by either agent leads to a reduction in social welfare.

5.3.2 Continuum of infinite number of correlated types

Results here are more opaque when derived analytically. However it is straightforward to generate expected utility values numerically by substituting necessary variance and RIC parameters into (34). The plots which we present in Figure 11 once again reveal the usefulness of RIC as

¹³An explanation of the concept can be found in Canto, Joines and Laffer (1982). The Laffer Curve suggests that the relation between tax revenue and tax rate is concave with an optimal tax rate in excess of which marginal revenue begins to decline. Eventually at a tax rate of 100% the revenue raised is zero since the high tax rate acts as the ultimate disincentive to work.

Figure 11: Certainty-Equivalent Wealth (dashed line) and RIC (solid line) as a function of *ex ante* IC

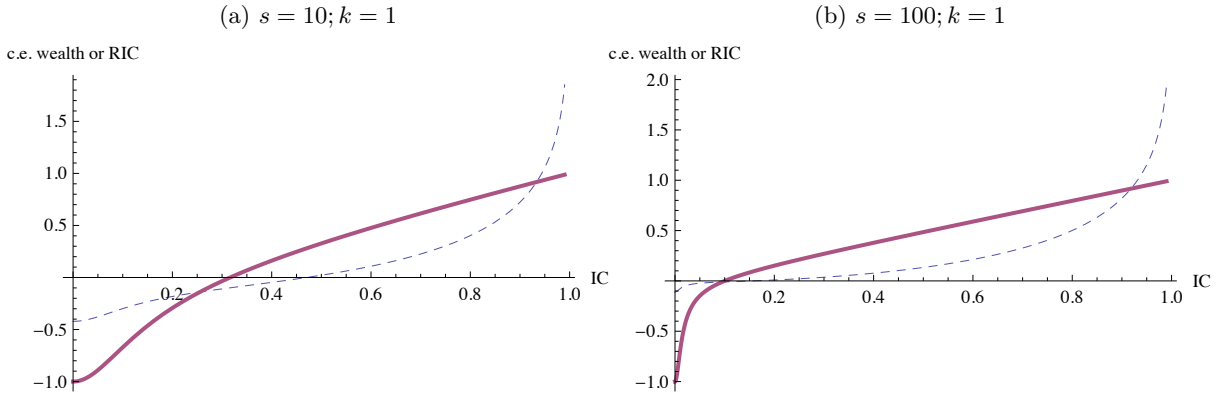
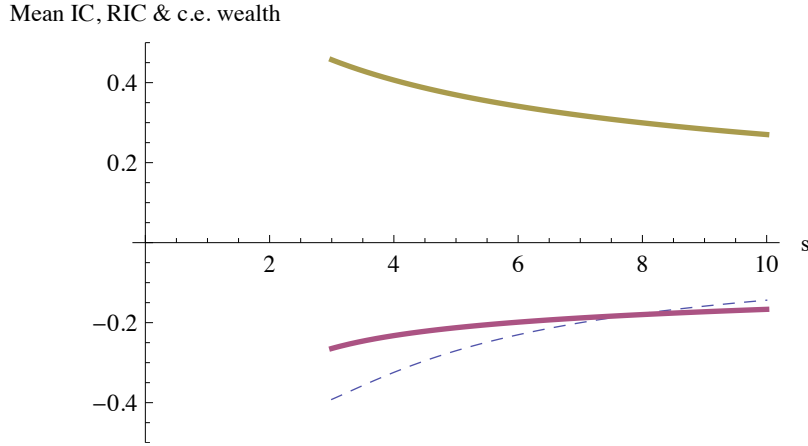


Figure 12: f -weighted average IC , RIC and certainty-equivalent wealth (from highest to lowest curve) against shape parameter of precision distribution



a crude proxy for expected utility.

In Figure 12 we extend the analysis to the level of aggregate social welfare. We achieve this numerically by calculating the average values of IC , RIC and certainty-equivalent wealth weighted by the f distribution (which represents the headcount of investors of each type in the population). These averages are plotted against increasing values of the s (shape) parameter of the precision distribution. Recall that increasing the s parameter will reduce mean values of IC and IC^2 according to the Propositions in Section 4.

Strikingly, however, Figure 12 shows that reducing mean IC^2 (our *ex ante* personal skill measurement) brings about an *increase* in both RIC and expected utility. After introducing the RIC formula in Proposition 4.8 we raised the concern that an increase in mean IC^2 under the f -density will be outpaced by the corresponding increase in the ‘threshold’ $\overline{IC^2}$ computed under the (precision-weighted) δ -density and this would tend to depress RIC values; in this plot we see evidence of exactly this. In particular it is noteworthy that the f -weighted mean RIC is negative despite the corresponding mean IC being positive.

6 Conclusions

The Information Coefficient (IC) is widespread in the literature, however it tends to be used in a partial equilibrium context. While this is an acceptable treatment when investors are atomistic and have no significant market impact we have shown that it has serious defects otherwise.

In this paper we have attempted to introduce some preliminary theory to bring IC into a general equilibrium setting. During the course of our work it became obvious that considerable richness and diversity of investor behaviour can be captured in this way and choosing illustrative examples is challenging in itself. Nevertheless, irrespective of the exact structural assumptions which we have made, we have found certain recurrent themes in our analysis. Often these have taken the form of precise analytic results which have either reinforced or challenged conventional wisdom. Occasionally we have come across deeper, more complex issues which deserve further scrutiny. In the paragraphs below we summarise our findings and attempt to tease out implications for practitioners and theorists.

Our first recurring theme is that connections between measured IC (which we term RIC) and innate personal forecasting skill may be tenuous at best. The measured IC number (RIC) is a noisy signal, corrupted by the influences of other competing (and potentially correlated) agents. Furthermore the relationship between signal and noise is highly non-linear. This has a range of implications:

(1) It provides a theoretical basis for the null hypothesis that positive in-sample investing performance (measured for instance by Information Ratio) will not persist out-of-sample. As a corollary to this, the value of analyst ratings based on past predictive accuracy may currently be over-estimated by the market. However we have been able to propose a way in which the skill/performance relationship can be modeled and parameterised by other potentially observable factors, specifically the overall profile of skill in the market and its distribution in terms of headcount. This may incrementally complicate future empirical analysis but it does not render it useless.

(2) Investors should make attempts to extract an estimate of their IC from available data using methodologies which go beyond simple correlation calculations. We have shown, for instance, that in a highly simplified two-type skilled/unskilled market $RIC \approx (1 - f)IC$, i.e. measured Information Coefficient RIC is *approximately* the product of innate skill IC and the proportion of the market which is unskilled. This shows the need to adjust empirical correlations (which would ordinarily be RIC values) to take account of market structure.

