Informational Efficiency and Asset Prices in Large Markets*

Georgy Chabakauri

London School of Economics and CEPR

August, 2024

Abstract

We study a noisy general rational expectations equilibrium in an economy with big trading data, populated by asymmetrically informed logarithmic investors. We show that the equilibrium can be either fully or partially revealing about macroeconomic shocks privately observed by informed investors, depending on economic parameters and the extent of endogenous information overlap across data sources. We find that trades in derivative securities reveal substantial information about the shocks; interest rates signal impending economic downturns; and asset prices jump when output volatility transitions across thresholds separating fully and partially revealing equilibria in the parameter space. In contrast to markets with logarithmic investors, in markets with CARA investors derivatives data reveals little information and the quantity of this information does not depend on the data set's dimension.

Journal of Economic Literature Classification Numbers: D53, D82, G12, G14.

Keywords: asymmetric information, big data, derivative securities, heterogeneous beliefs, informational efficiency, interest rates, learning from prices.

*Contact: g.chabakauri@lse.ac.uk. I am grateful to Ulf Axelson, Mike Chernov, Sergei Glebkin, Christian Julliard, Igor Makarov, Ian Martin, and seminar participants at the London School of Economics and FTG summer meeting (Barcelona) for helpful comments. All errors are my own.

1. Introduction

The informational efficiency of financial markets has been at the forefront of economic research ever since the pioneering works of Grossman (1976), Grossman and Stiglitz (1980), Hellwig (1980), Diamond and Verrecchia (1981), Admati (1985), and others. This early literature highlights how the superior knowledge of informed investors about economic fundamentals is incorporated into asset prices, and how uninformed investors can discern this information. In the age of big data, vast datasets generated from trading financial instruments in large markets offer unparalleled opportunities for uncovering information about macroeconomic and financial fundamentals. However, data sources for similar assets such as options with comparable strike prices, underlyings, and maturities, may disclose similar information. Therefore, despite the abundance of data, the amount of unique information revealed in large markets and the contribution of derivative securities to informational efficiency remain unclear.

In this paper, we study the information revealed by big trading data within an economy populated by privately informed and uninformed investors, and biased investors with uninformative (noisy) beliefs about fundamentals, all of whom have logarithmic preferences. We identify economic factors that determine the extent of informational efficiency of financial markets. We show that market data may fully or partially reveal private information about macroeconomic shocks depending on economic parameters such as output volatility and the extent of overconfidence among biased investors. The key to our results is a new methodology that enables the incorporation of investors with logarithmic preferences into a model with asymmetric information. These preferences are more prevalent in macroeconomic literature than constant absolute risk aversion (CARA) preferences favored in asymmetric information research for tractability. We find that models with CARA preferences underestimate the amount of the revealed information. For instance, in cases in which our model yields a fully revealing equilibrium, an otherwise identical model with CARA investors results in a partially revealing equilibrium with a striking property that the total amount of revealed information does not depend on the data set's dimension.

We apply our model to analyze the informational role of derivative securities and interest rates. Our analysis suggests that trades in non-redundant derivatives reveal substantial information, consistent with the empirical evidence (e.g., Easley, O'Hara, and Srinivas (1998); Chakravarty, Gulen, and Mayhew (2004); Pan and Poteshman (2006); Kacperczyk and Pagnotta (2019)) and the intuition that derivatives trading allows investors to express their views on market outlooks. In contrast, models with CARA investors (e.g., Brennan and Cao

(1996); Vanden (2008); among others) imply that derivatives are informationally redundant in the sense that they do not disclose new information beyond what is revealed by prices of underlying assets. Our model also endogenizes interest rates, which are commonly taken as given in the literature on asymmetric information. We show that low interest rates predict low output next period, in line with the evidence that yield curves contain information about impending recessions (e.g., Harvey (1988); Ang, Piazzesi, and Wei (2006)). Moreover, asset prices in the model experience upward or downward jumps as functions of model parameters, such as output volatility, when parameters cross boundaries of regions separating fully and partially revealing equilibria in the parameter space.

Next, we discuss the model and the intuition for the main results. Our analysis is set within a general equilibrium economy featuring two dates and multiple states of the world. There is a representative firm that produces exogenous output at the final date. The state probabilities depend on a random parameter interpreted as a productivity shock. This shock determines the expected output, and is drawn by Nature from a certain prior distribution. Informed investors have direct knowledge of the shock, while uninformed investors discern it from financial data, isolating it from the market noise. The noise is endogenously generated by multiple groups of biased investors with incorrect beliefs about the shock. These beliefs are i.i.d. across the groups of biased investors and are uninformative about the shock. All investors can trade a complete set of Arrow-Debreu securities written on the realizations of states and maximize expected discounted logarithmic utility over date-0 and date-1 consumptions. The limit in which the number of Arrow-Debreu securities approaches infinity is particularly tractable and corresponds to a big data economy.

The prices of Arrow-Debreu securities are non-linear functions of sufficient statistics that summarize the information extractable from prices and trading data. These statistics are given by the weighted average of probabilities of states according to the beliefs of informed and biased investors with weights equal to the masses of the respective investor groups in the economy. The statistics are separable in that the terms involving the productivity shock and noisy beliefs appear as separate additive components. This separability arises due to logarithmic preferences and market completeness, and is key to the model's tractability.

We identify three economic mechanisms that give rise to fully revealing equilibria and identify conditions for partially revealing equilibria. We focus on the case where the probabilities of states are from the exponential family, which means that log-likelihood ratios of informed and biased investors are linear functions of the productivity shock and i.i.d. belief parameters, respectively. This family encompasses normal, log-normal, gamma, binomial, and many other distributions, and is universally adopted in the literature on asymmetric in-

formation (see discussion in Breon-Drish, 2015). First, we show that with these probabilities, the equilibrium is fully revealing when the number of traded securities significantly exceeds the number of biased investors. The intuition is that the sufficient statistics, discussed above, can be viewed as a system of nonlinear equations for finding the productivity shock and the beliefs of biased investors. The number of these equations is one less than the number of traded securities. When the number of biased investors is much smaller than the number of traded securities, this system of equations is overdetermined, which allows identifying the shock and the beliefs of biased investors.

The fully revealing equilibrium, discussed above, breaks down when logarithmic investors are replaced with CARA investors. In the latter case, the sufficient statistics become degenerate in the sense that countless realizations of shocks and noisy beliefs produce identical statistics despite the system of equations being overdetermined. Moreover, the information revealed by these statistics is independent of the data set's dimension, given by the number of traded securities, and can be compressed into just two statistics. This degeneracy stems from the exponential structures of both CARA preferences and exponential family probabilities. Consequently, log-probabilities and wealth endowments act as perfect substitutes in investors' expected utilities, enabling them to offset the effects of log-probabilities on their utility due to market completeness. We show that deviating from either CARA preferences or exponential-family probabilities results in an increased volume of the revealed information.

The second type of fully revealing equilibria arises when the number of traded securities is fixed but the number of biased investors is large. In this case, idiosyncratic noisy beliefs in the sufficient statistics diversify away exposing the productivity shock. We rule out this simple case based on the empirical evidence that forecasting prowess is a rare skill (Tetlock, 2009), and hence, informed investors are likely scarce. Consequently, even when the number of biased investors is large, and hence, the noise is small, the equilibrium is not fully revealing because the signal in prices is weak. We show that when the number of informed investors is proportional to an inverse of the square root of the number of biased investors, the sufficient statistics are transformed into equivalent statistics that are partially revealing and conditionally normal (by the central limit theorem) when the number of biased investors is large. We label this the Gaussian-signal economy.

The equilibrium in the Gaussian-signal economy is partially revealing, holding the number of traded securities fixed. We then explore the big data limit in which the number of securities is very large and demonstrate that in such a limit the economy can be either fully or partially revealing. The intuition is that, on one hand, in large markets with many securities the investors have more sources of information. On the other hand, different securities may

contain similar information, and hence, the total amount of the revealed information is ambiguous. We derive necessary and sufficient conditions for the equilibrium to be fully or partially revealing. We show that the information is fully revealed when the true log-output volatility exceeds the biased investors' belief about this volatility. We interpret this situation as the case in which biased investors are overconfident about their beliefs. In this case, because output volatility is sufficiently small, the trades of Arrow-Debreu securities by informed investors when they observe good and bad productivity shocks are localized around mean log-outputs. Such elevated trading activity around localized segments of the market allows the uninformed investors to discern which shock was observed by the informed investors. The equilibrium is also fully revealing when the volatility of log-output is very large. In this case, informed investors actively trade across the whole spectrum of Arrow-Debreu securities and leave many easily discernible cues about their trades and information. The equilibrium is partially revealing for medium-range output volatilities.

For partially revealing equilibria, we characterize in closed form the measure of informational efficiency that quantifies how close an equilibrium is to being fully revealing. This measure, a multi-dimensional analogue of a signal-to-noise ratio, is infinite when the equilibrium is fully revealing and finite otherwise. It also exhibits a sharp gradient at the boundary between fully and partially revealing equilibria within the parameter space. As a result, asset prices jump, either upward or downward, in response to variations in model parameters.

The paper also sheds light on how interest rates aggregate information. We derive interest rates in closed form and show that they depend on posterior probabilities of productivity shocks embedded into these rates through the trades of uninformed investors. Therefore, interest rates are informative about economic shocks. Furthermore, lower interest rates predict lower output next period. Intuitively, when the informed investors observe a bad productivity shock, the uninformed investors update upward their posterior probability of a bad shock, after observing asset prices. Given the expectations of scant output next period, the prices of Arrow-Debreu securities rise and drag upward the price of a zero coupon bond, replicated by a portfolio of these securities, resulting in low interest rates.

The paper makes several methodological contributions. First, we present a new tractable framework for the analysis of asset prices under asymmetric information with non-CARA preferences. Another contribution is our approach to studying large data sets. We simplify the problem of quantifying informational efficiency in these data sets to solving a Fredholm integral equation of the second kind. This equation is solved in closed form by expanding the solution into a series with respect to the eigenfunctions of the related integral operator. Additionally, when closed-form solutions are not available, we show that finding the measure

of informational efficiency is an ill-posed problem. We address this challenge by finding an approximate measure of efficiency using Tikhonov regularization and showing that this approximate measure is always a low bound for the exact measure of efficiency. Consequently, high approximate measure of efficiency indicates that the equilibrium is near fully revealing.

Related Literature. This paper is related to a large literature on noisy rational expectation equilibrium (REE) models pioneered by Grossman (1976), Grossman and Stiglitz (1980), Hellwig (1980), and Diamond and Verrecchia (1981), among others. Admati (1985) extends these models to multi-asset economies. These early studies introduced models with CARA preferences and normally distributed asset payoffs, which remain prevalent in the literature on asymmetric information due to their tractability (e.g., Vives, 2008; Kurlat and Veldkamp, 2015; Davila and Parlatore, 2021; Glebkin, Gondhi, and Kuong, 2021; among others).

The paper is also related to the literature that relaxes the restrictive assumption of normally distributed asset payoffs while maintaining CARA preferences for investors. Breon-Drish (2015) studies an economy with CARA investors and asset returns drawn from distributions belonging to the exponential family. He establishes the existence and uniqueness of equilibrium and provides a characterization of asset prices in terms of inverse functions and closed-form solutions for cases with binomial distributions. Chabakauri, Yuan, and Zachariadis (2022) explore a multi-asset economy with multinomial distributions, and derive closed-form prices of securities with non-linear payoffs, such as debt, equity, and various derivative securities. They demonstrate that the effective completeness of asset markets makes REE models significantly more tractable, and that incorporating quadratic derivatives into an incomplete market economy renders it effectively complete. They also derive conditions under which derivative securities are informationally (non)redundant.

Pálvölgyi and Venter (2015) prove equilibrium uniqueness in multi-asset settings without normality. Han (2018) and Detemple, Rindisbacher, and Robertson (2020, 2022) study dynamic models in which asset payoffs are not normally distributed. Brennan and Cao (1996), Vanden (2008), Smith (2019), and Huang, Yueshen, and Zhang (2023), among others, study markets with derivative securities, and find that derivatives do not disclose new information which is not already in the prices of underlying assets. We diverge from the above literature in that all investors in our model have logarithmic utility along with non-normal asset payoffs and derivatives trading significantly facilitates informational efficiency.

The work closest to ours is Malamud (2015), which studies a noisy REE with Arrow-Debreu securities in a continuous-space and complete-market economy and demonstrates that equilibria with non-CARA preferences are fully revealing. Our paper differs in several

ways. In terms of model setup, Malamud (2015) studies a partial equilibrium economy with exogenous noisy asset supply and interest rates, whereas we study a general equilibrium economy with endogenous interest rates and noises in asset prices. In terms of economic results, the main differences are as follows. First, the non-CARA equilibria in Malamud (2015) are always fully revealing, while our model generates both fully and partially revealing equilibria. Second, we show that fully revealing equilibria can arise even when the number of traded securities is finite. Third, we examine the effects of information asymmetry on endogenous interest rates. Finally, the economic mechanisms leading to fully revealing equilibria in our model, discussed above, differ from the separation of smooth information from rough noise in Malamud (2015). Consequently, the economic channels in the two papers are complementary. Glebkin, Malamud, and Teguia (2023a,b) relax the assumptions of normality and CARA preferences in markets with strategic traders, albeit without asymmetric information.

The paper contributes to the limited literature on REE models with asymmetric information and non-CARA preferences. Albagli, Hellwig, and Tsyvinski (2023a,b) explore REE models with risk-neutral investors and position limit constraints. Peress (2004) investigates wealth effects in a noisy REE model with non-CARA preferences, employing small risk approximations. Avdis and Glebkin (2023) examine wealth distribution in a model with heterogeneously informed investors and non-CARA preferences. These studies primarily focus on economies with a single risky asset. In contrast, our paper considers an economy with multiple assets that have non-linear payoffs. Additionally, we analyze the effects of asymmetric information on interest rates and establish conditions under which the equilibria are fully or partially revealing. Another related paper is Breugem, Buss, and Peress (2020), which studies endogenous interest rates and their impact on stock market efficiency (as we do in this paper), albeit in a CARA-normal setting.

Radner (1979) shows that REE is generically fully revealing in models where investors receive distinct signals that take finite values. This finiteness is crucial for the results, and an equilibrium may not be fully revealing otherwise (Radner, 1979, p. 677). Our model, in contrast, allows investor beliefs drawn from distributions with continuous support, and hence, our model allows for partially revealing equilibria. Allen (1981, 1982) relaxes the finiteness assumption in Radner (1979) and shows that, under certain conditions, economies with non-revealing equilibria can be approximated (with any desired level of precision) by economies with fully revealing equilibria when the number of sources of uncertainty is smaller than the dimension of the vector of prices. Our paper differs in that we mainly focus on the case in which the latter dimensionality restriction is violated. Additionally, the latter articles do not provide criteria for determining whether a particular economy has a revealing or non-

revealing equilibrium, nor do they offer a general method for constructing fully revealing approximate economies. Consequently, the demonstration of the fully revealing equilibrium in our model does not follow from these articles. Rostek and Weretka (2012) study the effect of market size on information aggregation by prices in a setting where investors have correlated normally distributed asset values and quadratic preferences. In contrast, our model has log preferences, non-normal payoffs, identical asset values across investors, and only informed investors possess information.

The paper is also related to large literature studying the effects of heterogeneous investor beliefs on asset prices (e.g., Atmaz and Basak, 2018; Basak, 2000; Basak, 2005; Brunnermeier, Simsek, and Xiong, 2014; Chabakauri, 2015; Chabakauri and Han, 2020; Cvitanić and Malamud, 2011; Detemple and Murthy, 1997; Martin and Papadimitriou, 2022; among others). Similar to these studies, investors in our model hold different beliefs about output growth. Our contribution to this literature is twofold. First, we relax assumptions that investor beliefs are common knowledge and that they "agree to disagree" with each other. The uninformed investors in our model, in general, do not know the beliefs of other investors, and also recognize the superior information of the informed investors. Second, we provide conditions under which the uninformed investors can recover beliefs not only of the informed but also of biased investors. The latter result effectively endogenizes the assumption that beliefs are common knowledge.