(3) Changes in market structure (e.g. correlation due to herding) will inevitably have an impact on RIC values, but such effects can be partially estimated in advance as long as we have knowledge of innate skill embodied in IC . It is IC rather than RIC which is the exogenous factor and it is very much in investors' own interests to estimate its magnitude in order to facilitate this kind of analysis.

(4) An investor who has high RIC in one market may have much lower RIC in another market, even if he has exactly the same forecasting accuracy. This can arise due to differences in distribution of skill or inter-forecaster correlations. There is a conventional wisdom that diversification across markets is desirable in an aim to increase overall Information Ratios and

hence such forays are commonplace. However our theory suggests that investors might be better advised to remain exclusively in high-*RIC* markets, where they enjoy comparative advantage, and instead attempt to improve their Information Ratios by increasing breadth.

(5) Given the quantitative uncertainty surrounding the exact level of *IC*, investors may be curious as to how to adapt existing models which depend on cardinal estimates of future returns. Potentially attractive alternative paradigms include models based on ordered forecasts such as Almgren and Chris (2007).

Our second recurring theme is that, loosely speaking, *RIC* and utility depend on skill relative to the ‘average’ rather than absolute levels. However the subtlety here is that the key benchmark ‘average’ skill level is itself skill-weighted, i.e. the skill of higher skilled investors features more prominently in the average calculation than that of lower-skilled investors. Some implications are:

(1) From the point of view of the individual there may indeed be benefit in increasing skill as long as that investor is atomistic, however this may not be the case if the investor is locked in an ‘arms race’ scenario with other correlated agents who improve in unison. Again the need to identify both one’s own *IC* level and the structure of the market is compelling.

(2) Negative measured skill (*RIC*) does not necessarily imply a low level of innate forecasting ability (*IC*). Indeed it would not be surprising to find that the *RIC* of roughly half of market participants is below zero.

(3) From the social welfare point of view we can easily demonstrate plausible conditions where increasing average skill levels across the general population leads to a decline in welfare, measured in terms of headcount-weighted utility. (Apparent from the fact that the skill-weighted average skill increases more rapidly than the headcount-weighted average.) This has important policy consequences in areas of investor education and financial innovation and we discuss these matters in Satchell and Williams (2010).

(4) Measured *RIC* will typically *understate* an investor’s innate forecasting ability (as represented by *IC*) and the extent of this discrepancy depends on the distribution of headcount across types. If we take the example of the equiproportional skilled/unskilled model then skilled investors suffer a reduction in *RIC* when they, as a group, become a larger proportion of the population: ‘small is beautiful’ when it comes to skill. Active fund management over the last half century has had a history of innovation, e.g. the 1970’s hedge fund industry was dominated by just two style categories: event-driven and managed futures funds (essentially trend-followers), however this has more recently fragmented into many more diverse styles with a relatively uniform distribution of AUM.¹⁴ Although further research is required, we believe our model provides an embryonic formal framework for analysing this form of evolution.

In the course of this analysis we have tried to keep our model as simple as possible to explain our ideas clearly, but at the same time retain important features of reality. There are many theoretical aspects which can be expanded upon in further study. For instance although our model is static it shows the *RIC* which results *ex post* from a given profile of *IC* which agents

¹⁴This is apparent in estimated Assets Under Management (AUM) data compiled by various hedge fund researchers, e.g. Lo (2008).

use *ex ante*. If agents use historical measured *RIC* as a means of determining *IC* to use for their current period portfolio choice, then our model can easily be adapted to provide a dynamic evolution of *RIC* and associated variances. This could be combined with investor entry and exit dynamics (which characterise *f*) and the result would contribute to the literature on dynamic agent-based models.

The static nature of the current model precludes agents from learning about the structure of skill in the market around them. In reality we would expect to observe behaviour of this kind although it seems likely that the complexity and dynamic nature of market structure may make it a tall order for agents to determine this with great accuracy. Similarly there is great scope to extend the model into conventional frameworks of intertemporal portfolio choice.

In summary, we have presented a theoretical basis which we believe is a foundation for considerable further study. Our aim has been neither to bury the Information Coefficient nor to praise it, but to draw attention to its power and usefulness, provided it is properly understood.

A Proofs and Derivations

A.1 Derivations

A.1.1 Variances of actual and anticipated gains in (11) and (12)

$$\begin{aligned}
\sigma_Y^2 &= \text{Var}(p_{t+1} - p_t) \\
&= \text{Var}(p_{t+1}) + \text{Var}(p_t) - 2\text{Cov}(p_{t+1}, p_t) \\
&= \sigma^2 + V\sigma^2 - 2 \sum_{i=0}^{J-1} \delta^{(i)} IC^{(i)2} \sigma^2
\end{aligned}$$

$$\begin{aligned}
\sigma_{A^{(i)}}^2 &= \text{Var}(IC^{(i)} \sigma S^{(i)} - p_t) \\
&= \text{Var}(IC^{(i)} \sigma S^{(i)}) + \text{Var}(p_t) - 2\text{Cov}(IC^{(i)} \sigma S^{(i)}, p_t) \\
&= IC^{(i)2} \sigma^2 + V\sigma^2 - 2\text{Cov} \left(\sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)} \sigma S^{(j)}, IC^{(i)} \sigma S^{(i)} \right) \\
&= IC^{(i)2} \sigma^2 + V\sigma^2 - 2 \sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)} IC^{(i)} \phi_{ij} \sigma^2
\end{aligned}$$

A.1.2 Extension from discrete number of types to a continuum of types

In the discrete case we begin with the function $f^{(i)}$ which denotes the proportion of the population of type i . This group of investors has skill $IC^{(i)}$. For the continuum case we take the following approach: suppose that we represent the population of investors by the interval $(0, \infty]$ of the real line and we denote by F the probability measure representing the proportions of each type. F is defined over all subsets of $(0, \infty]$. This measure has density f with respect to Lebesgue measure and we will typically use f in further computations. Now we define the precision function $\frac{1}{1-IC^2}(x)$ and risk aversion function $\lambda(x)$ as positive measurable functions on

$(0, \infty]$ and hence subjective variance $\tilde{\sigma}^2(x) \equiv \sigma^2(1 - IC^2(x))$ will also be a measurable function. We now define the weighted harmonic mean risk-aversion variance product as

$$\bar{\sigma}^2 \equiv \left(\int_0^\infty f(x) \frac{1}{\tilde{\sigma}^2(x)} dx \right)^{-1} > 0 \quad (37)$$

and hence the δ density is obtained from:

$$\delta(x) = f(x) \frac{\bar{\sigma}^2}{\tilde{\sigma}^2(x)} \quad (38)$$

where in fact $\frac{\bar{\sigma}^2}{\tilde{\sigma}^2(x)}$ is the Radon-Nikodym derivative of the ‘market-presence’ measure with respect to the population headcount measure F . We denote by Δ the ‘market-presence’ measure itself, i.e.

$$\Delta(A) \equiv \int_A \delta dm \quad (39)$$

where m denotes Lebesgue measure.

We can now proceed to calculate integrals under the ‘market-presence’ measure. Formally we shall use Lebesgue integrals although for the structures which we consider they will be computationally equivalent to Riemann integrals. For example the mean value of IC^2 under the ‘market presence’ measure (with density δ) will be

$$\mathbb{E}_\delta [IC^2] = \int_0^\infty IC^2(x) \delta(x) dx \quad (40)$$

From one definition of the Lebesgue integral (see Billingsley (1995)) this can be thought of as the following limit:

$$\int_0^\infty IC^2(x) \delta(x) dx = \lim_n \sum_{|l| \leq n2^n} \frac{l-1}{2^n} \Delta \left[x : \frac{l-1}{2^n} \leq IC^2(x) < \frac{l}{2^n} \right] \quad (41)$$

Very informally one can think of this as calculating the expected values of various quantities by partitioning the population into increasingly finely-grained subsets. We group all investors in each subset together and ‘label’ them with the IC^2 value of the lowest member. For a given level of granularity (represented by n) we compute the sum of these ‘labels’ weighted by the size of the subset according to the market presence measure Δ . The value of the integral is then the limit as the size of subsets is reduced.

Finally we note that $\Delta(x) = 0$ when $x \in (0, \infty]$, or equivalently the set of investors such that x is precisely equal to a given real number has measure zero.

A.2 Derivation of future wealth density

We derive the density of wealth function by application of standard change of variable methods such as described by Billingsley (1995).

Recall that

$$W^{(i)} = YH^{(i)} = Y \frac{A^{(i)}}{\lambda \sigma^{(i)2}}$$

where we note that Y and $H^{(i)}$ are jointly normal with correlation coefficient RIC which we denote by ρ . We first consider the joint density of $(W^{(i)}, Y)$. Invoking the transformation method we have that $(H^{(i)}, Y)$ is bivariate normal with density denoted by $f(H^{(i)}, Y)$. We define the one-to-one transformation $\xi_1 = W^{(i)} = YH^{(i)}$ and $\xi_2 = Y$ and have that the joint density of ξ_1 and ξ_2 is

$$g(\xi_1, \xi_2) = f\left(\frac{\xi_1}{\xi_2}, \xi_2\right) |\mathcal{J}|$$

where

$$\begin{aligned} |\mathcal{J}| &= \det \begin{bmatrix} \frac{\partial H^{(i)}}{\partial \xi_1} & \frac{\partial H^{(i)}}{\partial \xi_2} \\ \frac{\partial Y}{\partial \xi_1} & \frac{\partial Y}{\partial \xi_2} \end{bmatrix} \\ &= \det \begin{bmatrix} \frac{1}{Y} & -\frac{H^{(i)}}{Y} \\ 0 & 1 \end{bmatrix} \\ &= \frac{1}{Y} \end{aligned}$$

is the Jacobian of the transformation. Hence we have

$$\begin{aligned} g(\xi_1, \xi_2) &= f\left(\frac{W^{(i)}}{Y}, Y\right) \frac{1}{Y} \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_{H^{(i)}}\sigma_Y} \frac{1}{Y} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{W^{(i)2}}{\sigma_{H^{(i)}}^2 Y^2} + \frac{Y^2}{\sigma_Y^2} - \frac{2\rho W^{(i)}}{\sigma_{H^{(i)}}\sigma_Y}\right)\right] \\ &= \frac{e^{\frac{\rho W^{(i)}}{(1-\rho^2)\sigma_{H^{(i)}}\sigma_Y}}}{2\pi\sqrt{1-\rho^2}\sigma_{H^{(i)}}\sigma_Y} \frac{1}{Y} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{W^{(i)2}}{\sigma_{H^{(i)}}^2 Y^2} + \frac{Y^2}{\sigma_Y^2}\right)\right] \end{aligned}$$

Now we require the marginal density of $W^{(i)}$ which we obtain by integration over Y as follows:

$$g(W^{(i)}) = \frac{e^{\frac{\rho W^{(i)}}{(1-\rho^2)\sigma_{H^{(i)}}\sigma_Y}}}{2\pi\sqrt{1-\rho^2}\sigma_{H^{(i)}}\sigma_Y} \int_{-\infty}^{\infty} \frac{1}{Y} \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{W^{(i)2}}{\sigma_{H^{(i)}}^2 Y^2} + \frac{Y^2}{\sigma_Y^2}\right)\right] dY$$

we achieve this by the substitution

$$\tau = \frac{Y^2}{2\sigma_Y^2(1-\rho^2)}$$

so we have

$$\begin{aligned} Y &= \sqrt{2\tau}\sigma_Y\sqrt{1-\rho^2} \\ \frac{dY}{d\tau} &= \frac{1}{\sqrt{2\tau}}\sigma_Y\sqrt{1-\rho^2} \end{aligned}$$

hence

$$\begin{aligned}
g(W^{1(i)}) &= \frac{e^{\frac{\rho W^{(i)}}{(1-\rho^2)\sigma_{H^{(i)}}\sigma_Y}}}{2\pi\sqrt{1-\rho^2}\sigma_{H^{(i)}}\sigma_Y} \int_{-\infty}^{\infty} \frac{1}{2\tau} \exp \left[- \left(\frac{W^{(i)2}}{4\sigma_{H^{(i)}}^2\sigma_Y^2(1-\rho^2)^2\tau} + \tau \right) \right] d\tau \\
&= \frac{e^{\frac{\rho W^{(i)}}{(1-\rho^2)\sigma_{H^{(i)}}\sigma_Y}}}{2\pi\sqrt{1-\rho^2}\sigma_{H^{(i)}}\sigma_Y} \int_0^{\infty} \frac{1}{\tau} \exp \left[- \left(\frac{W^{(i)2}}{4\sigma_{H^{(i)}}^2\sigma_Y^2(1-\rho^2)^2\tau} + \tau \right) \right] d\tau \\
&= \frac{e^{\frac{\rho W^{(i)}}{(1-\rho^2)\sigma_{H^{(i)}}\sigma_Y}}}{\pi\sqrt{1-\rho^2}\sigma_{H^{(i)}}\sigma_Y} K_0 \left(\frac{W^{(i)}}{\sigma_{H^{(i)}}\sigma_Y(1-\rho^2)} \right)
\end{aligned}$$

where $K_0(\cdot)$ is the modified Bessel function of the third kind with imaginary argument and the last step follows from

$$K_v(z) = \frac{1}{2} \int_{-\infty}^{\infty} \exp(-z \cosh t - vt) dt$$

if we make the substitution $\frac{1}{2}ze^t = \tau$ to obtain

$$K_v(z) = \frac{1}{2} \left(\frac{1}{2}z \right)^v \int_0^{\infty} \exp \left(-\tau - \frac{z^2}{4\tau} \right) \frac{d\tau}{\tau^{v+1}}$$

as per Watson (1944, page 182-3).