The paper also contributes to the emerging literature on the role of big data in economics and finance (e.g., Begenau, Farboodi, and Veldkamp (2018), Farboodi and Veldkamp (2020)) by exploring the implications of big data for informational efficiency. Our findings introduce a novel insight: the volume of information revealed through big data is influenced by the extent of information overlap across data sources.

2. Model

2.1. Securities Markets and Information Structure

We consider an economy with two dates, $t = 0$ and $t = 1$, and M states $\omega_1, \ldots, \omega_M$ at the terminal date, where $M \geq 2$. The economy is populated by competitive informed and uninformed investors, labeled I and U, respectively, and N different groups of misinformed biased investors who act as noise traders. The masses of informed and uninformed investors in the economy are μ_I and μ_U, respectively, and the mass of each group of biased investors is μ_B/N. The total mass of all investors is normalized to one, $\mu_I + \mu_U + \mu_B = 1$.

There is also one representative firm in the economy, interpreted as Lucas (1978) tree, that produces D_0 units of consumption good at date $t = 0$ and $D_m > 0$ units in state ω_m at the final date, per capita. The investors can trade M Arrow-Debreu securities with payoffs given by indicator functions $1_{\{\omega=\omega_m\}}$, equal to one in state ω_m and zero otherwise. As a result, the financial market is complete. For the simplicity of exposition, instead of introducing claims on the entire Lucas tree, we assume that state-m Arrow-Debreu security is in net supply of D_m units (per capita) for all states ω_m. Consequently, the investors can assemble a claim on the tree by buying a portfolio of Arrow-Debreu securities.

The probabilities of states $\pi_{I,m}(\varepsilon)$ depend on a macroeconomic *productivity shock* ε drawn by Nature before date-0 from a distribution with the probability density function (PDF) $\varphi_\varepsilon(x)$. This parameter may represent a realization of an economic shock that affects the probabilities of states so that when $\varepsilon > 0$ higher output is more likely than when $\varepsilon < 0$. In some applications below, ε determines the mean output growth. The informed investors know shock ε, whereas uninformed investors have a prior PDF $\varphi_\varepsilon(x)$ over shock ε. They update their prior PDF after observing publicly available information, as discussed below, and then trade based on their posterior probabilities $\pi_{U,m}$.

Biased investors are misinformed and hold exogenous heterogenous beliefs about the probabilities of states given by $\pi_{B,m}(\nu_n)$, where belief parameters ν_n are i.i.d. random variables across groups $n = 1, \ldots, N$ of these investors, and are drawn from a distribution with PDF $\varphi_\nu(x)$, which is common knowledge. The probability functions $\pi_{B,1}(z), \ldots, \pi_{B,M}(z)$ are linearly independent.

The beliefs are not directly observable and introduce noise into asset prices, similar to the related literature with CARA investors (e.g., Davila and Parlatore, 2021; Detemple, Rindisbacher, and Robertson, 2020; among others). Biased investors do not seek additional information and "agree to disagree" with other market participants, as in the extensive literature on heterogeneous beliefs (Atmaz and Basak, 2018; Basak, 2000, 2005; Brunnermeier, Simsek, and Xiong, 2014; Chabakauri, 2015; Chabakauri and Han, 2020; Cvitanić and Malamud, 2011; Detemple and Murthy, 1997; Martin and Papadimitriou, 2022). However, we depart from this literature by relaxing common assumptions that investors know beliefs of each other and agree to disagree. In particular, uninformed investors in our model do not observe beliefs of other investors and know that informed investors have superior information.

2.2. Investors' Optimization and Definition of Equilibrium

All investors start with an equal allocation of each Arrow-Debreu security and output D_0 at $t = 0$, and hence, all of them have the same initial wealth W_0. The investors allocate their initial wealth to date-0 consumption C_0 and a portfolio of Arrow-Debreu securities, which determines their state-ω consumption at the final date, $C_1(\omega)$. By $\theta_{i,m}$ we denote the fraction of the initial wealth that type-i investor allocates to security m and by p_m the date-0 price of state-ω_m Arrow-Debreu security. The investors solve the following optimization problem:

$$\max_{C_0, C_1(\omega), \theta_1, \ldots, \theta_M} \ln(C_0) + \beta \mathbb{E}^i[\ln(C_1(\omega))], \tag{1}$$

where $\mathbb{E}^i[\cdot]$ is the expectation under the probability distribution of type-i investor, subject to the self-financing budget constraints

$$W_0 = C_0 + (\theta_{i,1} + \cdots + \theta_{i,M})W_0, \tag{2}$$

$$C_1(\omega) = \frac{\theta_{i,1}W_0}{p_1} 1_{\{\omega=\omega_1\}} + \cdots + \frac{\theta_{i,M}W_0}{p_M} 1_{\{\omega=\omega_M\}}, \tag{3}$$

where $1_{\{\omega=\omega_m\}}$ is the payoff of state-ω_m security, and $\theta_{i,m}W_0/p_m$ is the number of state-ω_m securities in the investor's portfolio, and $i = I, U$ for informed and uninformed investors, respectively, and $i = 1, \ldots, N$ for biased investors.

The expected utility of uninformed investors in (1) is under their posterior probabilities conditional on their information set $\mathcal{F}_U$, and are given by

$$\pi_{U,m} = \mathbb{E}[\pi_{I,m}(\varepsilon)|\mathcal{F}_U]. \tag{4}$$

The information set $\mathcal{F}_U$ is constructed from asset prices and combined asset demands of informed and biased investors. Conditioning on these combined demands, as in the related literature (Breon-Drish, 2010; Banerjee and Green, 2015; Detemple, Rindisbacher, and Robertson, 2020; Kreps, 1977; Malamud, 2015), ensures the existence of REE even when asset prices are not invertible functions of shock ε and beliefs ν.[1] We discuss the role of the information set $\mathcal{F}_U$ in greater detail in Section 3.

[1]Following Malamud (2015), we note that investors can access information about asset demands through various public channels. Limit order books, for instance, provide real-time information about investor demands, while exchanges like the CBOE disclose data on open interest and trading volumes. Additionally, conditional orders enable investors to execute trades contingent on prices and trading volumes. Blume, Easley, and O'Hara (1994) and Schneider (2009) discuss evidence that trading volume offers additional information beyond price and present models in which investors base decisions on trading volume as well as price. Chabakauri, Fos, and Jiang (2023) argue that company insiders can detect activist trading by monitoring order flows of company shares.

Definition of Equilibrium. *A competitive noisy general rational expectations equilibrium is a vector of risky asset prices $P(\varepsilon, \nu)$, posterior probabilities (4) of the uninformed investors, and optimal consumptions $C^*_{0,i}$ and $C^*_{1,i}(\omega)$ and portfolios θ^*_i that solve optimization problem (1) subject to self-financing budget constraints (2) and (3), taking asset prices as given, and satisfy the following market clearing conditions in consumption and asset markets:*

$$\mu_I C^*_{0,I} + \mu_U C^*_{0,U} + \mu_B \left(\frac{C^*_{0,1} + C^*_{0,2} + \cdots + C^*_{0,N}}{N} \right) = D_0, \tag{5}$$

$$\mu_I C^*_{1,I}(\omega_m) + \mu_U C^*_{1,U}(\omega_m) + \mu_B \left(\frac{C^*_{1,1}(\omega_m) + C^*_{1,2}(\omega_m) + \cdots + C^*_{1,N}(\omega_m)}{N} \right) = D_m, \tag{6}$$

$$\mu_I \frac{\theta^*_{I,m} W_0}{p_m} + \mu_U \frac{\theta^*_{U,m} W_0}{p_m} + \frac{\mu_B}{N} \left(\frac{\theta^*_{1,m} W_0}{p_m} + \cdots + \frac{\theta^*_{N,m} W_0}{p_m} \right) = D_m, \tag{7}$$

for all $m = 1, \ldots, M$.

Equation (7) for the market clearing in the asset markets equates the total holdings of state-m Arrow-Debreu security by all investors to the total supply of that security that equals the state-m output D_m, as discussed in Section 2.1. The last terms on the left-hand sides of Equations (6) and (7) represent the random unobserved demands of biased investors for consumption and assets, respectively.

Standard models with CARA investors incorporate exogenous normally distributed price-inelastic noise traders. However, the latter approach is not feasible in an economy with logarithmic investors because consumption and asset markets do not clear when such noise traders submit exogenous demands exceeding supply D_m. This is because, in contrast to CARA investors, logarithmic investors have infinite disutility of zero consumption, and hence, will not match the demand of noise traders even at infinitely high prices. Another challenge with exogenous noise trading in multi-asset markets is that there are many degrees of freedom for determining correlations of noise across markets. This abundance of choices presents difficulty in selecting an economically plausible structure for noise. Introducing endogenous and price-elastic noise trading by biased investors circumvents these problems in our model.

3. Characterization of equilibrium

As the first step of our analysis, we provide general characterization of investors' optimal consumptions, portfolios, and asset prices taking posterior probabilities (4) of the uninformed investors as given and assuming the existence of the equilibrium. As in Chabakauri, Yuan,

and Zachariadis (2022), we observe that the financial market is effectively complete because the investors trade M securities in an economy with M states, taking time-0 asset prices as given. There are no limits to arbitrage, and hence, there exist unique risk-neutral probabilities π_m^{RN} [e.g., Duffie (2001, p. 4)] such that the prices of Arrow-Debreu securities are given by $p_m = \pi_m^{\text{RN}}/(1+r_f)$, where r_f is the endogenous riskless interest rate in the economy. Furthermore, the investors' first order conditions (FOCs) equate their marginal utilities and state price densities $\pi_m^{\text{RN}}/((1+r_f)\pi_{i,m})$, where $\pi_{i,m}$ are physical probabilities from investor i's perspective [e.g., Duffie (2001, p. 5)]. We solve for the optimal consumptions from the FOCs, and then find optimal portfolios θ_i^* from the budget constraint (3), taking risk-neutral probabilities as given. Proposition 1 summarizes the results.

Proposition 1 [Optimal consumptions and portfolios]. *Optimal date-0 consumptions of all investors are given by $C_{0,i} = W_0/(1+\beta)$. Optimal date-1 consumptions of the informed, uninformed, and type-n biased investors in state ω_m are given by*

$$\begin{aligned} C_{1,I}^*(\omega_m) &= \ell W_0 \frac{(1+r_f)\pi_{I,m}(\varepsilon)}{\pi_m^{\text{RN}}}, \\ C_{1,U}^*(\omega_m) &= \ell W_0 \frac{(1+r_f)\pi_{U,m}}{\pi_m^{\text{RN}}}, \\ C_{1,n}^*(\omega_m) &= \ell W_0 \frac{(1+r_f)\pi_{B,m}(\nu_n)}{\pi_m^{\text{RN}}}, \end{aligned} \tag{8}$$

respectively, and their optimal fractions of wealth invested in asset m are given by

$$\theta_{I,m}^* = \ell \pi_{I,m}(\varepsilon), \quad \theta_{U,m}^* = \ell \pi_{U,m}, \quad \theta_{n,m}^* = \ell \pi_{B,m}(\nu_n), \tag{9}$$

where $\ell = \beta/(1+\beta)$, and $n = 1, \ldots, N$ are types of biased investors.

Proposition 1 demonstrates that portfolios (9) are proportional to the investors' physical probabilities and are independent of prices. This independence generates sufficient statistics, employed by uninformed investors to update their beliefs, that are separable in shock ε and beliefs ν, as elaborated below. Such separability is key to the tractability of our model. Portfolios (9) also imply that the asset demands of investors I and B, given by $\theta_{I,m}^* W_0/p_m$ and $\theta_{n,m}^* W_0/p_m$, are downward sloping functions of prices p_m.

Next, we derive the sufficient statistics that summarize the information of the uninformed investors. Portfolios (9) imply that the combined asset demand of informed and biased investors for security m is given by $\ell W_0(\mu_I \pi_{I,m}(\varepsilon) + \mu_B(\pi_{B,m}(\nu_1) + \cdots + \pi_{B,m}(\nu_N))/N)/p_m$. Conditioning on the latter demands and prices p_m reveals the following sufficient statistics:

$$s_m = \mu_I \pi_{I,m}(\varepsilon) + \mu_B \frac{\pi_{B,m}(\nu_1) + \cdots + \pi_{B,m}(\nu_N)}{N}, \tag{10}$$

We observe that these statistics sum up to $\mu_I + \mu_B$ because probabilities sum to one one. Therefore, statistic s_M is redundant and will be omitted in subsequent discussions. We denote the vector of the remaining $M-1$ statistics by $s = (s_1, \ldots, s_{M-1})^\top$. The sufficient statistics (10) are separable in shock ε and noise, given by the second term in (10), which can potentially obfuscate the information about shock ε. The uninformed investors can use these statistics to update their beliefs about ε as $\pi_{m,U} = \mathbb{E}[\pi_{m,I}(\varepsilon)|s]$.

We observe that if asset prices are invertible functions of sufficient statistics, then the information in combined demands is redundant. However, if prices are non-invertible, then the latter information is additionally required to ensure the existence of REE, following the related literature (Breon-Drish, 2010; Banerjee and Green, 2015; Detemple, Rindisbacher, and Robertson, 2020; Kreps, 1977; Malamud, 2015). To show that non-invertible prices are incompatible with REE, substituting portfolios (9) into the market clearing condition (7), we find that

$$\mu_I \pi_{I,m}(\varepsilon) + \mu_B \frac{\pi_{B,m}(\nu_1) + \cdots + \pi_{B,m}(\nu_N)}{N} = \frac{(1+\beta) p_m D_m}{\beta W_0} - \pi_{U,m}(p). \tag{11}$$

The right-hand side of Equation (11) depends on prices and not on shock ε or beliefs ν, and hence, this equation implies that identical prices $p(s) = p(\tilde{s})$ imply identical statistics $s = \tilde{s}$.

Next, we derive the equilibrium. Substituting optimal consumptions and portfolios in Proposition 1 into market clearing conditions, we find the prices of Arrow-Debreu securities, risk-neutral probabilities, and the riskless interest rate. Proposition 2 reports the results.

Proposition 2 [Equilibrium]. *The equilibrium prices of Arrow-Debreu securities, risk-neutral probabilities, and the riskless rate are given by*

$$p_m = \frac{\beta \bar{\pi}_m(s)}{D_m / D_0}, \tag{12}$$

$$\pi_m^{\mathrm{RN}} = \frac{\bar{\pi}_m(s)/D_m}{\sum_{m=1}^M \bar{\pi}_m(s)/D_m}, \tag{13}$$

$$r_f = \frac{1/\beta}{\sum_{m=1}^M \bar{\pi}_m(s)(D_0/D_m)} - 1, \tag{14}$$

where $\bar{\pi}_m(s)$ is the average probability across investors as a function of statistics s, given by

$$\bar{\pi}_m(s) = s_m + \mu_U \mathbb{E}[\pi_{I,m}(\varepsilon)|s], \tag{15}$$

where $s = (s_1, \ldots, s_{M-1})^\top$.

Proposition 2 provides the equilibrium prices in terms of a conditional expectation $\mathbb{E}[\pi_{I,m}(\varepsilon)|s]$. Evaluating this conditional expectation requires knowing the conditional distribution $\varphi_{\varepsilon|s}(x)$ using Bayes's law. The application of Bayes's law in turn requires knowing the joint distribution of noise components in the expressions for the sufficient statistics in (10). This distribution, in general, is not available in the closed form. However, the conditional expectation in (15) can be evaluated numerically as a fitted value from a regression of simulated probabilities $\pi_{m,I}(\varepsilon)$ on multivariate polynomials of simulated signals s_m.

4. Fully revealing equilibria with few biased investors

In this section, we outline conditions for a fully revealing equilibrium in our economy when the number of investor types is low compared to the number of states. We show that the equilibrium in a similar economy but with CARA preferences is not fully revealing. This distinction highlights how models with CARA and non-CARA utilities yield conflicting insights into the informational role of asset prices.

An equilibrium is considered *fully revealing* if for any two vectors of shocks and beliefs $(\varepsilon, \nu_1, \ldots, \nu_N)$ and $(\tilde{\varepsilon}, \tilde{\nu}_1, \ldots, \tilde{\nu}_N)$ that give rise to identical sufficient statistics (10) it follows that $\varepsilon = \tilde{\varepsilon}$. Therefore, the sufficient statistics based on shock ε cannot be replicated by any alternative shock $\tilde{\varepsilon}$. Consequently, the shock can be uniquely identified from the data.

We now provide an intuitive explanation for the emergence of fully revealing equilibria. Suppose, there exist two vectors of shock and noises, $(\varepsilon, \nu_1, \ldots, \nu_N)$ and $(\tilde{\varepsilon}, \tilde{\nu}_1, \ldots, \tilde{\nu}_N)$, that result in identical sufficient statistics (10). Consequently, they satisfy the following system of $M-1$ equations with $2N+2$ unknown variables:

$$\mu_I \pi_{I,m}(\varepsilon) + \mu_B \frac{\pi_{B,m}(\nu_1) + \cdots + \pi_{B,m}(\nu_N)}{N} = \mu_I \pi_{I,m}(\tilde{\varepsilon}) + \mu_B \frac{\pi_{B,m}(\tilde{\nu}_1) + \cdots + \pi_{B,m}(\tilde{\nu}_N)}{N}, \quad (16)$$

$m = 1, \ldots, M-1$. This system is overdetermined when $2N+2 < M-1$, and hence, generically unsolvable unless $(\varepsilon, \nu_1, \ldots, \nu_N) = (\tilde{\varepsilon}, \tilde{\nu}_1, \ldots, \tilde{\nu}_N)$ (up to a permutation of beliefs ν_i and $\tilde{\nu}_i$ in Equation (16)).

We prove the fully revealing nature of equilibrium in an economy in which probability distributions belong to the exponential family with log-likelihood ratios that are linear in the productivity shock. This family includes normal, gamma, binomial, and many other distributions. The distributions from this family are widely used in models with asymmetric information and CARA investors, and in the most general form have been studied in Breon-Drish (2015). Proposition 3 establishes the main result of this section.

Proposition 3 [Fully revealing equilibrium]. *Suppose the informed and biased investors have probabilities from the same exponential family, given by*

$$\pi_{I,m}(z) = \pi_{B,m}(z) = \exp\{a(x_m) + zx_m - c(z)\}, \tag{17}$$

albeit with different parameters z across investors ($z = \varepsilon$ for the informed and $z = \nu_n$ for n-th biased investors), where $a(x_m)$ and $c(z)$ are functions of the state variable x_m and parameter z, respectively. If $2N + 2 < M$ and $\mu_I \neq \mu_B n/N$, where n is an integer such that $1 \leq n \leq N$, then the equilibrium is fully revealing.

First, we discuss probabilities (17). In the equation for probabilities, x_m stands for the outcomes of a random variable such that $x(\omega_m) = x_m$. This variable may, for example, represent log-output growth, $x(\omega) = \ln(D(\omega)/D_0)$. Negative or positive shock z then captures whether the states with high output become more or less likely, respectively. The scaling factor $c(z)$ ensures that probabilities total to one. Notably, $c(z)$ is related to the cumulant generating function of the distribution, and $c'(z) = \mathbb{E}[x]$.

Proposition 3 shows that the equilibrium is fully revealing when the number of biased investors is below the threshold $N < 0.5M - 1$. In this case, the system of equations for shocks and beleifs is overdetermined, allowing for the unique identification of the shock. Noises ν_n are also revealed in equilibrium but with ambiguity about their order; the prices remain the same even if noises ν_n and ν_k interchange their positions in Equation (16).

Remark 1. Proposition A.1 in the Appendix shows that the results of Proposition 3 remain valid under a weaker definition of rational expectations equilibrium in which uninformed investors learn only from prices. Consequently, the information contained in the residual demands of informed and biased investors is redundant.

4.1. Comparison with economies with CARA investors

We now juxtapose our fully revealing market with logarithmic investors against markets with CARA investors and probabilities from the exponential family that are prevalent in the literature on asymmetric information. Contrary to our conclusions in Proposition 3, equilibria under CARA preferences are not fully revealing. Moreover, the amount of information discernible from asset prices does not depend on the number of traded securities.

We study an economy that maintains the same investor groups, probability structures, and output as described in Section 2.1. However, the investors now have CARA utility with

absolute risk aversion parameter γ and solve the following intertemporal optimization:

$$\max_{C_0,\, C_1(\omega)} -e^{-\gamma C_0} + \beta \mathbb{E}^i\Big[-e^{-\gamma C_1(\omega)}\Big], \tag{18}$$

subject to self-financing budget constraints (2) and (3). Proposition A.2 in the Appendix shows that the optimal final-date consumptions and the optimal quantities $n^*_{i,m}$ of securities in investor i's portfolio are determined as follows:

$$C^*_{1,i}(\omega_m) = n^*_{i,m} = C^*_0 + \frac{\ln(\pi_{i,m}) - \ln(p_m/\beta)}{\gamma}. \tag{19}$$

Next, we find the sufficient statistics. As in Section 3, substituting consumptions (19) into date-1 consumption clearing (6), using date-0 consumption clearing (5), and simplifying, we find the following sufficient statistics in the economy with CARA investors:

$$s_m = \mu_I \ln(\pi_{I,m}(\varepsilon)) + \mu_B \frac{\ln(\pi_{B,m}(\nu_1)) + \cdots + \ln(\pi_{B,m}(\nu_N))}{N}, \tag{20}$$

where $m = 1, \ldots, M$. Substituting exponential-family probabilities (17) into Equation (20), the statistics (20) simplify to:

$$s_m = x_m\Big(\mu_I\varepsilon + \mu_B\frac{\nu_1 + \cdots + \nu_N}{N}\Big) - \Big(\mu_I c(\varepsilon) + \mu_B\frac{c(\nu_1) + \cdots + c(\nu_N)}{N}\Big) + a(x_m).$$

We observe that all these statistics s_m linearly depend on the same two functions of shock ε and noises ν, and hence, the information revealed by these statistics can be further compressed to only two statistics

$$\tilde{s}_1 = \mu_I\varepsilon + \mu_B\frac{\nu_1 + \cdots + \nu_N}{N}, \tag{21}$$

$$\tilde{s}_2 = \mu_I c(\varepsilon) + \mu_B\frac{c(\nu_1) + \cdots + c(\nu_N)}{N}. \tag{22}$$

Shock ε cannot be, in general, identified from the sufficient statistics (21) and (22), and hence, the equilibrium is not fully revealing. Striking property of the latter statistics is that they do not depend on the number of traded securities M, which is also the dimension of the data set consisting of observed sufficient statistics (20). Therefore, markets with either a small or large number of traded securities reveal the same volume of information. The latter result is in sharp contrast with fully revealing equilibria under log utilities in Proposition 3.

We note that the degeneracy of sufficient statistics (20) is due to the special property that both CARA utility and distributions (17) have exponential structures. Consequently,

the uncertainty factor $x_m z + c(z)$ in the expression for probabilities (17), where $z = \varepsilon$ for the informed and $z = \nu_n$ for the n-th biased investors, and an investor's wealth are substitutes in the expected utility function. That is, the optimal portfolios and consumptions are the same as in an economy where investors receive wealth endowments $-(x_m z - c(z) + c(0))/\gamma$ at the final date but have shock $z = 0$ in probabilities (17). This substitutability and the market completeness allow investors to construct portfolios that offset the uncertainty factor, similar to Chabakauri, Yuan, and Zachariadis (2022). The investors' holdings are then linear functions of z and the scaling parameter $c(z)$, giving rise to sufficient statistics (21)–(22).

The sufficient statistic $\tilde{s}_1$ in (21) is linear in shock ε and noises ν_n, similar to models that use CARA preferences, exponential family distributions, and noise traders (Breon-Drish, 2015; Chabakauri, Yuan, and Zachariadis, 2022). Nonlinear statistic $\tilde{s}_2$ in (22), to our best knowledge, is new to the literature. It emerges because the model incorporates date-0 consumption and endogenous interest rates, which standard models with CARA preferences usually do not have. An additional market clearing condition for date-0 consumption then reveals additional information represented by statistic $\tilde{s}_2$.

We summarize the results on equilibrium with CARA investors in Proposition 4 below.

Proposition 4 [Equilibrium with CARA investors]. *Suppose all investors have CARA preferences and beliefs from the exponential family of distributions. Then, the equilibrium is partially revealing and the total amount of the revealed information does not depend on the data set's dimension M.*

Equation (20) also implies that the sufficient statistics reveal more information when the probabilities are not from the exponential family (17). Consider, for example, more general probabilities with quadratic log-likelihood ratios, given by $\pi_{I,m}(z) = \pi_{B,m}(z) = \exp\{a(x_m) + z x_m + z^2 x_m^2 - c(z)\}$. Substituting these probabilities into Equation (20) yields three sufficient statistics, two of which are given by (21) and (22), and the third is $\tilde{s}_3 = \mu_I \varepsilon^2 + \mu_B(\nu_1^2 + \cdots + \nu_N^2)/N$. Similarly, probabilities $\pi_{I,m}(z) = \pi_{B,m}(z) = \exp\{a(x_m) + P_k(z x_m) - c(z)\}$, where $P_k(x)$ is a polynomial of degree k, yield $k+1$ sufficient statistics. Therefore, more information is revealed in equilibrium, and the resulting system of equations becomes overdetermined when $k > N$. As a result, deviating from either CARA preferences or probabilities with linear log-likelihood ratios (17) results in a higher volume of the revealed information.

5. Revealing and non-revealing equilibria with many biased investors and big data

In this section, we investigate economies populated by a substantial number of biased investors. We derive a closed-form characterization of the financial market's informational efficiency as N approaches infinity. The limit in which the number of traded securities M also approaches infinity is particularly tractable and is called *big data economy*. Our findings suggest that in the big data economy the market can be either fully or partially revealing depending on economic conditions.

First, we examine the straightforward case in which the number N of biased investor types becomes large, while holding the mass of informed investors fixed. From Equation (10) for the sufficient statistics, we note that the second term, which represents noise, converges to a deterministic cross-sectional average belief $\overline{\mathbb{E}}[\pi_{B,m}(\nu)]$ by the law of large numbers. The statistics then converge to $s_m = \mu_I \pi_{I,m}(\varepsilon) + \mu_B \overline{\mathbb{E}}[\pi_{B,m}(\nu)]$, revealing the probabilities $\pi_m(\varepsilon)$. Intuitively, as N grows, noise diversifies away, leading to a fully revealing equilibrium.

We now explore our main, more complex, case in which the mass of informed investors is asymptotically small, reflecting the reality that accurate forecasting is a rare skill even among experts (Tetlock, 2009). Specifically, we set $\mu_I = \widehat{\mu}_I/\sqrt{N}$. Then, we replace sufficient statistics s_m with new statistics given by linear transformations $\widehat{s}_m = \sqrt{N}(s_m/\mu_B - \overline{\mathbb{E}}[\pi_{B,m}(\nu)])$, which reveal the same information as s_m. After simplifying, we find that

$$\widehat{s}_m = \frac{\widehat{\mu}_I}{\mu_B}\pi_{I,m}(\varepsilon) + \frac{(\pi_{B,m}(\nu_1) - \overline{\mathbb{E}}[\pi_{B,m}(\nu)]) + \cdots + (\pi_{B,m}(\nu_N) - \overline{\mathbb{E}}[\pi_{B,m}(\nu)])}{\sqrt{N}}, \tag{23}$$

where $m = 1, \ldots, M-1$, and $\overline{\mathbb{E}}[\pi_{B,m}(\nu)]$ are cross-sectional expectations with respect to ν.

Next, we apply the central (or local) limit theorem to (23). We observe that probability vectors $(\pi_{B,1}(\nu_n), \ldots, \pi_{B,M-1}(\nu_n))^\top$ are i.i.d. across biased investors, due to independence of belief parameters ν_n. Consequently, the second terms of statistics (23) have joint normal distribution in the limit $N \to \infty$, allowing to simplify statistics (23) as follows:

$$\widehat{s}_m = \eta \pi_{I,m}(\varepsilon) + \xi_m, \quad \xi \sim \mathcal{N}(0, \Sigma_\xi), \tag{24}$$

where $m = 1, \ldots, M-1$, $\eta = \widehat{\mu}_I/\mu_B$ is the measure parameter of informed investors relative to biased investors, $\xi = (\xi_1, \ldots, \xi_{M-1})^\top$, $\mathcal{N}(0, \Sigma_\xi)$ is multivariate normal distribution, and $\Sigma_\xi \in \mathbb{R}^{M-1,M-1}$ is a symmetric cross-sectional covariance matrix of the probabilities of biased investors with elements equal to covariances $\overline{\mathrm{cov}}(\pi_{B,m}(\nu), \pi_{B,k}(\nu))$ with respect to beliefs ν.

The equilibrium with statistics (24) is not fully revealing for finite η. Intuitively, when N is large, the noise introduced by biased investors is reduced through diversification. Nevertheless, because the mass of informed investors is small, discerning information about ε becomes more challenging. It is also noteworthy that Equations (24) encompass two extremes as special cases, fully revealing and non-revealing equilibria when $\eta = +\infty$ and $\eta = 0$, respectively. The former equilibrium emerges when μ_I does not approach zero as N grows, whereas the latter arises when $\mu_I = o(1/\sqrt{N})$, implying that μ_I converges to zero too fast.

We call the economy in which uninformed investors make inference from statistics (24) a *Gaussian-signal economy*. Lemma A.3 in the Appendix demonstrates that as $N \to \infty$, asset prices converge to those in the Gaussian-signal economy. Our subsequent focus is on the equilibrium in this economy. Specifically, we probe into the total amount of information revealed in this economy and its relation with the aggregate number of traded securities, M. This relationship, inherent to our model, is absent in economies featuring CARA investors and exponential-family distributions, as demonstrated in Section 4.1.

We show that as the number of traded securities M increases, the equilibrium converges to either fully or partially revealing, contingent on economic conditions. The character of the equilibrium in the limit is determined by the balance between the following two effects. On the one hand, economies with a larger number of securities offer more sources of information given by the sufficient statistics (24). On the other hand, when the number of states M is so large that our discrete state-space is close to a continuous one, the consecutive statistics $\widehat{s}_m$ and $\widehat{s}_{m+1}$ tend to converge in their informational content, diminishing the incremental value of additional statistics. Consider, as an example, call options with different strikes written on log-output. As strikes become more and more dense so that the number of available options increases, on one hand, the number of data points from which investors infer information increases. On the other hand, options with similar strikes have similar prices and trading volumes, and hence, reveal similar information.

To streamline the exposition, we demonstrate our results in a parsimonious setup in which shock ε can take only two values, normalized to zero and one, $\varepsilon_L = 0$ and $\varepsilon_H = 1$. These values might correspond to situations in which mean output is growing ($\mathbb{E}[\ln(D_1/D_0)] > 0$) or not ($\mathbb{E}[\ln(D_1/D_0)] = 0$), respectively. The corresponding prior probabilities of ε_L and ε_H are given by φ_L and $\varphi_H = 1 - \varphi_L$, respectively.

We now derive the *measure of informational efficiency*. Let $\widehat{\pi}_L(\widehat{s}) = \text{Prob}(\varepsilon = \varepsilon_L|\widehat{s})$. Given that the signals depend on the random beliefs of biased investors, this probability is inherently random. Consider the expected posterior probability conditioned on the true

shock being ε_L, denoted by $\mathbb{E}[\widehat{\pi}_L(\widehat{s})|\varepsilon = \varepsilon_L]$. Latter expected probability being close to one indicates that the uninformed investors, on average, correctly assign high probability to shock $\varepsilon = \varepsilon_L$. Consequently, we will say that the *equilibrium is close to fully revealing* when $\mathbb{E}[\widehat{\pi}_L(\widehat{s})|\varepsilon = \varepsilon_L] \geq 1 - q$ and $\text{var}[\widehat{\pi}_L(\widehat{s})|\varepsilon = \varepsilon_L] \leq q$ for a sufficiently small q. We note that the latter condition on variance, added for completeness, is redundant. Because $\widehat{\pi}_L(\widehat{s}) \in [0, 1]$, the expected value of $\widehat{\pi}_L(\widehat{s})$ close to one necessarily implies that its variance is close to zero.[2] Hence, the expected posterior probability measures *proximity of equilibrium to fully revealing*. Lemma below derives the expected probability $\mathbb{E}[\widehat{\pi}_L(\widehat{s})|\varepsilon = \varepsilon_L]$ analytically, and defines the quantity that we adopt as a convenient measure of informational efficiency.