A.3 Proofs of Section 3 Propositions

A.3.1 Proposition 3.1

Proof. (a) The proof of $f^{(i)} \geq 0 \forall i \Rightarrow \delta^{(i)} \geq 0 \forall i$ follows directly from (4) and (5). (b) To prove $\delta^{(i)} \geq 0 \forall i \Rightarrow f^{(i)} \geq 0 \forall i$ we proceed by contradiction: suppose $\delta^{(i)} \geq 0 \forall i$ but $\exists j : f^{(j)} < 0 \Rightarrow \bar{\sigma}^2 < 0 \Rightarrow f^{(i)} \leq 0 \forall i$ (since $\delta^{(i)} \geq 0 \forall i \Rightarrow \sum_{i=0}^{J-1} f^{(i)} < 0$ which contradicts the requirement

$$\sum_{i=0}^{J-1} f^{(i)} = 1. \quad \square$$

A.3.2 Realised Information Coefficient formula

Proof.

$$\begin{aligned}
RIC^{(i)} &\equiv \frac{\sigma_{A^{(i)}Y}}{\sigma_{A^{(i)}}\sigma_Y} \\
&= \frac{Cov(p_{t+1} - p_t, IC^{(i)}\sigma S^{(i)} - p_t)}{\sigma_{A^{(i)}}\sigma_Y} \\
&= \frac{Cov(p_{t+1}, IC^{(i)}\sigma S^{(i)}) - Cov(p_{t+1}, p_t) - Cov(p_t, IC^{(i)}\sigma S^{(i)}) + V\sigma^2}{\sigma_{A^{(i)}}\sigma_Y} \\
&= \frac{IC^{(i)2}\sigma^2 - \sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)2}\sigma^2 - Cov\left(\sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)}\sigma S^{(j)}, IC^{(i)}\sigma S^{(i)}\right) + V\sigma^2}{\sigma_{A^{(i)}}\sigma_Y} \\
&= \frac{IC^{(i)2} - \sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)2} - \sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)} IC^{(i)} \phi_{ij} + V}{\sqrt{\left[1 + V - 2 \sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)2}\right] \left[IC^{(i)2} + V - 2 \sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)} IC^{(i)} \phi_{ij}\right]}} \tag{42}
\end{aligned}$$

□

A.3.3 Proposition 3.3

Proof.

$$\begin{aligned}
RIC^{(i)} &= \frac{IC^{(i)2} - \sum_{j=0}^{J-1} \delta^{(j)} IC^{(j)2} - \sum_{j=0}^{J-1} \delta^{(j)} IC^{(i)} IC^{(j)} \phi_{ij} + V}{\sigma_Y \sigma_{A^{(i)}}} \\
&= \frac{IC^2 - J\delta IC^2 - (J-1)\delta IC^2 \sum_{j=0}^{J-1} \phi_{ij} - \delta IC^2 + J(J-1)\delta^2 IC^2 \sum_{j=0}^{J-1} \phi_{ij} + J\delta^2 IC^2}{\sigma_Y \sigma_{A^{(i)}}} \\
&= \frac{IC^2 - IC^2 - \frac{(J-1)}{J} IC^2 \sum_{j=0}^{J-1} \phi_{ij} - \frac{IC^2}{J} + \frac{(J-1)}{J} IC^2 \sum_{j=0}^{J-1} \phi_{ij} + \frac{IC^2}{J}}{\sigma_Y \sigma_{A^{(i)}}} \\
&= 0
\end{aligned}$$

since $\delta^{(i)} = \delta = \frac{1}{J}$ due to the assumption of equiprequent types with equal constant absolute risk aversion. □

A.4 Proofs of Section 4 Propositions

A.4.1 Proposition 4.1

Proof. RIC Approximation: we know that $V = \delta^{(skilled)^2} IC^{(skilled)^2}$, so we have

$$\begin{aligned} RIC &= \frac{\sqrt{IC^2 - 2\delta^{(skilled)} IC^2 + \delta^{(skilled)^2} IC^2}}{\sqrt{1 + \delta^{(skilled)^2} IC^2 - 2\delta^{(skilled)} IC^2}} \\ &= IC^{(skilled)} \frac{(1 - \delta^{(skilled)})}{\sqrt{1 + IC^2 \delta^{(skilled)} [\delta^{(skilled)} - 2]}} \end{aligned}$$

Furthermore we have

$$\delta = \frac{f}{1 - IC^2(1 - f)} \quad (43)$$

Now we proceed by first order Taylor approximation around $f = 0$, differentiating with respect to f . This is facilitated by noting that $\frac{\partial RIC}{\partial f} = \frac{\partial RIC}{\partial \delta} \frac{\partial \delta}{\partial f}$ and

$$\begin{aligned} \frac{\partial \delta}{\partial f} &= \frac{\partial}{\partial f} \left[\frac{f}{1 - IC^2(1 - f)} \right] \\ &= \frac{(1 - IC^2(1 - f)) - f IC^2}{(1 - IC^2(1 - f))^2} \\ &= \frac{1}{1 - IC^2} \text{ if } f = 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial RIC}{\partial f} &= IC \frac{-\frac{\partial \delta}{\partial f} (1 + IC^2 \delta (\delta - 2))^{\frac{1}{2}} - (1 - \delta)^{\frac{1}{2}} (1 + IC^2 \delta (\delta - 2))^{-\frac{1}{2}} (2IC^2 \delta - 2IC^2) \frac{\partial \delta}{\partial f}}{1 + IC^2 \delta (\delta - 2)} \\ &= IC \left[-\frac{\partial \delta}{\partial f} + IC^2 \frac{\partial \delta}{\partial f} \right] \text{ if } f = 0 \\ &= IC \left[\frac{IC^2 - 1}{1 - IC^2} \right] \\ &= -IC \end{aligned}$$

Hence

$$\begin{aligned} RIC &\approx RIC(f = 0) - IC f \\ &= IC(1 - f) \end{aligned}$$

□

A.4.2 Proposition 4.2

Proof. First we have some intermediate results. Since we have only one skilled type we use IC and δ to represent skill level and market presence of that type in the derivation below; clearly $IC = 0$ for the unskilled type and their market presence is given by $1 - \delta$.