Lemma 1 [Informational efficiency]. *Let $f \in \mathbb{R}^M$ be a vector with elements $f_m = \eta(\pi_{I,m}(\varepsilon_H) - \pi_{I,m}(\varepsilon_L))$ and $\mathcal{K} \in \mathbb{R}^{M\times M}$ be a matrix with elements $\overline{\mathbb{E}}[\pi_{B,m}(\nu)\pi_{B,k}(\nu)]$, $m, k = 1, \ldots, M$, where $\overline{\mathbb{E}}[\cdot]$ is an expectation with respect to beliefs PDF $\varphi_\nu(x)$. Then,*

$$\mathbb{E}[\widehat{\pi}_L(\widehat{s})|\varepsilon = \varepsilon_L] = g(\sqrt{J_M^*}), \tag{25}$$

where $\widehat{\pi}_L(\widehat{s}) = \text{Prob}(\varepsilon = \varepsilon_L|\widehat{s})$, $g(y)$ is a monotonically increasing function such that $g(0) = \varphi_L$ and $g(y) \to 1$ as $y \to \infty$, given by Equation (A.29) in the Appendix, and $J_M^* = f^\top \mathcal{K}^{-1} f$ is a measure of informational efficiency.

Lemma 1 establishes that a high value of the measure of informational efficiency J_M^* indicates that the equilibrium is close to fully revealing. This measure, given by $J_M^* = f^\top \mathcal{K}^{-1} f$, captures the quantity of information relative to noise. Vector f has elements $f_m = \eta(\pi_{I,m}(\varepsilon_H) - \pi_{I,m}(\varepsilon_L))$. Small norm $\|f\|_2$ of this vector indicates that the informed investors' probability distributions for good and bad shocks ε are close. As a result, these investors trade similarly in both states, making it challenging to identify the shock from the data, and hence, the measure of informativeness is low. Matrix $\mathcal{K}$ has elements given by cross-sectional expectations $\overline{\mathbb{E}}[\pi_{B,m}(\nu)\pi_{B,k}(\nu)]$ and measures the average co-movement of probabilities. The informational efficiency tends to be higher when this matrix is closer to singular, indicating similarity of probabilities across states. This is because such similarity implies that the noises ξ_m in Equation (24) have a large common component. Adjusting for this component diminishes noise, thereby improving informational efficiency.

Determining the measure of informational efficiency $J_M^* = f^\top \mathcal{K}^{-1} f$ reduces to solving a system of linear equations, given by $\mathcal{K}v = f$, where $\mathcal{K}$ is a symmetric, positive-definite, and invertible matrix (by Lemma A.4 in the Appendix). Consequently, there exists unique vector

[2] More formally, Suppose, $\mathbb{E}[\widehat{\pi}_L(\widehat{s})|\varepsilon = \varepsilon_L] \geq 1 - d$, where d is small. Then, $\text{var}[\widehat{\pi}_L(\widehat{s})|\varepsilon = \varepsilon_L] = \mathbb{E}[\widehat{\pi}_L(\widehat{s})^2|\varepsilon = \varepsilon_L] - (\mathbb{E}[\widehat{\pi}_L(\widehat{s})|\varepsilon = \varepsilon_L])^2 \leq 1 - (1 - d)^2 < 2d$. Then, we define $q = 2d$.

v solving that equation. Hence, the informational efficiency can be expressed as $J_M^* = f^\top v$. Finding solution v in practice, however, is an ill-posed problem due to near-singularity of matrix $\mathcal{K}$ when the number of states M is large. Such near-singularity often arises because, in cases with a large M, the probabilities of consecutive states m and $m+1$ are similar, causing rows of the matrix $\mathcal{K}$ to be nearly collinear. Furthermore, because of tail events ω_m with small probabilities, matrix $\mathcal{K}$ has many near-zero elements even on the main diagonal, resulting in a near-singularity. As a result, the set of approximate solutions v_ϵ such that $\|\mathcal{K}v_\epsilon - f\| \leq \epsilon$, for a sufficiently small ϵ, appears to be extremely wide.

We tackle the challenge of determining the solution v of the ill-posed system $\mathcal{K}v = f$ by employing Tikhonov regularization (Tikhonov and Arsenin, 1977). This method is commonly adopted in machine learning and empirical asset pricing, where it is also known as ridge regression. In particular, we look for v_τ that solves the following variation of the standard optimization problem in ridge regression:

$$v_\tau = \underset{v}{\operatorname{argmax}} \, \|\mathcal{K}v - f\|_2^2 + \tau v^\top \mathcal{K} v, \tag{26}$$

where $\tau > 0$. The first order condition for optimization (26) is given by

$$\mathcal{K}v_\tau + \tau v_\tau = f. \tag{27}$$

System of equations (27) is significantly easier to solve than the original system with $\tau = 0$. We then introduce the *Tikhonov-regularized measure of informational efficiency* $J_M(\tau) = f^\top v_\tau$, such that $J_M(0) = J_M^*$. Lemma 2 below establishes important properties of $J_M(\tau)$.

Lemma 2 [Low bound on informational efficiency]. *The regularized informational efficiency $J_M(\tau)$ is a continuous and decreasing function of τ for all $\tau \geq 0$. Therefore, $J_M(\tau)$ is a low bound on the informational efficiency J_M^*:*

$$J_M(\tau) \leq J_M^*. \tag{28}$$

Lemma 2 offers an easily computable low bound $J_M(\tau)$ for the measure of informational efficiency. If this bound is large for some $\tau > 0$, then the equilibrium is close to fully revealing by Lemma 1. Additionally, as regularization parameter τ shrinks to zero, the regularized measure $J_M(\tau)$ converges to the efficiency measure J_M^*. Below, we study two economies in which the efficiency measure can be characterized in closed firm. The first economy has continuous state-space in the limit as M becomes very large, and in the second economy the state-space is discrete and the number of types of biased investors is finite.

5.1. Big data economy

Below, we characterize the regularized measure of informational efficiency $J_{M}(\tau)$ in a big data economy with a large number of states M. We find this measure of efficiency in closed form for a particular example of state probabilities. Then, we take limit $\tau \to 0$ and find conditions on the model parameters under which the equilibrium is fully or partially revealing.

We define the probabilities of states as discretizations of continuous-space PDFs $\pi_{I}(x;\varepsilon)$ and $\pi_{B}(x;\nu)$. To simplify the exposition, we assume that $x \in \mathbb{R}$. The state probabilities of informed and biased investors in terms of these PDFs are given by

$$\pi_{I,m}(\varepsilon) = \frac{\pi_{I}(x_m;\varepsilon)\Delta x}{\sum_{k=1}^{M} \pi_{I}(x_k;\varepsilon)\Delta x}, \quad \pi_{B,m}(\nu) = \frac{\pi_{B}(x_m;\nu)\Delta x}{\sum_{k=1}^{M} \pi_{B}(x_k;\nu)\Delta x}, \tag{29}$$

where $x_m = -A_{M} + m\Delta x$ and $\Delta x = 2A_{M}/M$ is the *granularity of state-space* Δx. We assume that $A_{M} \to \infty$ and $\Delta x \to 0$ as $M \to \infty$, so that the probabilities (29) converge to distributions $\pi_{I}(x;\varepsilon)$ and $\pi_{B}(x;\nu)$, respectively. As the dimension M of the data set of sufficient statistics increases, investors gain access to more sources of information. However, adjacent statistics (10) become increasingly similar, thereby revealing similar information. Therefore, the net impact on informational efficiency is unclear without additional analysis.

As the number of states M becomes very large, the system of linear equations (27), which we use for computing the measure of informational efficiency, is replaced with the following *Fredholm integral equation of the second kind*:

$$\int_{-\infty}^{\infty} K(x,y)v(y;\tau)dy + \tau v(x;\tau) = f(x), \tag{30}$$

where $f(x) = \eta(\pi_{I}(x;\varepsilon_{H}) - \pi_{I}(x;\varepsilon_{L}))$ and $K(x,y) = \overline{\mathbb{E}}[\pi_{B}(x;\nu)\pi_{B}(y;\nu)]$ is the expectation with respect to belief parameter ν.

Next, we solve Equation (30) analytically when the distributions are given by

$$\pi_{I}(x;\varepsilon) = \frac{\exp(-0.5(x-\mu\varepsilon)^2/\sigma_{I}^2)}{\sqrt{2\pi\sigma_{I}^2}}, \quad \pi_{B}(x;\nu) = \frac{\exp(-0.5(x-\nu)^2/\sigma_{B}^2)}{\sqrt{2\pi\sigma_{B}^2}}, \tag{31}$$

where $\varepsilon \in \{\varepsilon_{L} = 0, \varepsilon_{H} = 1\}$. In this specification, we interpret state variable x as the log-output growth, so that $x = \ln(D_1/D_0)$. Therefore, belief parameters ν represent expected log-output growth next period. The shocks ε_{L} and ε_{H} have unconditional probabilities φ_{H} and φ_{L}, respectively. We also assume that beliefs are normally distributed in the cross-section of biased investors, $\nu \sim \mathcal{N}(0,\sigma_{\nu}^2)$. Proposition below presents the kernel $K(x,y)$ and the solution of Equation (30) in closed form.

Proposition 5 [Fredholm integral equation and informational efficiency]. *Suppose the state probabilities from the viewpoints of informed and biased investors are given by Equations (31), respectively, and beliefs ν are normally distributed in the cross-section of biased investors, $\nu \sim \mathcal{N}(0, \sigma_\nu^2)$. Then, the following statements hold.*

1) *Fredholm integral equation (30) has Gaussian kernel given by*

$$K(x,y) = \frac{\widehat{\sigma}}{2\pi\sigma_B^2\sigma_\nu} \exp\left(-\frac{\widehat{\sigma}^2}{2\sigma_B^4}(x-y)^2 - \frac{\widehat{\sigma}^2}{2\sigma_B^2\sigma_\nu^2}(x^2+y^2)\right), \tag{32}$$

where $\widehat{\sigma} = (2/\sigma_B^2 + 1/\sigma_\nu^2)^{-1/2}$.

2) *The corresponding integral operator $\mathcal{K}$ in (30) has the following set of orthonormal eigenfunctions and eigenvalues:*

$$\psi_n(x) = \sqrt{a}\exp(-0.5a^2x^2)\frac{H_n(ax)}{\sqrt{2^n n!\sqrt{\pi}}}, \tag{33}$$

$$\lambda_n = \frac{1}{\sqrt{\pi\sigma_B^2}}\frac{\widehat{\sigma}}{\sigma_\nu+\widehat{\sigma}}\left(\frac{\sigma_\nu-\widehat{\sigma}}{\sigma_\nu+\widehat{\sigma}}\right)^n, \tag{34}$$

where $\lambda_n > 0$, $a = \sqrt{\widehat{\sigma}/(\sigma_\nu\sigma_B^2)}$, $\widehat{\sigma} = (2/\sigma_B^2 + 1/\sigma_\nu^2)^{-1/2}$, and $H_n(x)$ are Hermite polynomials. Eigenfunctions $\psi_n(x)$ form an orthonormal basis in $\mathcal{L}_2(\mathbb{R})$.

3) *Equation (30) has unique solution $v(x;\tau) \in \mathcal{L}_2(\mathbb{R})$. This solution and the regularized measure of informational efficiency $J_\infty(\tau)$ are given by*

$$v(x;\tau) = \sum_{n=0}^{\infty}\frac{f_n\psi_n(x)}{\tau+\lambda_n}, \tag{35}$$

$$J_\infty(\tau) = \sum_{n=0}^{\infty}\frac{f_n^2}{\tau+\lambda_n}, \tag{36}$$

where f_n are Fourier coefficients of function $f(x)$ in (30) with respect to orthonormal basis $\psi_n(x)$, given by Equation (A.39) in the Appendix in closed form. The series (35) converges with respect to a norm in $\mathcal{L}_2(\mathbb{R})$ and series (36) converges to $J_\infty(\tau)$, for all $\tau > 0$.

Proposition 5 provides in closed-form the solution of Equation (30) and the lower bound on the measure of informational efficiency (36). This lower bound depends on the Fourier coefficients f_n of the function $f(x)$, which represents the difference of the informed investors' PDFs under good and bad productivity shocks. Given that the eigenvalues λ_n are strictly positive, the proposition also demonstrates that the lower bound (36) is a decreasing function of the regularization parameter τ, consistent with the findings in Lemma 2.

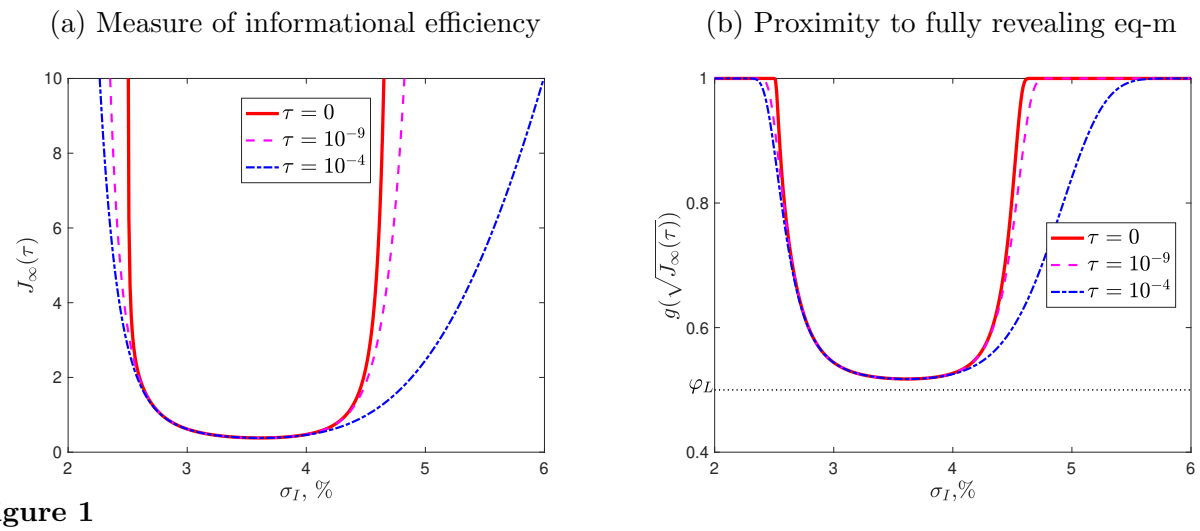


Figure 1
Informational efficiency and proximity to fully revealing equilibrium
Panel (a) shows the measure of informational efficiency $J^*_\infty = J_\infty(0)$ (solid line) and its low bounds $J_\infty(\tau)$ for $\tau = 10^{-9}$ and $\tau = 10^{-4}$ as functions of log-output volatility σ_I of informed investors' PDF $\pi_I(x;\varepsilon)$ in (31). The equilibrium is partially revealing when measure J^*_∞ is finite and fully revealing otherwise. Panel (b) shows how close the equilibrium is to fully revealing. The proximity to fully revealing equilibrium is given by $\mathbb{E}[\widehat{\pi}_L(\widehat{s})|\varepsilon = \varepsilon_L] = g(\sqrt{J^*_\infty})$ (solid line) and its lower bounds are given by $g(\sqrt{J_\infty(\tau)})$ (dash and dash-dot lines), and $g(\sqrt{J^*_\infty}) = 1$ (< 1) indicates a fully (partially) revealing equilibrium. The parameters of PDFs (31) are $\sigma_B = 2.5\%$ and $\mu = 2\%$, noise ν has standard deviation $\sigma_\nu = 3\%$, and mass parameter of informed investors is $\eta = 0.5$.

Next, we take limit $\tau \to 0$ in Equation (36) and establish conditions under which the limiting value is bounded or unbounded. As demonstrated in Lemma 1, unbounded (bounded) $J_\infty(0)$ implies that the equilibrium is fully (partially) revealing. Proposition 6 provides the conditions for determining the nature of equilibrium and derives the limiting value $J^*_\infty = J_\infty(0)$ in analytical form.