Initially we recall (43) and hence

$$\frac{\partial \delta}{\partial IC} = \frac{2IC}{(1-IC^2)}\delta(1-\delta) > 0$$

We now proceed to compute

$$\begin{aligned}\sigma_A^2 &= \sigma^2 IC^2 (1-\delta)^2 \\ \frac{\partial \sigma_A^2}{\partial IC} &= \sigma^2 \left[2IC(1-\delta)^2 - 2IC^2(1-\delta) \frac{\partial \delta}{\partial IC} \right] \\ &= \sigma^2 \left[2IC(1-\delta)^2 - \frac{4IC^3}{1-IC^2} \delta(1-\delta)^2 \right]\end{aligned}$$

and now the key argument:

$$\begin{aligned}\frac{\partial [RIC^2]}{\partial IC} &= 0 \\ \frac{\partial}{\partial IC} \left[\frac{\sigma_A^2}{\sigma_Y^2} \right] &= 0 \\ \frac{\partial}{\partial IC} \left[\frac{\sigma_A^2}{\sigma_A^2 - \sigma^2 (IC^2 - 1)} \right] &= 0 \\ \frac{[\sigma_A^2 - \sigma^2 (IC^2 - 1)] \frac{\partial \sigma_A^2}{\partial IC} - \sigma_A^2 \left[\frac{\partial \sigma_A^2}{\partial IC} - 2IC\sigma^2 \right]}{[\sigma_A^2 - \sigma^2 (IC^2 - 1)]^2} &= 0 \\ \frac{\partial \sigma_A^2}{\partial IC} \left\{ \frac{\sigma^2 [1-IC^2]}{[\sigma_A^2 - \sigma^2 (IC^2 - 1)]^2} \right\} + \frac{2IC\sigma^2\sigma_A^2}{[\sigma_A^2 - \sigma^2 (IC^2 - 1)]^2} &= 0 \\ \frac{\partial \sigma_A^2}{\partial IC} &= -2 \frac{IC}{1-IC^2} \sigma_A^2 \\ 2IC(1-\delta)^2 - \frac{4IC^3}{1-IC^2} \delta(1-\delta)^2 &= -2 \frac{IC^3}{1-IC^2} (1-\delta)^2 \\ (1-IC^2) - 2IC^2\delta &= -IC^2 \\ IC^2\delta &= \frac{1}{2} \\ \frac{fIC^2}{1-IC^2(1-f)} &= \frac{1}{2} \\ 1-IC^2(1-f) &= 2IC^2f \\ IC^2 &= \frac{1}{1+f}\end{aligned}$$

At this point we have

$$\begin{aligned}\delta &= \frac{1+f}{2} \\ 1-\delta &= \frac{1-f}{2} \\ \frac{1-IC^2}{IC^2} &= \frac{1-\frac{1}{1+f}}{\frac{1}{1+f}} = f\end{aligned}$$

$$\begin{aligned}
RIC_{max} &= \sqrt{\frac{\sigma_A^2}{\sigma_Y^2}} \\
&= \sqrt{\frac{IC^2(1-\delta)^2}{IC^2(1-\delta)^2 - IC^2 + 1}} \\
&= \sqrt{\frac{(1-\delta)^2}{(1-\delta)^2 + \frac{(1-IC^2)}{IC^2}}} \\
&= \sqrt{\frac{\left(\frac{1-f}{2}\right)^2}{\left(\frac{1-f}{2}\right)^2 + f}} \\
&= \frac{1-f}{1+f}
\end{aligned}$$

Hence the turning point in RIC will come where $IC = \frac{1}{\sqrt{1+f}}$ at which point RIC will be at its maximum value of $\frac{1-f}{1+f}$. \square

A.4.3 Proposition 4.3

Proof. We denote the unskilled and skilled types by the superscripts (0) and (1) respectively (and hence $IC^{(0)} = 0$).

$$\begin{aligned}
RIC^{(0)} &= \frac{IC^{(0)2} - \sum_{i=0}^{J-1} \delta^{(i)} IC^{(i)2} - \sum_{i=0}^{J-1} \delta^{(i)} IC^{(i)} IC^{(0)} \phi_{0,1} + V}{\sqrt{\left[1 + V - 2 \sum_{i=0}^{J-1} \delta^{(i)} IC^{(i)2}\right] \left[IC^{(0)2} + V - 2 \sum_{i=0}^{J-1} \delta^{(i)} IC^{(i)} IC^{(0)} \phi_{0,1}\right]}} \\
&= \frac{-\delta^{(1)} IC^{(1)2} + \delta^{(1)2} IC^{(1)2}}{\sqrt{\left[1 + \delta^{(1)2} IC^{(1)2} - 2\delta^{(1)} IC^{(1)2}\right] \left[\delta^{(1)2} IC^{(1)2}\right]}} \\
&= -IC^{(1)} \frac{(1 - \delta^{(1)})}{\sqrt{1 + IC^{(1)2} \delta^{(1)} [\delta^{(1)} - 2]}} \\
&= -RIC^{(1)}
\end{aligned} \tag{44}$$

\square

A.4.4 Proposition 4.4

Proof. We require a simple functional form which we can impose on ϕ_{ij} such that we ensure positive definiteness of the overall matrix of forecasts and prices $(p_{t+1}, S_t^{(i)})_{i \in [0, J-1]}$ which we denote by $\mathbf{\Gamma}$.

$$\mathbf{\Gamma} = \begin{bmatrix} \sigma^2 & IC^{(0)}\sigma & IC^{(1)}\sigma & \dots & IC^{(J-1)}\sigma \\ IC^{(0)}\sigma & 1 & \phi_{0,1} & \dots & \phi_{0,J-1} \\ IC^{(1)}\sigma & \phi_{1,0} & 1 & \dots & \phi_{1,J-1} \\ \dots & \dots & \dots & \dots & \dots \\ IC^{(J-1)}\sigma & \phi_{J-1,0} & \phi_{J-1,1} & \dots & 1 \end{bmatrix} \tag{45}$$

From Kreindler and Jameson (1972) the condition for positive definiteness can equivalently be stated as requiring positive definiteness of:

$$\begin{bmatrix} 1 & \phi_{0,1} & \dots & \phi_{0,J-1} \\ \phi_{J-1,0} & 1 & \dots & \phi_{1,J-1} \\ \dots & \dots & \dots & \dots \\ \phi_{J-1,0} & \phi_{J-1,1} & \dots & 1 \end{bmatrix} - \begin{bmatrix} IC^{(0)2} & IC^{(0)}IC^{(1)} & \dots & IC^{(0)}IC^{(J-1)} \\ IC^{(1)}IC^{(0)} & IC^{(1)2} & \dots & IC^{(1)}IC^{(J-1)} \\ \dots & \dots & \dots & \dots \\ IC^{(J-1)}IC^{(0)} & IC^{(J-1)}IC^{(1)} & \dots & IC^{(J-1)2} \end{bmatrix} \quad (46)$$

$$= \begin{bmatrix} 1 - IC^{(0)2} & \phi_{0,1} - IC^{(0)}IC^{(1)} & \dots & \phi_{0,J-1} - IC^{(0)}IC^{(J-1)} \\ \phi_{1,0} - IC^{(1)}IC^{(0)} & 1 - IC^{(1)2} & \dots & \phi_{1,J-1} - IC^{(1)}IC^{(J-1)} \\ \dots & \dots & \dots & \dots \\ \phi_{J-1,0} - IC^{(J-1)}IC^{(0)} & \phi_{J-1,1} - IC^{(J-1)}IC^{(1)} & \dots & 1 - IC^{(J-1)2} \end{bmatrix} \quad (47)$$

(Alternatively this condition can be thought of as checking positive definiteness in the conditional covariance matrix of individual scores given the actual asset price, which is what the above matrix is).