Proposition 6 [Informational efficiency]. *The equilibrium is partially revealing if*

$$\sigma_B < \sigma_I < \sqrt{\sigma_B^2 + 2\sigma_\nu^2}, \tag{37}$$

and is fully revealing otherwise. In the fully revealing equilibrium the measure of informational efficiency is $J^*_\infty = \infty$ and in the partially revealing equilibrium it is given by:

$$\begin{aligned} J^*_\infty \;=\; & \frac{\eta^2 a}{\sqrt{\pi(1-\rho^2)\lambda_0(1+a^2\sigma_I^2)}}\left[\exp\left\{-\frac{\mu^2a^2}{1+a^2\sigma_I^2}\left(1-\frac{2\rho}{(1+\rho)(1-a^2\sigma_I^2)}\right)\right\}\right. \\ & \left. -2\exp\left\{-\frac{1}{2}\frac{\mu^2a^2}{1+a^2\sigma_I^2}\left(1+\frac{2\rho^2}{(1-\rho^2)(1-a^2\sigma_I^2)}\right)\right\}+1\right], \end{aligned} \tag{38}$$

where a and ρ are given by:

$$a = \frac{1}{\sigma_B}\sqrt{\frac{\widehat{\sigma}}{\sigma_\nu}}, \quad \rho = \frac{1 - a^2\sigma_I^2}{1 + a^2\sigma_I^2} \cdot \frac{\sigma_\nu + \widehat{\sigma}}{\sigma_\nu - \widehat{\sigma}}, \tag{39}$$

and $\widehat{\sigma} = (2/\sigma_B^2 + 1/\sigma_\nu^2)^{-1/2}$.

Proposition 6 establishes the conditions determining whether the equilibrium is partially or fully revealing. Specifically, the equilibrium is fully revealing if the volatility of log-output σ_I from the informed investor's PDFs in (31) satisfies inequalities (37). The equilibrium is fully revealing if these inequalities are not satisfied, which occurs when the volatility σ_I is either excessively low or high. The intuition behind this finding is as follows. As shown in Proposition 1, the holdings of Arrow-Debreu securities by informed investors are directly proportional to the probabilities of respective states. When the volatility parameter σ_I is relatively low, the informed investors' PDF in (31) is concentrated around its mean, given by $\mu > 0$ when $\varepsilon = 1$ and zero when $\varepsilon = 0$. As a result, the informed investors predominantly hold Arrow-Debreu securities of states x clustering around their distribution's mean. For instance, when $\varepsilon = 1$, the uninformed investors would notice a heightened buying activity around states x near μ compared to those near zero. This allows them to deduce that the informed investors have observed good signal $\varepsilon = 1$. When the volatility parameter σ_I exceeds $\sqrt{\sigma_B^2 + 2\sigma_\nu^2}$, the PDFs of informed investors, and consequently their asset holdings, exhibit a broader dispersion around the distribution's mean than the dispersion of trades by biased investors, captured by $\sqrt{\sigma_B^2 + 2\sigma_\nu^2}$. This leads to heightened trading activity even at positions distant from the mean and beyond plausible trading range of biased investors. Therefore, uninformed investors are able to discern shock ε from the noise.

Figure 1 plots the measure of efficiency J_∞^* and its lower (regularized) bounds $J_\infty(\tau)$ for different values of the regularization parameter τ as functions of the informed investor's log-output volatility in (31).[3] When $\sigma_B < \sigma_I < \sqrt{\sigma_B^2 + 2\sigma_\nu^2}$ and σ_I is farther away from the boundaries, the informational efficiency is low and the equilibrium is partially revealing. In this case, the information overlap across sufficient statistics prevents uninformed investors from perfectly learning shock ε. The informational efficiency is infinite outside of the latter interval, in which case the equilibrium is fully revealing. We note that the inequality $\sigma_B < \sigma_I$

[3]The volatility parameter σ_I ranges from 2% to 6%, and the mean growth rate (conditional on a realization of shock ε_H) is set at $\mu = 2\%$, consistent with the evidence that the unconditional volatility and mean consumption growth are around 3.2% and 1.8%, respectively (e.g., Basak and Cuoco (1998), Campbell (2003), among others). The consumption growth volatility from the viewpoint of biased investors is set at $\sigma_B = 2.5\%$. The standard deviation of biased investors mean-growth beliefs is $\sigma_\nu = 3\%$ and is a single-digit number so that investors beliefs have the same order of magnitude as unconditional mean growth of 1.8%.

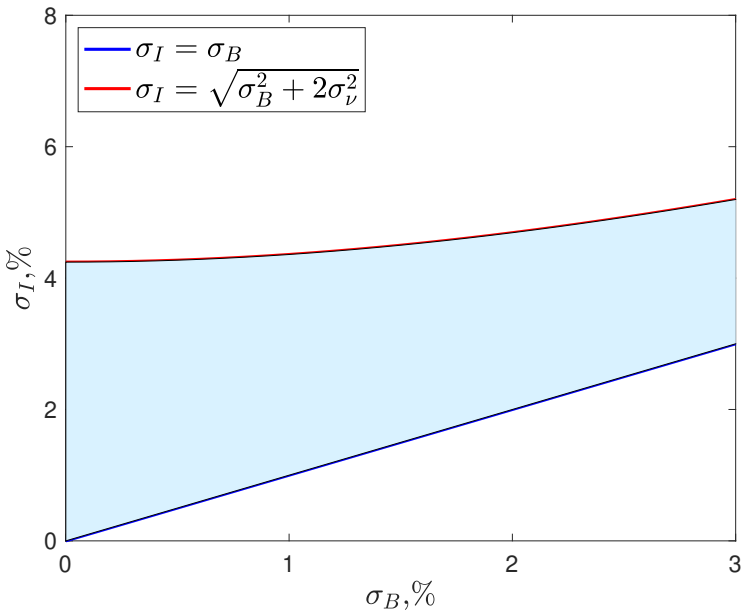


Figure 2
Set of parameters leading to partially revealing equilibria
This Figure shows the shaded region of parameter space in which inequality $\sigma_B < \sigma_I < \sqrt{\sigma_B^2 + 2\sigma_\nu^2}$ from Proposition 6 is satisfied, and hence, the equilibrium is partially revealing. The equilibrium is fully revealing outside of this region. We set $\sigma_\nu = 3\%$.

means that biased investors assign lower volatility to log-output under their beliefs than informed investors. This scenario is consistent with overconfidence bias among biased investors, similar to Daniel, Hirshleifer, and Subrahmanyam (1998). Consequently, partially revealing equilibrium is likely to arise when overconfidence is prevalent among biased investors.

Our findings imply that derivative securities in financial market reveal valuable information about economic fundamentals, in line with empirical literature on price discovery in financial markets showing that prices and trading volumes of derivatives reveal information about the payoffs of underlying assets (e.g., Easley, O'Hara, and Srinivas (1998); Chakravarty, Gulen, and Mayhew (2004); Pan and Poteshman (2006), Kacperczyk and Pagnotta (2019)). The result is also in line with finings in Bai, Philippon, and Savov (2016), who show that price informativeness is higher for firms with higher levels of option turnover.

Next, we explore how the asymmetry of information impacts interest rates and show that interest rates predict future economic conditions. Specifically, we show that interest rates tend to be higher (lower) when informed investors anticipate high (low) output growth rate next period. The following proposition characterizes the interest rates when the informed and biased investors have normal PDFs (31) over log-output growth $x = \ln(D_1/D_0)$.

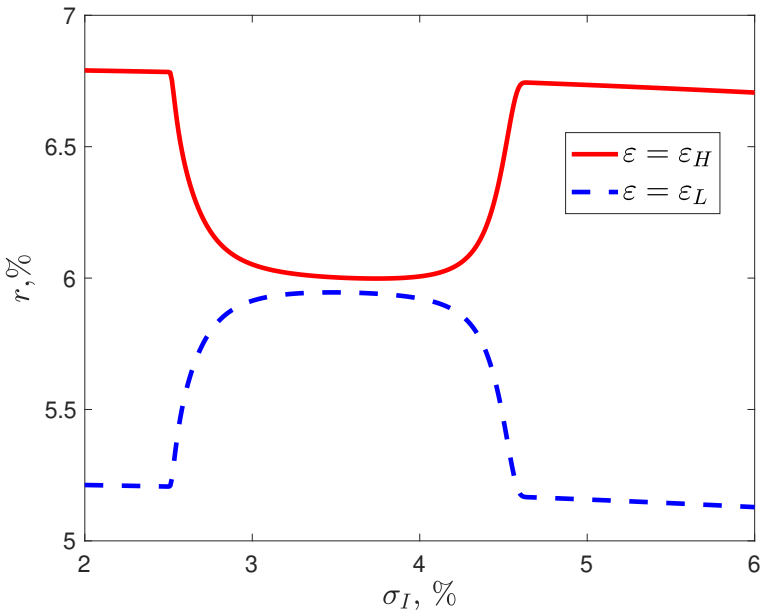


Figure 3
Expected conditional interest rates $\mathbb{E}[r_f(\widehat{s})|\varepsilon]$
This Figure shows the expected interest rates conditional of the realization of shock ε. Solid line corresponds to $\varepsilon = \varepsilon_H$ and dashed line to $\varepsilon = \varepsilon_L$. The model parameters are as for Figure 1.

Proposition 7 [Interest rate]. *The riskless interest rate in the Gaussian-signal economy with binary shock* $\varepsilon \in \{\varepsilon_H = 1, \varepsilon_L = 0\}$ *is given by*

$$r_f(\widehat{s}) = \frac{1/\beta}{\mu_B \exp\{0.5(\sigma_\nu^2 + \sigma_B^2)\} + \mu_U \left[\exp\{-\mu + 0.5\sigma_I^2\}(1 - \widehat{\pi}_L(\widehat{s})) + \exp\{0.5\sigma_I^2\}\widehat{\pi}_L(\widehat{s})\right]} - 1, \tag{40}$$

where $\widehat{\pi}_L(\widehat{s}) = \text{Prob}(\varepsilon = \varepsilon_L|\widehat{s})$ *is given by Equation* (A.28) *in the Appendix.*

Proposition 7 sheds light on how interest rates aggregate information. Equation (40) reveals that the interest rate is a decreasing function of a bad shock conditional probability $\widehat{\pi}_L(\widehat{s}) = \text{Prob}(\varepsilon = \varepsilon_L|\widehat{s})$. When the informed investors receive bad signal ε_L, this probability goes up, leading to lower interest rates, and vice versa for a good signal ε_H. Therefore, interest rates are high when output growth is expected to be high, and low otherwise. This is because the abundance of consumption in the future reduces the value of Arrow-Debreu securities, and hence, also of riskless bonds, driving interest rates up. This relationship between interest rates and economic conditions is consistent with the evidence in Harvey (1988) and Ang, Piazzesi, and Wei (2006) that information in yields predicts GDP growth.

Figure 3 plots conditional expected interest rates $\mathbb{E}[r_f(\widehat{s})|\varepsilon_L]$ and $\mathbb{E}[r_f(\widehat{s})|\varepsilon_H]$ as functions

of volatility σ_I (a critical parameter determining whether the equilibrium is fully revealing). We observe that, consistent with Proposition 7, interest rates tend to be higher when economic conditions are better (i.e., $\varepsilon = \varepsilon_H$). Figure 3 also illustrates new predictions of the model that the interest rates are highly nonlinear functions of the output volatility σ_I and jump when volatility σ_I crosses the boundaries separating regions in the parameter space corresponding to fully and partially revealing equilibria, depicted on Figure 2. For example, for the case of $\varepsilon = \varepsilon_H$, the interest rates are lower in the region in which the equilibrium is not fully revealing, and hence, the uncertainty is high. As volatility σ_I crosses into the region where the equilibrium is fully revealing, the conditional probability of bad shock ε_L jumps down to zero, causing interest rates to increase. We note that risk-neutral probabilities also depend on conditional probability $\widehat{\pi}_L(\widehat{s})$, as shown by Equation (13). Therefore, various securities priced using these probabilities also experience similar jumps as interest rates.

5.2. Economy with finite number of investor types.

We consider an economy with M states in which the number of biased investors is large, $N \to \infty$, but these investors have one of M types of beliefs. In particular, each biased investor's beliefs are represented by a probability vector that is an i.i.d. draw from one of M probability vectors $\Pi_1, \ldots, \Pi_M \in \{(\pi_1, \ldots, \pi_M)^\top \in \mathbb{R}^M_+ : \pi_1 + \cdots + \pi_M = 1\}$, where Π_m is drawn with probability $\tilde{\pi}_m$. Vectors Π_m are linearly independent, but general otherwise. By law of large numbers, probability $\tilde{\pi}_m$ equals the mass of investors with beliefs Π_m.

The true probabilities are known to the informed investors and are given by vector $\Pi_I(\varepsilon) \in \mathbb{R}^M_+$, where $\varepsilon = \varepsilon_L$ or $\varepsilon = \varepsilon_H$ with unconditional probabilities φ_L and φ_H, respectively, as in the baseline analysis. The uninformed investors learn about shock ε from sufficient statistics (24), in which $\pi_{I,m}(\varepsilon)$ are the elements of vector $\Pi_I(\varepsilon)$. Because the probability vectors Π_m are linearly independent, they form a basis in $\mathbb{R}^M$. Consequently, the probability vectors of informed investors $\Pi_I(\varepsilon)$ can be expanded into this basis as follows:

$$\Pi_I(\varepsilon) = d_1(\varepsilon)\Pi_1 + \cdots + d_M(\varepsilon)\Pi_M. \tag{41}$$

Proposition 8 below derives closed-form expression for the measure of informational efficiency.

Proposition 8 [Measure of informational efficiency]. *The measure of informational efficiency in the economy with M types of biased investors is given by:*

$$J^*_M = \eta^2 \sum_{m=1}^{M} \frac{(d_m(\varepsilon_H) - d_m(\varepsilon_L))^2}{\tilde{\pi}_m}, \tag{42}$$

where $d_m(\varepsilon)$ are coefficients from the expansion (41) of informed investors' probabilities, $\eta = \widehat{\mu}_I/\mu_B$, and $\tilde{\pi}_m$ are masses of type-m biased investors.

Proposition 8 shows that the measure of informational efficiency (42) is inversely related to the masses $\tilde{\pi}_m$ of biased investor types and positively related to squared differences $(d_m(\varepsilon_H) - d_m(\varepsilon_L))^2$ of coefficients in the expansions of the informed investors probabilities (41). Next, we discuss the intuition of the latter result.

We start with a case in which the informed investors probabilities $\Pi_I(\varepsilon)$ are drawn from the same set of vectors $\Pi_1, \ldots, \Pi_M$ as the probabilities of biased investors. Consequently, for some indices i and j, the informed investors probabilities are given by $\Pi_I(\varepsilon_H) = \Pi_i$ and $\Pi_I(\varepsilon_L) = \Pi_j$. This means that some biased investors have correct beliefs by coincidence. These investors should not, however, be interpreted as informed because, e.g., type-i biased investors may hold beliefs $\Pi_i = \Pi_I(\varepsilon_H)$ even when the shock is $\varepsilon = \varepsilon_L$. In this example, the expansion coefficients in equation (41) are given by $d_m(\varepsilon_H) = 0$ if $m \neq i$ and $d_i(\varepsilon_H) = 1$, and $d_m(\varepsilon_L) = 0$ if $m \neq j$ and $d_j(\varepsilon_L) = 1$. Consequently, the measure of efficiency is given by

$$J_M^* = \eta^2 \Big(\frac{1}{\tilde{\pi}_i} + \frac{1}{\tilde{\pi}_j} \Big). \tag{43}$$

Equation (43) shows that the amount of information revealed in the market is inversely related to masses $\tilde{\pi}_i$ and $\tilde{\pi}_j$ of type i and j investors. This is because it is more difficult to detect how informed investors trade when many biased investors trade using the same probabilities but at wrong times (e.g., type-i investors use probabilities Π_i even when the true probabilities are Π_j, and vice versa). As the number of investor types M increases so that the distribution of biased investor types approaches a continuous distribution, masses of investor types $\tilde{\pi}_i$ and $\tilde{\pi}_j$ shrink to zero. Consequently, the efficiency measure (43) becomes infinite, and the equilibrium is fully revealing in the limit.