Now suppose we set ϕ as follows:

$$\phi_{ij} = kIC^{(i)}IC^{(j)} \text{ where } i \neq j \quad (48)$$

$$\phi_{ij} = 1 \text{ where } i = j \quad (49)$$

Then we require positive definiteness of

$$\begin{bmatrix} 1 - IC^{(0)2} & (k-1)IC^{(0)}IC^{(1)} & \dots & (k-1)IC^{(0)}IC^{(J-1)} \\ (k-1)IC^{(1)}IC^{(0)} & 1 - IC^{(1)2} & \dots & (k-1)IC^{(1)}IC^{(J-1)} \\ \dots & \dots & \dots & \dots \\ (k-1)IC^{(J-1)}IC^{(0)} & (k-1)IC^{(J-1)}IC^{(1)} & \dots & 1 - IC^{(J-1)2} \end{bmatrix} \quad (50)$$

Which can be written as an outer product plus a diagonal matrix as follows:

$$(k-1) \begin{bmatrix} IC^{(0)} \\ IC^{(1)} \\ \dots \\ IC^{(J-1)} \end{bmatrix} \begin{bmatrix} IC^{(0)} & IC^{(1)} & \dots & IC^{(J-1)} \end{bmatrix} + \begin{bmatrix} 1 - kIC^{(0)2} & 0 & \dots & 0 \\ 0 & 1 - kIC^{(1)2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 - kIC^{(J-1)2} \end{bmatrix}$$

A sufficient condition for positive definiteness is therefore:

$$1 < k < \frac{1}{IC_{max}^2} \quad (51)$$

where IC_{max}^2 denotes the largest IC value in the population. \square

A.4.5 Proposition 4.5

We recall the definition of δ for the discrete case from (4) and (5) and use the following approach for a continuum of types (see A.1.2 for further clarification):

Proof.

$$\delta(\pi) = \frac{\pi}{\mathbb{E}_f[\pi]} f(\pi) \quad (52)$$

$$= \frac{s-1}{s} \frac{1 - \pi_{max}^{-s}}{1 - \pi_{max}^{1-s}} \frac{s}{1 - \pi_{max}^{-s}} \pi^{-s} \quad (53)$$

$$= \frac{(s-1)}{1 - \pi_{max}^{1-s}} \pi^{-s} ; 1 \leq \pi < \pi_{max} \quad (54)$$

□

A.4.6 Proposition 4.6

Proof.

$$\begin{aligned} \mathbb{E}_\delta [IC^2] &= \int_1^{\pi_{max}} \left(1 - \frac{1}{\pi}\right) \delta(\pi) d\pi \\ &= \frac{(s-1)}{1 - \pi_{max}^{1-s}} \int_1^{\pi_{max}} (\pi^{-s} - \pi^{-s-1}) d\pi \\ &= \frac{(s-1)}{1 - \pi_{max}^{1-s}} \left[\frac{1}{1-s} \pi^{1-s} + \frac{1}{s} \pi^{-s} \right]_1^{\pi_{max}} \\ &= \frac{(s-1)}{1 - \pi_{max}^{1-s}} \left[\frac{1}{s-1} - \frac{1}{s} + \frac{1}{1-s} \pi_{max}^{1-s} + \frac{1}{s} \pi_{max}^{-s} \right] \\ &= \frac{1}{1 - \pi_{max}^{1-s}} \left[\frac{1}{s} - \pi_{max}^{1-s} + \frac{s-1}{s} \pi_{max}^{-s} \right] \\ &= \frac{1}{1 - \pi_{max}^{1-s}} \left[\frac{1}{s} (1 - \pi_{max}^{-s}) - \pi_{max}^{-s} (\pi_{max} - 1) \right] \\ &\rightarrow \frac{1}{s} \text{ as } \pi_{max} \rightarrow \infty \end{aligned}$$

□

A.4.7 Variance Components

In the discrete types case we have to evaluate summations of the form

$$\sum_{j=0}^{J-1} \delta^{(i)} IC^{(i)} IC^{(j)} \phi_{ij}$$

where $IC^{(i)}$ is fixed. If we impose the correlation structure $\phi_{ij} = kIC^{(i)}IC^{(j)}$ then the equivalent continuum form of the summation is

$$\begin{aligned} & \mathbb{E}_\delta \left[IC^{(i)} IC \phi_{ij} \right] \\ &= kIC^{(i)2} \mathbb{E}_\delta \left[IC^2 \right] \end{aligned}$$

and also, by a similar method, the continuum version of V (the Variance Factor of equilibrium prices) becomes

$$\begin{aligned} V &= \int_1^{\pi_{max}} \int_1^{\pi_{max}} IC(\pi_1) IC(\pi_2) \phi(\pi_1, \pi_2) \delta(\pi_1) \delta(\pi_2) d\pi_1 d\pi_2 \\ &= k \int_1^{\pi_{max}} \int_1^{\pi_{max}} \left(1 - \frac{1}{\pi_1} \right) \left(1 - \frac{1}{\pi_2} \right) \delta(\pi_1) \delta(\pi_2) d\pi_1 d\pi_2 \\ &= k \left(\mathbb{E}_\delta \left[IC^2 \right] \right)^2 \end{aligned}$$

where we use π_1 and π_2 to denote the two identical ranges of precision over which we double integrate. As explained in A.1.2 the set of investors where $\pi_1 = \pi_2$ has measure zero. Hence in the continuum we always integrate over cases analogous to those in the discrete case where we have $i \neq j$ and $\phi_{ij} = kIC^{(i)}IC^{(j)}$; this results in the simple expressions above.

A.4.8 Corollary 4.9

Proof.

$$\begin{aligned} RIC &= \frac{\left[IC^2 - \overline{IC^2} \right] \left[1 - k\overline{IC^2} \right]}{\sqrt{\left[1 - \overline{IC^2} \right] \left(2 - k\overline{IC^2} \right)} \left[IC^2 \left(1 - 2k\overline{IC^2} \right) + k\overline{IC^2} IC^2 \right]} \\ &= \frac{IC^2 - \overline{IC^2}}{\sqrt{IC^2 + k \frac{\overline{IC^2} IC^2}{1 - 2k\overline{IC^2}}}} \frac{\left[1 - k\overline{IC^2} \right]}{\sqrt{\left[1 - \overline{IC^2} \right] \left(2 - k\overline{IC^2} \right)} \left(1 - 2k\overline{IC^2} \right)} \\ \theta RIC &= \frac{IC^2 - \overline{IC^2}}{\sqrt{IC^2 + k \frac{\overline{IC^2} IC^2}{1 - 2k\overline{IC^2}}}} \\ \theta^2 RIC^2 &= \frac{IC^4 - 2\overline{IC^2} IC^2 + \left(\overline{IC^2} \right)^2}{IC^2 + k \frac{\overline{IC^2} IC^2}{1 - 2k\overline{IC^2}}} \end{aligned}$$

where

$$\theta^2 = \frac{\left[1 - \overline{IC^2} \right] \left(2 - k\overline{IC^2} \right) \left(1 - 2k\overline{IC^2} \right)}{\left[1 - k\overline{IC^2} \right]^2}$$

Now

$$IC^4 - IC^2 \left(2\overline{IC^2} + \theta^2 RIC^2 \right) + \left(\overline{IC^2} \right)^2 - k \left(\frac{\overline{IC^2} \overline{IC^2}}{1 - 2k\overline{IC^2}} \right) \theta^2 RIC^2 = 0$$

We now apply the quadratic formula to determine IC^2 as a function of RIC . We must be careful here to distinguish between two cases: $RIC \leq 0$ and $RIC > 0$. These correspond, respectively, to the cases $IC^2 \leq \overline{IC^2}$ and $IC^2 > \overline{IC^2}$ and the appropriate IC^2 functions are given for each case by choosing the appropriate sign in the quadratic formula (i.e. selecting the appropriate root). By judicious rearrangement, however, we can exploit the fact that the relevant sign is also the sign of RIC itself, and so IC^2 can be stated in a convenient format as shown below:

$$IC^2 = \frac{2\overline{IC^2} + \theta^2 RIC^2 \pm \sqrt{4\overline{IC^2} \theta^2 RIC^2 + \theta^4 RIC^4 + 4k \left(\frac{\overline{IC^2} \overline{IC^2}}{1 - 2k\overline{IC^2}} \right) \theta^2 RIC^2}}{2}$$

$$IC^2 = \overline{IC^2} + \frac{1}{2} \theta^2 RIC^2 + \theta RIC \sqrt{\frac{1}{4} \theta^2 RIC^2 + \overline{IC^2} \left(\frac{1 - k\overline{IC^2}}{1 - 2k\overline{IC^2}} \right)}$$

□

A.4.9 Proposition 4.10

Proof. We apply the standard change of variable method as described by Billingsley (1995).

$$\begin{aligned} f_{IC^2}(IC^2) &= f(\pi) \frac{\partial}{\partial IC^2} \left[\frac{1}{1 - IC^2} \right] \\ &= \frac{s}{1 - \pi_{max}^{-s}} \left(\frac{1}{1 - IC^2} \right)^{-(s+1)} \left[\frac{1}{1 - IC^2} \right]^2 \\ &= \frac{s}{1 - \pi_{max}^{-s}} \left(\frac{1}{1 - IC^2} \right)^{-s+1} ; 0 \leq IC^2 < IC_{max}^2 \end{aligned}$$

□

A.4.10 Proposition 4.11

Proof. We follow the same method as in Proposition 4.10 above. For this we require IC^2 as a function of RIC , which is provided by Corollary 4.9, and also the derivative of this function with respect to RIC which we obtain as follows:

$$\frac{\partial IC^2(RIC)}{\partial RIC} = \theta^2 RIC + \theta \sqrt{\frac{1}{4} \theta^2 RIC^2 + \overline{IC^2} \left(\frac{1 - k\overline{IC^2}}{1 - 2k\overline{IC^2}} \right)} + \frac{\frac{1}{4} \theta^3 RIC^2}{\sqrt{\frac{1}{4} \theta^2 RIC^2 + \overline{IC^2} \left(\frac{1 - k\overline{IC^2}}{1 - 2k\overline{IC^2}} \right)}}$$

It is now straightforward to compute the density as the product:

$$f_{RIC}(RIC) = f_{IC^2} [IC^2(RIC)] \frac{\partial IC^2(RIC)}{\partial RIC}$$

□

A.4.11 Proposition 4.12

Proof.

$$\begin{aligned} f_{IC}(IC) &= f_{IC^2}(IC^2) \frac{\partial}{\partial IC} IC^2 \\ &= \frac{2s}{1 - \pi_{max}^{-s}} \left(\frac{1}{1 - IC^2} \right)^{-s+1} IC ; 0 \leq IC < IC_{max} \end{aligned}$$

□

A.4.12 Proposition 4.13

Proof.

$$f_{IC}(IC) = \frac{2s}{1 - \pi_{max}^{-s}} \left(\frac{1}{1 - IC^2} \right)^{-s+1} IC ; 0 \leq IC < IC_{max}$$

For the mode we require:

$$\begin{aligned} \frac{\partial f_{IC}(IC)}{\partial IC} &= 0 \\ \frac{2s}{1 - \pi_{max}^{-s}} \left(\frac{1}{1 - IC_{mode}^2} \right)^{-s+1} + \frac{2s}{1 - \pi_{max}^{-s}} IC_{mode}(s-1) \left(\frac{1}{1 - IC_{mode}^2} \right)^{2-s} (-2IC_{mode}) &= 0 \\ 1 - 2IC_{mode}^2(s-1) \left(\frac{1}{1 - IC_{mode}^2} \right) &= 0 \end{aligned}$$

$$\begin{aligned} 1 - IC_{mode}^2 &= 2(s-1)IC_{mode}^2 \\ IC_{mode} &= \frac{1}{\sqrt{2s-1}} \end{aligned}$$

□

References

- AGARWAL, V. AND N. NAIK (2000): “Generalised Style Analysis of Hedge Funds,” *Journal of Asset Management*, 1, 93–109.
- ALLEN, F. AND D. GALE (1988): “Optimal Security Design,” *Review of Financial Studies*, 1, 229–263.
- ALMGREN, R. AND N. CHRISS (2007): “Optimal portfolios from ordering information,” in *Forecasting Expected Returns in the Financial Markets*, ed. by S. Satchell, Academic