Next, we explore a more general case in which the coefficients in the expansion (41) of informed probabilities are more spread out. Suppose, the distribution of biased investor types is given by $\tilde{\pi}_m = \tilde{\pi}(x_m)\Delta x$, where $\tilde{\pi}(x_m)$ is some probability density function, and $\Delta x = x_{m+1} - x_m$. Furthermore, it would be convenient to express weights $d_m(\varepsilon_H)$ and $d_m(\varepsilon_L)$ in Equation (41) as proportions of probabilities of types so that, by definition, $d_m(\varepsilon_H) := \tilde{d}_H(x_m)\tilde{\pi}(x_m)\Delta x$ and $d_m(\varepsilon_L) := \tilde{d}_L(x_m)\tilde{\pi}(x_m)\Delta x$. In this case, as $\Delta x \to 0$ and the distribution of types approaches a continuous function, substituting $d_m(\varepsilon_H)$ and $d_m(\varepsilon_L)$ into the efficiency measure (42), we find that this measure converges to an integral

$$J_M^* = \eta^2 \int_X \Big(\tilde{d}_H(x) - \tilde{d}_L(x) \Big)^2 \tilde{\pi}(x) dx, \tag{44}$$

where X is the support of the density function of types.

The expression on the right-hand side of Equation (44) can be interpreted as the distance between the probability distributions of informed investors for shocks ε_H and ε_L, respectively. Measure J_M^* is low when the distributions are similar, and hence, it is more difficult for the uninformed investors to identify whether the informed investors observe shock ε_H or ε_L. Conversely, when the distributions diverge significantly, informed investors exhibit markedly different trading behaviors in response to shocks ε_H and ε_L, and it becomes easier for the uninformed investors to identify shock ε correctly from the trading data. As a result, the measure of informational efficiency increases.

6. Conclusion

In this paper, we develop a theory of informational efficiency and asset pricing in large markets with many tradable securities in the presence of asymmetric information. Our findings reveal that such markets can potentially disclose a significant amount of information regarding economic shocks, which are known only to a subset of informed investors. We establish specific conditions that lead to either fully or partially revealing equilibria, depending on economic parameters. The paper emphasizes the critical role of investors' utility functions and the structure of probability distributions in determining the extent of information revealed in financial markets. We find that models with CARA investors and log-likelihood ratios linearly related to economic shocks might not fully capture all the information disclosed in financial markets. We address this issue by introducing a novel approach to analyzing asset prices under asymmetric information based on logarithmic utilities and market completeness. The paper also contributes to the literature by developing new methodology for recovering signals about economic shock from large data sets.

Appendix: Proofs

Proof of Proposition 1. Due to market completeness, investor i's budget constraint can be rewritten in the following static form:

$$C_{0,i} + \mathbb{E}^i\Big[\frac{1}{1+r_f}\frac{\pi^{\text{RN}}(\omega)}{\pi_i(\omega)}C_{1,i}(\omega)\Big] = W_0, \tag{A.1}$$

where $\pi^{\text{RN}}(\omega)/((1+r_f)\pi(\omega))$ is the stochastic discount factor. We then maximize the expected utility function (1) subject to the budget constraint (A.1). Using the method of Lagrange multipliers, we find that

$$C^*_{0,i} = \frac{1}{\lambda}, \quad C^*_{1,i}(\omega_m) = \frac{\beta}{\lambda}\frac{(1+r_f)\pi_{i,m}}{\pi^{\text{RN}}_m}, \tag{A.2}$$

where λ is the Lagrange multiplier for the constraint (A.1). Substituting consumptions (A.2) into the budget constraint (A.1), we find the Lagrange multiplier $\lambda = (1+\beta)/W_0$. Substituting this multiplier into (A.2), we obtain consumptions reported in Proposition 1.

Next, substituting consumptions (8) into the time-1 budget constraint (3), we obtain the following equations for determining portfolio weights θ^*_i:

$$\ell W_0 \frac{(1+r_f)\pi_{I,m}(\varepsilon)}{\pi^{\text{RN}}_m} = \frac{W_0\theta^*_{i,m}}{p_m}. \tag{A.3}$$

Equation (A.3) follows from the fact that state-ω_m consumption is financed by the payoffs of state-ω_m Arrow-Debreu securities held by the investor. Noting that for the state-m Arrow-Debreu security $p_m = \pi^{\text{RN}}_m/(1+r_f)$, cancelling terms in (A.3), we find portfolios (9). ■

Proof of Proposition 2. Substituting optimal date-0 consumptions $C^*_{0,i} = W_0/(1+\beta)$, derived in Proposition 1, into the market clearing condition (5), we find that $W_0 = (1+\beta)D_0$. Moreover, the uninformed investors observe the sufficient statistics s, and hence, compute posterior probability $\pi_{U,m} = \mathbb{E}[\pi_{I,m}(\varepsilon)|s]$. Substituting the latter expressions for W_0 and $\pi_{U,m}$ into the rewritten market clearing condition (11), we find the expression for prices (12).

The prices of Arrow-Debreu securities can be also expressed in terms of risk-neutral probabilities as $p_m = \pi^{\text{RN}}_m/(1+r_f)$. Because the risk-neutral probabilities sum up to one, from the latter equation we find that $(1+r_f) = 1/\sum_{m=1}^M p_m$. Substituting prices (12) into the latter equation, we find the riskless rate (14). Risk-neutral probabilities (13) are found by substituting prices (12) and the riskless rate (14) into the equation $\pi^{\text{RN}}_m = p_m(1+r_f)$. ■

Proof of Proposition 3. Suppose, there exist two vectors of shocks and beliefs $(\varepsilon, \nu_1, \ldots, \nu_N)$ and $(\tilde{\varepsilon}, \tilde{\nu}_1, \ldots, \tilde{\nu}_N)$ such that $\varepsilon \neq \tilde{\varepsilon}$ and the corresponding sufficient statistics are the same. Consequently, the equilibrium is not fully revealing, and the shocks and noises satisfy the following system of $M-1$ equations with $2N+2$ unknowns:

$$\begin{aligned} \mu_I \pi_{I,m}(\varepsilon) + (\mu_B/N)(\pi_{B,m}(\nu_1) + \cdots + \pi_{B,m}(\nu_N)) &= \\ \mu_I \pi_{I,m}(\tilde{\varepsilon}) + (\mu_B/N)(\pi_{B,m}(\tilde{\nu}_1) + \cdots + \pi_{B,m}(\tilde{\nu}_N)), & \end{aligned} \tag{A.4}$$

for $m = 1, \ldots, M-1$. Substituting probabilities $\pi_{I,m}(z)$ and $\pi_{B,m}(z)$ from (17) into the system (A.4), after cancelling $\exp(a(x_m))$ on both sides, we obtain the system of equations

$$\begin{aligned} \mu_I e^{\varepsilon x - c(\varepsilon)} + (\mu_B/N)(e^{\nu_1 x - c(\nu_1)} + \cdots + e^{\nu_N x - c(\nu_N)}) &= \\ \mu_I e^{\tilde{\varepsilon} x - c(\tilde{\varepsilon})} + (\mu_B/N)(e^{\tilde{\nu}_1 x - c(\tilde{\nu}_1)} + \cdots + e^{\tilde{\nu}_N x - c(\tilde{\nu}_N)}), & \end{aligned} \tag{A.5}$$

where $x = x_1, \ldots, x_{M-1}$. After fixing ε, $\tilde{\varepsilon}$, ν_n, $\tilde{\nu}_n$, (A.5) can be viewed as one equation with one unknown x that has $M-1$ solutions $x = x_1, \ldots, x_{M-1}$. However, Lemma A.1 on zeros of exponential sums below shows that, unless the left- and right-hand sides of (A.5) are identical for all x, this equation has at most $2N+1$ solutions, contradicting the condition $2N+1 < M-1$. Hence, system (A.4) either does not have a solution $(\varepsilon, \nu_1, \ldots, \nu_N)$ and $(\tilde{\varepsilon}, \tilde{\nu}_1, \ldots, \tilde{\nu}_N)$ or the left- and right-hand sides of (A.5) coincide for all x. It remains to exclude the latter possibility.

Suppose, the exponential sums on the left- and right-hand sides of (A.5) coincide for all x. Because the exponential functions are linearly independent, this means that the two sides of the equation contain the same exponents but possibly in different order. Suppose that $\exp(\varepsilon x - c(\varepsilon)) = \exp(\tilde{\varepsilon} x - c(\tilde{\varepsilon}))$ for all x. The latter equality implies that $\varepsilon = \tilde{\varepsilon}$, contradicting the initial assumption that $\varepsilon \neq \tilde{\varepsilon}$. Suppose, $\mu_I \exp(\varepsilon x - c(\varepsilon)) = (\mu_B n/N) \exp(\tilde{\nu}_k x - c(\tilde{\nu}_k))$ for all x and some k. Linear independence of exponents implies that $\varepsilon = \tilde{\nu}_k$, and hence, $\mu_I = \mu_B n/N$, contradicting the assumptions of Proposition 3. We note that if $\mu_I = \mu_B n/N$, the equilibrium might not be fully revealing when, for example, $\varepsilon = \tilde{\nu}_1 = \cdots = \tilde{\nu}_n$, $\tilde{\varepsilon} = \nu_1 = \cdots = \nu_n$, $\nu_{n+1} = \tilde{\nu}_{n+1}$, $\ldots$, $\nu_{N+1} = \tilde{\nu}_{N+1}$. However, the set of beliefs satisfying $\nu_1 = \nu_2$ for a non-revealing equilibrium when $n > 1$ has zero probability measure when beliefs are drawn from distributions with continuous support. ■

Lemma A.1 [Zeros of sums of exponents]. *Consider the following equation:*

$$A_1 e^{a_1 x} + A_2 e^{a_2 x} + \cdots + A_N e^{a_N x} = 0, \tag{A.6}$$

where A_n and a_n are some real numbers and at least one of the coefficients A_n is non-zero. Then, Equation (A.6) has at most $N-1$ distinct real roots.

Proof of Lemma A.1. This Lemma is Corollary 3.2 in Jameson (2006, p. 228). ■

Proposition A.1 [Fully revealing equilibrium with learning from prices]. *Suppose the uninformed investors learn about shock ε only from prices $P(\varepsilon, \nu)$ and both informed and uninformed investors have beliefs from the exponential family, given by (17). Then, the equilibrium is fully revealing if $2N + 2 < M$ and $\mu_I + \mu_U \neq \mu_B/N$.*

Proof of Proposition A.1. When the equilibrium is fully revealing, the beliefs of uninformed and informed investors coincide, so that $\pi_{U,m} = \pi_{I,m}(\varepsilon)$. Consequently, the prices of Arrow-Debreu securities (12) are given by

$$p_m = \beta \frac{(\mu_I + \mu_U)\pi_{I,m}(\varepsilon) + \mu_B(\pi_{B,m}(\nu_1) + \cdots + \pi_{B,m}(\nu_N))/N}{D_m/D_0}. \tag{A.7}$$

Next, we verify that prices (A.7) are fully revealing. Suppose, prices (A.7) are not fully revealing. Consequently, there exist two vectors $(\varepsilon, \nu_1, \ldots, \nu_N)$ and $(\tilde{\varepsilon}, \tilde{\nu}_1, \ldots, \tilde{\nu}_N)$ such that $\varepsilon \neq \tilde{\varepsilon}$ and $P(\varepsilon, \nu_1, \ldots, \nu_N) = P(\tilde{\varepsilon}, \tilde{\nu}_1, \ldots, \tilde{\nu}_N)$, where $P(\cdot)$ is the vector of prices. Shocks and noises then satisfy the following system of $M - 1$ equations with $2N + 2$ unknowns:

$$\begin{aligned} (\mu_I + \mu_U)\pi_{I,m}(\varepsilon) + (\mu_B/N)(\pi_{B,m}(\nu_1) + \cdots + \pi_{B,m}(\nu_N)) \quad &= \\ (\mu_I + \mu_U)\pi_{I,m}(\tilde{\varepsilon}) + (\mu_B/N)(\pi_{B,m}(\tilde{\nu}_1) + \cdots + \pi_{B,m}(\tilde{\nu}_N)), & \end{aligned} \tag{A.8}$$

for $m = 1, \ldots, M - 1$. System of equations (A.8) is similar to system (A.4) with the only difference being the coefficient in front of $\pi_{I,m}$. Therefore, the rest of the proof that the prices are fully revealing is analogous to the proof of Proposition 3. If prices are given by (A.7), the uninformed investors learns shock ε and update their probabilities to $\pi_n(\varepsilon)$. The latter probabilities lead to prices (A.7), and hence, these are the equilibrium prices. ■

Proposition A.2 [Optimal consumptions and portfolios of CARA investors]. *Suppose, the market is complete. Consider an investor who maximizes intertemporal expected CARA utility (18) subject to a self-financing budget constraint and invests in all available Arrow-Debreu securities. Then, the optimal consumptions C_0^* and $C_1(\omega_m)$ and numbers of units n_m^* of each security in the optimal portfolio are given by:*

$$C_0^* = \left(W_0 - \frac{\ln(\beta)}{\gamma(1 + r_f)} + \frac{\sum_{m=1}^{M} p_m \Big(\ln(p_m) - \ln(\pi_m) \Big)}{\gamma} \right) \frac{1 + r_f}{2 + r_f}, \tag{A.9}$$

$$C_1^*(\omega_m) = n_m^* = C_0^* + \frac{\ln(\pi_m) - \ln(p_m/\beta)}{\gamma}, \tag{A.10}$$

where p_m is the price of security m and π_m is the physical probability of state ω_m.

Proof of Proposition A.2. Because the market is complete, the investor's budget constraint is given in the following static form:

$$W_0 = C_0 + \mathbb{E}\Big[\frac{1}{1+r_f}\frac{\pi^{\text{RN}}(\omega)}{\pi(\omega)}C_1(\omega)\Big]. \tag{A.11}$$

Using the method of Lagrange multipliers, we find that the first order conditions for optimization (18) subject to the latter budget constraint are then given by

$$C_0^* = \tilde{\ell}, \tag{A.12}$$

$$C_1^*(\omega_m) = \tilde{\ell} + \frac{\ln(\beta)}{\gamma} + \frac{\ln(\pi_m) - \ln(p_m)}{\gamma}, \tag{A.13}$$

where $\tilde{\ell}$ is some function of the Lagrange multiplier. Substituting (A.12) into (A.13), gives the expression for the optimal consumption in (A.10). The value of the portfolio of Arrow-Debreu securities in state ω_m at date 1 is equal to the number of state-ω_m securities purchased at date-0, given by n_m^*. Moreover, the consumption at the terminal date should equal the portfolio value at that date, and hence, $C_1^*(\omega_m) = n_m^*$. Finally, we find constant $\tilde{\ell}$ by substituting consumptions (A.12) and (A.13) into the static budget constraint (A.11). ■

Proof of Proposition 4. For the case when investor beliefs are given by equations (17), the proof is presented in Subsection 4.1. More general version in which the beliefs of informed and uninformed investors have different belief functions $\pi_{I,m}(z) = \exp\{a_I(x_m) + zx_m - c_I(z)\}$ and $\pi_{B,m}(z) = \exp\{a_B(x_m) + zx_m - c_B(z)\}$ is established analogously. ■

Lemma A.2 [Convergence of conditional PDFs]. *Let $\widehat{s}_m = \eta\pi_m(\varepsilon) + \widehat{\xi}$, where $\widehat{\xi} \in \mathbb{R}^{M-1}$, where $\widehat{\xi}$ has PDF $\phi_N(y)$. Let $\phi_{N,\varepsilon|\widehat{s}}(x|y)$ and $\phi_{\varepsilon|\widehat{s}}(x|y)$ denote conditional PDFs of shock ε given signals $\widehat{s}$ when the PDFs of $\widehat{\xi}$ are given by $\phi_N(y)$ and $\phi(y)$, respectively, where $\phi(y)$ is normal PDF. Moreover, $\phi_N(y)$ converges to $\phi(y)$ uniformly, that is, $\phi_N(y) \rightrightarrows \phi(y)$. Then, for fixed y, the following convergence holds:*

$$\lim_{N\to\infty}\int_\Omega |\phi_{N,\varepsilon|\widehat{s}}(x|y) - \phi_{\varepsilon|\widehat{s}}(x|y)|dx = 0. \tag{A.14}$$

Proof of Lemma A.2. We will use the following notation throughout this proof:

$$\begin{aligned}
\tilde{\phi}_N(y,x) &\equiv \phi_N(y_1 - \eta\pi_{B,1}(x), \ldots, y_{M-1} - \eta\pi_{B,M-1}(x)),\\
\tilde{\phi}(y,x) &\equiv \phi(y_1 - \eta\pi_{B,1}(x), \ldots, y_{M-1} - \eta\pi_{B,M-1}(x)),\\
I_N(y) &= \textstyle\int_\Omega \phi_N(y_1 - \eta\pi_{B,1}(x), \ldots, y_{M-1} - \eta\pi_{B,M-1}(x))\varphi_\varepsilon(x)dx,\\
I(y) &= \textstyle\int_\Omega \phi(y_1 - \eta\pi_{B,1}(x), \ldots, y_{M-1} - \eta\pi_{B,M-1}(x))\varphi_\varepsilon(x)dx.
\end{aligned}$$

The conditional PDF of $\varepsilon|\widehat{s}$ is given by:

$$\phi_{N,\varepsilon|\widehat{s}}(x|y) = \frac{\tilde{\phi}_N(y,x)\varphi_\varepsilon(x)}{I_N(y)}. \tag{A.15}$$

Next, we prove the uniform convergence of denominators in (A.15).