- Press, chap. 4, 55–100.
- ASQUITH, P., M. MIKHAIL, AND A. AU (2005): “Information content of equity analyst reports,” *Journal of Financial Economics*, 75, 245–282.
- BILLINGSLEY, P. (1995): *Probability and Measure*, Wiley Inter-Science.
- BLACK, F. AND R. LITTERMAN (1992): “Global Portfolio Optimization,” *Financial Analysts Journal*.
- BOUCHAUD, J.-P., J. FARMER, AND F. LILLO (2009): “How Markets Slowly Digest Changes in Supply and Demand,” in *Handbook of Financial Markets Dynamics and Evolution*, ed. by T. Hens and K. Schenk-Hoppé, North-Holland.
- BRANCH, W. (2002): “Local convergence properties of a cobweb model with rationally heterogeneous expectations,” *Journal of Economic Dynamics and Control*, 27, 63–85.
- BREALEY, R. AND S. MYERS (2002): *Principles of Corporate Finance*, McGraw Hill.
- BROCK, W. AND C. HOMMES (1997): “A Rational Route to Randomness,” *Econometrica*, 65, 1059–1160.
- BROWN, S. AND W. GOETZMANN (1997): “Mutual Fund Styles,” *Journal of Financial Economics*, 43, 373–399.
- (2003): “Hedge Funds with Style,” *Journal of Portfolio Management*, 29, 101–112.
- BURGESS, K. (2006): “FSA clears hedge fund trading systems,” *Financial Times*, 26th September.
- CANTO, V., D. JOINES, AND A. LAFFER (1982): *Foundations of Supply Side Economics - Theory and Evidence*, New York: Academic Press.
- CLARKE, R., H. DE SILVA, AND S. THORLEY (2002): “Portfolio Constraints and the Fundamental Law of Active Management,” *Financial Analysts Journal*, 58, 48–66.
- COGGIN, T. AND J. HUNTER (1983): “Problems in Measuring the Quality of Investment Information: The Perils of the Information Coefficient,” *Financial Analysts Journal*.
- CRAIG, C. (1936): “On the Frequency Function of xy ,” *Annals of Mathematical Statistics*, 7, 1–15.
- DEMANGE, G. AND G. LAROQUE (1995): “Private Information and the Design of Securities,” *Journal of Economic Theory*, 65, 233–257.
- DIMSON, E. AND P. MARSH (1984): “An Analysis of Brokers’ and Analysts’ Unpublished Forecasts of UK Stock Returns,” *Journal of Finance*, 39, 1257–1292.
- FERSON, W. AND R. W. SCHADT (1996): “Measuring Fund Strategy and Performance in Changing Economic Conditions,” *Journal of Finance*, 51, 425–461.
- FUNG, W. AND D. HSIEH (2002): “Asset-based style factors for hedge funds,” *Financial Analysts Journal*, 58, 16–27.
- GOODWIN, T. (1998): “The Information Ratio,” *Financial Analysts Journal*, 54, 34–43.
- GRENE, S. (2008): “Capturing ideas for the good of all,” *Financial Times*, 28th April.

- GRINOLD, R. (1989): “The Fundamental Law of Active Management,” *Journal of Portfolio Management*.
- GRINOLD, R. AND R. KAHN (1999): *Active Portfolio Management*, McGraw Hill, second ed.
- HIRSHLEIFER, D. AND S. TEOH (2009): “Thought and Behavior Contagion in Capital Markets,” in *Handbook of Financial Markets Dynamics and Evolution*, ed. by T. Hens and K. Schenk-Hoppé, North-Holland.
- KOSOWSKI, R., N. NAIK, AND M. TEO (2007): “Do hedge funds deliver alpha? A Bayesian and bootstrap analysis,” *Journal of Financial Economics*, 84, 229–264.
- KOSOWSKI, R., A. TIMMERMANN, R. WERMERS, AND H. WHITE (2006): “Can Mutual Fund ”Stars” Really Pick Stocks? New Evidence from a Bootstrap Analysis,” *Journal of Finance*, 61, 2551–2595.
- KOUTSOYANNIS, C. AND S. SATCHELL (2007): “Robust optimization for utilizing forecasted returns in institutional investment,” in *Forecasting Expected Returns in the Financial Markets*, ed. by S. Satchell, Academic Press, chap. 8, 177–189.
- KREINDLER, E. AND A. JAMESON (1972): “Conditions for Nonnegativeness of Partitioned Matrices,” *IEEE Transactions on Automatic Control*, 17, 147–148.
- KYLE, A. (1985): “Continuous Auctions and Insider Trading,” *Econometrica*, 53, 1315–1335.
- LEE, C.-F. AND S. RAHMAN (1990): “Market timing, selectivity and mutual fund performance: an empirical investigation,” *Journal of Business*, 63, 261–278.
- LEE, W. (2000): *Theory and Methodology of Tactical Asset Allocation*, Wiley.
- LEFEVRE, E. (1923): *Reminiscences of a Stock Operator*, George H. Doran and Company.
- LELAND, H. AND D. PYLE (1977): “Informational Asymmetries, Financial Structure and Financial Intermediation,” *Journal of Finance*, 32, 371–387.
- LITTERMAN, R. (2009): “When and how is it appropriate to use mathematical models in finance?” Speech at Quant Invest 2009 conference, December 1st, Paris.
- LO, A. (2008): *Hedge Funds - An Analytic Perspective*, Princeton University Press.
- MACKINTOSH, J. (2006): “GLG launches its new ‘alpha capture’ fund,” *Financial Times*, 9th December.
- MADAN, D., P. CARR, AND E. CHANG (1998): “The Variance Gamma Process and Option Pricing Model,” *European Finance Review*, 2.
- MADAN, D. AND E. SENETA (1990): “The Variance Gamma (V.G.) Model for Share Market Returns,” *Journal of Business*, 63, 511–524.
- MALKIEL, B. (1995): “Returns from investing in equity mutual funds 1971 to 1991,” *Journal of Finance*, 50, 549–572.

- RAMNATH, S., S. ROCK, AND P. SHANE (2008): “The financial analyst forecasting literature: A taxonomy with suggestions for further research,” *International Journal of Forecasting*, 24, 34–75.
- SATCHELL, S. AND O. WILLIAMS (2010): “Social Welfare Issues of Financial Literacy,” Cambridge Working Papers in Economics.
- SCHERER, B. (2007): *Portfolio Construction and Risk Budgeting*, Risk Books.
- SÉNÉCHAL, E. (2004): “The Challenges of Declining Cross-Sectional Volatility,” *Horizon - The Barra Newsletter*, Autumn, 2–10.
- SHARPE, W. (1981): “Decentralized Investment Management,” *Journal of Finance*, 36.
- (1992): “Asset Allocation: Management Style and Performance Measurement,” *Journal of Portfolio Management*, 18, 7–19.
- TREYNOR, J. AND F. BLACK (1973): “How to Use Security Analysis to Improve Portfolio Selection,” *Journal of Business*, 46, 66–86.
- TREYNOR, J. L. AND K. MAZUY (1966): “Can mutual funds outguess the market?” *Harvard Business Review*, 44, 131–136.
- WATSON, G. (1944): *A Treatise on The Theory of Bessel Functions*, Cambridge University Press, second ed.