By uniform convergence of PDFs, $\forall\epsilon > 0\ \exists N_0$ such that $\forall N \geq N_0$ we have

$$|\tilde{\phi}_N(y,x) - \tilde{\phi}(y,x))| < \epsilon,$$

for all x and y. From the latter inequality, by integrating, we find that $|I_N(y) - I(y)| < \epsilon \int \varphi_\varepsilon(x)dx = \epsilon$, which implies uniform convergence $I_N(y) \rightrightarrows I(y)$.

Because $\phi(y)$ is a bounded normal PDF and $\pi_{B,m}(x) \in [0,1]$, for a fixed vector y there exist bounds $A_1(y)$ and $A_2(y)$ such that

$$0 < A_1(y) < \tilde{\phi}(y,x) < A_2(y).$$

Multiplying the latter inequality by $\varphi_\varepsilon(x)$ and integrating, we find that $0 < A_1(y) < I(y) < A_2(y)$. Due to the uniform convergence of integrals, for a sufficiently large N and fixed y, we have: $0.5A_1(y) < I_N(y)$.

Then, for a fixed y consider the following integral:

$$\begin{aligned}
\int_\Omega |\phi_{N,\varepsilon|\widehat{s}}(x|y) - \phi_{\varepsilon|\widehat{s}}(x|y)|dx &= \int_\Omega \left|\frac{\tilde{\phi}_N(x,y)}{I_N(y)} - \frac{\tilde{\phi}(x,y)}{I(y)}\right| \varphi_\varepsilon(x)dx \\
&= \int_\Omega \left|\frac{\tilde{\phi}_N(x,y)I(y) - \tilde{\phi}(x,y)I_N(y)}{I_N(y)I(y)}\right| \varphi_\varepsilon(x)dx \\
&= \int_\Omega \left|\frac{(\tilde{\phi}_N(x,y) - \tilde{\phi}(x,y))I(y) - \tilde{\phi}(x,y)(I_N(y) - I(y))}{I_N(y)I(y)}\right| \varphi_\varepsilon(x)dx.
\end{aligned} \tag{A.16}$$

Now, pick arbitrarily small $\epsilon > 0$. Then, $\exists N_0$ such that $\forall N \geq N_0$ $|\tilde{\phi}_N(x,y) - \tilde{\phi}(x,y)| < \epsilon$ and $|I_N(y) - I(y)| < \epsilon$. Using the bounds for the densities and the integrals, we obtain the following bound on integral (A.16):

$$\int_\Omega |\phi_{N,\varepsilon|\widehat{s}}(x|y) - \phi_{\varepsilon|\widehat{s}}(x|y)|dx < \frac{4\epsilon A_2(y)}{A_1(y)^2}\int_\Omega \varphi_\varepsilon(x)dx = \frac{4\epsilon A_2(y)}{A_1(y)^2}.$$

The latter inequality implies the convergence (A.14) for a fixed y. ■

Lemma A.3 [Convergence of prices]. *Asset prices in the baseline economy converge point-wise to prices in the Gaussian-signal economy as functions of sufficient statistics s, when $N \to \infty$.*

Proof of Lemma A.3. Let $\widehat{\xi}_m = (\pi_{B,m}(\nu_1) - \overline{\mathbb{E}}[\pi_{B,m}(\nu)) + \cdots + (\pi_{B,m}(\nu_N) - \overline{\mathbb{E}}[\pi_{B,m}(\nu))/\sqrt{N}$ and $\widehat{\xi} = (\widehat{\xi}_1, \ldots, \widehat{\xi}_{M-1})^\top$. By $\phi_N(y)$ we denote the PDF function of vectors $\widehat{\xi}$, and by $\phi(y)$ we denote the corresponding multivariate normal PDF with zero mean and covariance matrix Σ_ξ. Because our random variables are probabilities in the interval $[0, 1]$ and are i.i.d., they have bounded moments, and hence, by the multivariate local limit theorem (e.g., Shervashidze, 1971) there is uniform convergence $\phi_N(y) \rightrightarrows \phi(y)$, as $N \to \infty$.

The PDF of ε conditional of observing signals $\widehat{s}_m = \eta\pi_{I,m}(\varepsilon) + \widehat{\xi}_m$ is given by:

$$\phi_{N,\varepsilon|\widehat{s}}(x|y) = \frac{\phi_N(y_1 - \eta\pi_{B,1}(x), \ldots, y_{M-1} - \eta\pi_{B,M-1}(x))\varphi_\varepsilon(x)}{\int \phi_N(y_1 - \eta\pi_{B,1}(x), \ldots, y_{M-1} - \eta\pi_{B,M-1}(x))\varphi_\varepsilon(x)dx}. \tag{A.17}$$

Next, we show that the posterior probabilities with finite N converge to those in the Gaussian-signal economy. These probabilities in the baseline economy are given by:

$$\mathbb{E}[\pi_{I,m}(\varepsilon)|\widehat{s} = y] = \int_\Omega \pi_{I,m}(x)\phi_{N,\varepsilon|\widehat{s}}(x|y)dx. \tag{A.18}$$

The convergence follows from the following inequality and Lemma A.2:

$$\left| \int_\Omega \pi_{I,m}(x)\phi_{N,\varepsilon|\widehat{s}}(x|y)dx - \int_\Omega \pi_{I,m}(x)\phi_{\varepsilon|\widehat{s}}(x|y)dx \right| \leq \int_\Omega \left|\phi_{N,\varepsilon|\widehat{s}}(x|y) - \phi_{\varepsilon|\widehat{s}}(x|y)\right| dx \xrightarrow[N\to\infty]{} 0.$$

Taking the limit $N \to \infty$ in the equation for prices (12), we then find that prices converge to those in the Gaussian-signal economy with posterior probabilities $\mathbb{E}[\pi_{I,m}(\varepsilon)|\widehat{s}] = \int_\Omega \pi_{I,m}(x)\phi_{\varepsilon|\widehat{s}}(x|y)dx$. ■

Lemma A.4. *Suppose probabilities $\pi_{B,m}(\nu)$ are linearly independent functions of ν for $m = 1, \ldots, M$. Consider two systems of linear algebraic equations:*

$$\Sigma_\xi \tilde{v} = \tilde{f}, \tag{A.19}$$

$$\mathcal{K}v = f, \tag{A.20}$$

where $\Sigma_\xi \in \mathbb{R}^{(M-1)\times(M-1)}$ has elements given by cross-sectional covariances $\overline{\text{cov}}(\pi_{B,m}(\nu), \pi_{B,k}(\nu))$ with respect to cross-sectional distribution of belief parameters $\varphi_\nu(x)$ for $m, k = 1, \ldots M-1$, $\mathcal{K} \in \mathbb{R}^{M\times M}$ is a larger matrix with elements $\overline{\mathbb{E}}[\pi_{B,m}(\nu)\pi_{B,k}(\nu)]$ for $m, k = 1, \ldots M$ also computed using cross-sectional distribution $\varphi_\nu(x)$. Vector $f = (f_1, \ldots, f_M)^\top \in \mathbb{R}^M$ is such that its components sum up to zero, $f_1 + \cdots + f_M = 0$, and vector $\tilde{f}$ is given by $\tilde{f} = (f_1, \ldots, f_{M-1})^\top \in \mathbb{R}^{M-1}$. Then, matrix $\mathcal{K}$ is nonsingular, positive-definite and symmetric, and each system (A.19) and (A.20) has unique solution, and these solutions satisfy the following equations:

$$\tilde{v} = (v_1 - v_M, v_2 - v_M, \ldots, v_{M-1} - v_M)^\top, \tag{A.21}$$

$$\tilde{v}^\top \tilde{f} = v^\top f. \tag{A.22}$$

Proof of Lemma A.4. First, we show that matrix Σ_ξ is invesrtible, and hence, equation (A.19) has unique solution. Matrix Σ_ξ is a symmetric Gram matrix. Suppose, this matrix is singular, and hence, there exists vector $\tilde{\ell}$ such that $\Sigma_\xi \tilde{\ell} = 0$. Consequently,

$$\tilde{\ell}^\top \Sigma_\xi \tilde{\ell} = \overline{\text{var}}(\tilde{\ell}_1 \pi_{B,1}(\nu) + \cdots + \tilde{\ell}_1 \pi_{B,M-1}(\nu)) = 0,$$

where $\overline{\text{var}}[\cdot]$ is variance with respect to beliefs ν. The latter equation implies that $\tilde{\ell}_1 \pi_{B,1}(\nu) + \cdots + \tilde{\ell}_1 \pi_{B,M-1}(\nu) := \gamma$ is constant. Because $\pi_{B,M}(\nu) = 1 - \pi_{B,1}(\nu) - \cdots - \pi_{B,M-1}(\nu)$, we can express it as $\pi_{B,M}(\nu) = (\tilde{\ell}_1/\gamma - 1)\pi_{B,1}(\nu) + \cdots + (\tilde{\ell}_{M-1}/\gamma - 1)\pi_{B,M-1}(\nu)$, contradicting linear independence of probabilities. Therefore, Σ_ξ is non-singular and positive-definite.

Matrix $\mathcal{K}$ is symmetric by construction. We prove that it is non-singular and positive definite. For all vectors $\ell \neq 0$, we have:

$$\ell^\top \mathcal{K} \ell = \overline{\mathbb{E}}[(\ell_1 \pi_{B,1}(\nu) + \cdots + \ell_M \pi_{B,M}(\nu))^2] > 0.$$

We note that $\ell^\top \mathcal{K} \ell = 0$ is not possible, because otherwise, $\ell_1 \pi_{B,1}(\nu) + \cdots + \ell_M \pi_{B,M}(\nu) = 0$, contradicting the linear independence of probabilities. Therefore, $\mathcal{K}$ is invertible and equation (A.20) has a unique solution.

Rewrite the system of equations (A.20) as follows:

$$\overline{\mathbb{E}}[\pi_{B,m}(\nu)\pi_{B,1}(\nu)]v_1 + \cdots + \overline{\mathbb{E}}[\pi_{B,m}(\nu)\pi_{B,M-1}(\nu)]v_{M-1} + \overline{\mathbb{E}}[\pi_{B,m}(\nu)\pi_{B,M}(\nu)]v_M = f_m, \quad \text{(A.23)}$$

for all $m = 1, \ldots, M$. Summing up all equations in (A.23) and using the fact that the probabilities sum up to one, and the components of vector f sum up to zero, we find that

$$\overline{\mathbb{E}}[\pi_{B,1}(\nu)]v_1 + \cdots + \overline{\mathbb{E}}[\pi_{B,M-1}(\nu)]v_{M-1} + \overline{\mathbb{E}}[\pi_{B,M}(\nu)]v_M = 0. \quad \text{(A.24)}$$

Substituting $\pi_{B,M}(\nu) = 1 - \pi_{B,1}(\nu) - \cdots - \pi_{B,M-1}(\nu)$ into (A.23) and (A.24), we obtain:

$$\overline{\mathbb{E}}[\pi_{B,m}(\nu)\pi_{B,1}(\nu)]\tilde{v}_1 + \cdots + \overline{\mathbb{E}}[\pi_{B,m}(\nu)\pi_{B,M-1}(\nu)]\tilde{v}_{M-1} + \overline{\mathbb{E}}[\pi_{B,m}(\nu)]v_M = f_m, \quad \text{(A.25)}$$

$$-\overline{\mathbb{E}}[\pi_{B,1}(\nu)]\tilde{v}_1 - \cdots - \overline{\mathbb{E}}[\pi_{B,M-1}(\nu)]\tilde{v}_{M-1} = v_M, \quad \text{(A.26)}$$

where $\tilde{v}_m := v_m - v_M$. Substituting v_M from (A.26) into (A.25), we obtain:

$$\overline{\text{cov}}(\pi_{B,m}(\nu), \pi_{B,1}(\nu))\tilde{v}_1 + \cdots + \overline{\text{cov}}(\pi_{B,m}(\nu), \pi_{B,M-1}(\nu))\tilde{v}_{M-1} = f_m, \quad \text{(A.27)}$$

where $m = 1, \ldots, M-1$. The latter system of equations is the same as system (A.19), which proves Equation (A.21).

Finally, we prove Equation (A.22). Substituting (A.21) into $\tilde{v}^\top \tilde{f}$, we obtain:

$$\begin{aligned} \tilde{v}^\top \tilde{f} &= v_1 f_1 + \cdots + v_{M-1} f_{M-1} - v_M(f_1 + \cdots + f_{M-1}) \\ &= v_1 f_1 + \cdots + v_{M-1} f_{M-1} + v_M f_M = v^\top f, \end{aligned}$$

where we used the fact that $f_M = -(f_1 + \cdots + f_{M-1})$. ■

Proof of Lemma 1. The vector of sufficient statistics is given by $\widehat{s} = \tilde{f} \cdot \varepsilon + \xi$, where $\xi \sim \mathcal{N}(0, \Sigma_\xi)$ and $\tilde{f} \in \mathbb{R}^{M-1}$ is a vector with elements $f_m = \eta(\pi_{I,m}(\varepsilon_H) - \pi_{I,m}(\varepsilon_L))$, and ε takes two values, $\varepsilon_L = 0$ and $\varepsilon_H = 1$. Using Bayes's Theorem, taking into account that $\widehat{s}$ are conditionally normally distributed, after simple algebra, we obtain:

$$\begin{aligned}\widehat{\pi}_L(\widehat{s}) := \text{Prob}(\varepsilon = \varepsilon_L|\widehat{s}) &= \frac{\exp(-0.5\widehat{s}^\top \Sigma_\xi^{-1}\widehat{s})\varphi_L}{\exp(-0.5\widehat{s}^\top \Sigma_\xi^{-1}\widehat{s})\varphi_L + \exp(-0.5(\widehat{s}-\tilde{f})^\top \Sigma_\xi^{-1}(\widehat{s}-\tilde{f}))\varphi_H} \\ &= \frac{1}{1 + (\varphi_H/\varphi_L)\exp(-0.5\tilde{f}^\top \Sigma_\xi^{-1}\tilde{f} + \tilde{f}^\top \Sigma_\xi^{-1}\widehat{s})}.\end{aligned} \tag{A.28}$$

Next, we find $\mathbb{E}[\widehat{\pi}_L(\widehat{s})|\varepsilon = \varepsilon_L]$. From the structure of the signal, we observe that the distribution of $\widehat{s}$ conditional on knowing that $\varepsilon = \varepsilon_L$ is $\mathcal{N}(0, \Sigma_\xi)$. Consequently, $\tilde{f}^\top \Sigma_\xi^{-1}\widehat{s} = x\sqrt{\tilde{f}^\top \Sigma_\xi^{-1}\tilde{f}}$, where $x \sim \mathcal{N}(0,1)$. Using the latter change of variable and (A.28), we obtain:

$$\mathbb{E}[\widehat{\pi}_L(\widehat{s})|\varepsilon = \varepsilon_L] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \frac{\exp(-0.5x^2)}{1 + (\varphi_H/\varphi_L)\exp(-0.5\tilde{f}^\top \Sigma_\xi^{-1}\tilde{f} + x\sqrt{\tilde{f}^\top \Sigma_\xi^{-1}\tilde{f}})}dx.$$

Denoting $J_M^* = \tilde{f}^\top \Sigma_\xi^{-1}\tilde{f}$, we find that $\mathbb{E}[\widehat{\pi}_L(\widehat{s})|\varepsilon = \varepsilon_L]$ is given by Equation (25) in which function $g(y)$ is given by

$$g(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \frac{\exp(-0.5x^2)}{1 + (\varphi_H/\varphi_L)\exp(-0.5y^2 + xy)}dx. \tag{A.29}$$

We now show that $J_M^* := \tilde{f}^\top \Sigma_\xi^{-1}\tilde{f} = f^\top \mathcal{K}^{-1} f$. Let $\tilde{v} := \Sigma_\xi^{-1}\tilde{f}$, and hence, $\Sigma_\xi \tilde{v} = \tilde{f}$. By Lemma A.4, $\tilde{v}^\top \tilde{f} = v^\top f$, where v solves $\mathcal{K}v = f$. Therefore, $J_M^* = f^\top \mathcal{K}^{-1} f$.

Next, we show that $g(y) \to 1$, as $y \to \infty$. Consider the following equation:

$$\begin{aligned} g(y) - 1 &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \frac{\exp(-0.5x^2)}{1 + (\varphi_H/\varphi_L)\exp(-0.5y^2 + yx)} - \exp(-0.5x^2)dx \\ &= -\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \frac{\exp(-0.5x^2)}{1 + (\varphi_L/\varphi_H)\exp(0.5y^2 - yx)}dx. \end{aligned} \tag{A.30}$$

Denote the function under the latter integral by $h_0(x; y)$. There is uniform convergence $h_0(x; y) \rightrightarrows 0$ as $y \to \infty$ for all $x \in [a, b]$, where $[a, b] \in \mathbb{R}$ is an arbitrary finite interval. Moreover, $h_0(x; y) < \exp(-0.5x^2)$, and hence, the improper integral (A.30) converges uniformly by Weierstrass convergence test (Zorich, 2004, p. 420). Consequently, we can take limit under the integral (A.30) (Zorich, 2004, p. 423), and hence, $g(y) - 1 \to 0$ as $y \to \infty$.

Denote the function under the integral in (A.29) by $h(x;y)$. Differentiating function $h(x;y)$ with respect to y, we obtain:

$$\begin{aligned}\frac{\partial h(x,y)}{\partial y} &= \frac{\varphi_H}{\varphi_L}\frac{\exp(-0.5x^2-0.5y^2+yx)}{(1+(\varphi_H/\varphi_L)\exp(-0.5y^2+yx))^2}(y-x)\\ &= \frac{\varphi_H}{\varphi_L}\frac{\exp(-0.5(y-x)^2)}{(1+(\varphi_H/\varphi_L)\exp(-0.5(y-x)^2+0.5x^2))^2}(y-x).\end{aligned} \tag{A.31}$$

Then, we have the following uniform bound on the partial derivative in (A.31) with respect to $y \in [0,d]$, where $d>0$ is an arbitrary number:

$$\left|\frac{\partial h(x,y)}{\partial y}\right| \leq \frac{\varphi_H}{\varphi_L}\exp(-0.5x^2+d|x|)(|x|+d)),$$

The integral $\int_{-\infty}^{\infty}\exp(-0.5x^2+d|x|)(|x|+d)dx$ converges and is finite, and hence, by Weierstrass convergence test (Zorich, 2004, p. 420), the integral of $\partial h(x,y)/\partial y$ converges uniformly for all $y\in[0,d]$. Therefore, the derivative of function $g(y)$ in (A.29) is equal to the integral of the partial derivative $\partial h(x,y)/\partial y$ (Zorich, 2004, p. 426), given by:

$$g'(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{\varphi_H}{\varphi_L}\frac{\exp(-0.5(y-x)^2)}{(1+(\varphi_H/\varphi_L)\exp(-0.5(y-x)^2+0.5x^2))^2}(y-x)dx.$$

In the latter integral, we change the variable of integration from x to $t=y-x$ and split the integral in two, as follows:

$$\begin{aligned}g'(y) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{\varphi_H}{\varphi_L}\frac{\exp(-0.5t^2)t}{(1+(\varphi_H/\varphi_L)\exp(-0.5t^2+0.5(t-y)^2))^2}dt\\ &= \frac{1}{\sqrt{2\pi}}\int_{0}^{\infty}\frac{\varphi_H}{\varphi_L}\frac{\exp(-0.5t^2)t}{(1+(\varphi_H/\varphi_L)\exp(-0.5t^2+0.5(t-y)^2))^2}dt\\ &\quad -\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0}\frac{\varphi_H}{\varphi_L}\frac{\exp(-0.5t^2)t}{(1+(\varphi_H/\varphi_L)\exp(-0.5t^2+0.5(t-y)^2))^2}dt\\ &= \frac{1}{\sqrt{2\pi}}\int_{0}^{\infty}\frac{\varphi_H}{\varphi_L}\frac{\exp(-0.5t^2)t}{(1+(\varphi_H/\varphi_L)\exp(-0.5t^2+0.5(t-y)^2))^2}dt\\ &\quad -\frac{1}{\sqrt{2\pi}}\int_{0}^{\infty}\frac{\varphi_H}{\varphi_L}\frac{\exp(-0.5t^2)t}{(1+(\varphi_H/\varphi_L)\exp(-0.5t^2+0.5(t+y)^2))^2}dt,\end{aligned} \tag{A.32}$$

where the last integral is obtained by another change of variable, from t to $-t$. Inequality $g'(y)>0$ then follows from the fact that the difference of functions under the corresponding integrals is positive due to the fact that for all $t>0$ and $y\geq 0$ we have that $\exp(-0.5t^2+0.5(t-y)^2) < \exp(-0.5t^2+0.5(t+y)^2)$. ■

Proof of Lemma 2. Matrix $\mathcal{K}$ is nonsingular and positive definite by Lemma A.4. Hence, it has M eigenvalues $\lambda_m > 0$ and eigenvectors ψ_m that form an orthonormal basis in $\mathbb{R}^M$. It can be directly verified that the solution of the regularized equation (27) is given by

$$v_\tau = \sum_{m=1}^{M} \frac{f_m}{\lambda_m + \tau} \psi_m,$$

where $f_m = (f, \psi_m)$ are inner products. Moreover, $f = \sum_{m=1}^{M} f_m \psi_m$. Therefore, the regularized measure of efficiency $J_{\scriptscriptstyle M}(\tau) = f^\top v_\tau$ is then given by

$$J_{\scriptscriptstyle M}(\tau) = \sum_{m=1}^{M} \frac{f_m^2}{\lambda_m + \tau}.$$

Function $J_{\scriptscriptstyle M}(\tau)$ is decreasing and continuous for all $\tau \geq 0$ because $\lambda_m > 0$ for all m. ■

Proof of Proposition 5.

1) The kernel $K(x, y) = \overline{\mathbb{E}}[\pi_{\scriptscriptstyle B}(x; \nu)\pi_{\scriptscriptstyle B}(y; \nu)]$ of Equation (30) with $\nu \sim \mathcal{N}(0, \sigma_\nu^2)$ is given by:

$$K(x, y) = \frac{1}{2\pi\sigma_{\scriptscriptstyle B}^2\sqrt{2\pi\sigma_\nu^2}} \int_{-\infty}^{\infty} \exp\Big(-0.5(x-\nu)^2/\sigma_{\scriptscriptstyle B}^2 - 0.5(y-\nu)^2/\sigma_{\scriptscriptstyle B}^2 - 0.5\nu^2/\sigma_\nu^2\Big) d\nu. \quad \text{(A.33)}$$

The argument of the exponent in (A.33) can be simplified as follows:

$$\frac{(x-\nu)^2 + (y-\nu)^2}{\sigma_{\scriptscriptstyle B}^2} + \frac{\nu^2}{\sigma_\nu^2} = \frac{\widehat{\sigma}^2}{\sigma_{\scriptscriptstyle B}^4}(x-y)^2 + \frac{\widehat{\sigma}^2}{\sigma_{\scriptscriptstyle B}^2\sigma_\nu^2}(x^2+y^2) + \frac{(\nu - \widehat{\mu}_\nu)^2}{\widehat{\sigma}^2},$$

where $\widehat{\mu}_\nu$ does not depend on ν. Substituting the latter expression into (A.33) and integrating, we find

$$\begin{aligned} K(x, y) &= \frac{1}{2\pi\sigma_{\scriptscriptstyle B}^2\sqrt{2\pi\sigma_\nu^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{\widehat{\sigma}^2}{2\sigma_{\scriptscriptstyle B}^4}(x-y)^2 - \frac{\widehat{\sigma}^2}{2\sigma_{\scriptscriptstyle B}^2\sigma_\nu^2}(x^2+y^2) - \frac{(\nu-\widehat{\mu}_\nu)^2}{2\widehat{\sigma}^2}\right) d\nu, \\ &= \frac{\widehat{\sigma}}{2\pi\sigma_{\scriptscriptstyle B}^2\sigma_\nu} \exp\left(-\frac{\widehat{\sigma}^2}{2\sigma_{\scriptscriptstyle B}^4}(x-y)^2 - \frac{\widehat{\sigma}^2}{2\sigma_{\scriptscriptstyle B}^2\sigma_\nu^2}(x^2+y^2)\right). \end{aligned} \quad \text{(A.34)}$$

2) Functions (33) are orthogonal Hermite functions:

$$\begin{aligned} \int_{-\infty}^{\infty} \psi_n(x)\psi_k(x)dx &= \frac{a}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp(-a^2x^2) \frac{H_n(ax)H_k(ax)}{\sqrt{2^n n!}\sqrt{2^k k!}} dx \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \exp(-x^2) \frac{H_n(x)H_k(x)}{\sqrt{2^n n!}\sqrt{2^k k!}} dx. \end{aligned} \quad \text{(A.35)}$$

The latter integral is equal to 1 when $n = k$ and zero otherwise (Lebedev, 1972, pp. 65-66).

Next, we find eigenfunctions and the eigenvalues of the integral operator using the following Mehler's formula (Lebedev, 1972, p. 61), slightly rewritten in a convenient form:

$$\sum_{n=0}^{\infty} \frac{H_n(x)H_n(y)}{n!}\left(\frac{u}{2}\right)^n = \frac{1}{\sqrt{1-u^2}}\exp\left\{\frac{u}{1+u}(x^2+y^2) - \frac{u}{1-u^2}(x-y)^2\right\}, \tag{A.36}$$

where $H_n(x)$ are Hermite polynomials, and the series on the left-hand side converges if and only if $|u| < 1$. We rewrite formula (A.36) for arguments ax and ay, express Hermite polynomials via eigenfunctions (33), multiply both sides by λ_0, and take $u = (\sigma_\nu - \widehat{\sigma})/(\sigma_\nu + \widehat{\sigma})$. We note that because $\widehat{\sigma} = (2/\sigma_B^2 + 1/\sigma_\nu^2)^{-1/2} < \sigma_\nu$, we have $0 < u < 1$. Then, using some algebra, we find that

$$\sum_{n=0}^{\infty} \lambda_n \psi_n(x)\psi_n(y) = K(x,y), \tag{A.37}$$

where $K(x,y)$ is given by (A.34). Due to Cramer's inequality $|\psi_n(x)| < 1.1\sqrt{\pi}$ (Gradshteyn and Ryzhik, 2007, Eq. 8.954.2) and the fact that λ_n is a decaying exponential function given by (34), the series on the left-hand side of (A.37) converges uniformly, and hence, summation and integration are interchangeable. Multiplying Equation (A.37) by $\psi_k(y)$, integrating, and using the orthogonality of Hermite functions, we find that

$$\int_{-\infty}^{\infty} K(x,y)\psi_k(y)dy = \lambda_k \psi_k(x),$$

and hence, $\psi_n(x)$ and λ_n are the eigenfunctions and eigenvalues of the integral operator $\mathcal{K}$. Functions $\psi_n(x)$ form an orthonormal basis in $\mathcal{L}_2(R)$ (Johnston, 2014).

3) Operator $\mathcal{K}$ is positive-definite by construction. Moreover, $K(x,y) \in \mathcal{L}_2(R^2)$ because

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K(x,y)^2 dxdy < \frac{\widehat{\sigma}^2}{4\pi^2\sigma_B^4\sigma_\nu^2}\left(\int_{-\infty}^{\infty}\exp\left(-\frac{\widehat{\sigma}^2}{\sigma_B^2\sigma_\nu^2}x^2\right)dx\right)^2 < \infty.$$

Consequently, operator $\mathcal{K}$ is completely continuous (Kolmogorov and Fomin, 1999, p. 120) and is self adjoint because $K(x,y) = K(y,x)$. Moreover, because the operator is positive definite, the operator equation has only zero solution when $f(x) = 0$. Consequently, the solution of the integral equation (30) exits and is unique by Fredholm's theorem (Kolmogorov and Fomin, 1999, pp. 117-118). The solution of this equation can then be expanded into series (35), where f_n are Fourier coefficients given by:

$$f_n = \frac{\eta}{\sqrt{2\pi\sigma_I^2}}\int_{-\infty}^{\infty}\left(\exp(-0.5(x-\mu)^2/\sigma_I^2) - \exp(-0.5x^2/\sigma_I^2)\right)\psi_n(x)dx. \tag{A.38}$$

We evaluate the latter integral using equation 7.374.10 in Gradshteyn and Ryzhik (2007), and after additional algebra, we find that[4]

$$f_n = \frac{\eta\sqrt{a}}{\sqrt{1+a^2\sigma_I^2}}\left(\frac{1-a^2\sigma_I^2}{1+a^2\sigma_I^2}\right)^{n/2}\left(\exp\left(-\frac{0.5\mu^2a^2}{1+a^2\sigma_I^2}\right)\frac{H_n(a\mu/\sqrt{1-a^4\sigma_I^4})}{(2^n n!\sqrt{\pi})^{1/2}} - \frac{H_n(0)}{(2^n n!\sqrt{\pi})^{1/2}}\right). \tag{A.39}$$

To prove the convergence of series (35) in $L_2(R)$, using the orthonormality of eigenfunctions $\psi_n(x)$ we find that

$$\left\|\sum_{n=m}^{m+k}\frac{f_n\psi_n(x)}{\tau+\lambda_n}\right\|_{\mathcal{L}_2}^2 = \sum_{n=m}^{m+k}\frac{f_n^2}{(\tau+\lambda_n)^2} < \frac{1}{\tau^2}\sum_{n=m}^{m+k}f_n^2,$$

The convergence then follows from Cauchy's convergence test and the convergence of $\sum_{n=0}^{+\infty}f_n^2$, which in turn follows from the fact that $f(x)\in\mathcal{L}_2(R)$, and hence, by Parseval's identity $\sum_{n=0}^{\infty}f_n^2 = \|f(x)\|_{\mathcal{L}_2}^2$. Applying Parseval's identity once more, we obtain the following expression for the measure of informational efficiency:

$$J_\infty(\tau) := \int_{-\infty}^{\infty} v(x;\tau)f(x)dx = \sum_{n=0}^{\infty}\frac{f_n^2}{\tau+\lambda_n}. \tag{A.40}$$

The convergence of the latter series follows from the convergence of $\sum_{n=0}^{+\infty}f_n^2$, demonstrated above, and inequality $\sum_{n=0}^{\infty}f_n^2/(\tau+\lambda_n) < (1/\tau)\sum_{n=0}^{\infty}f_n^2$. ■

Proof of Proposition 6. We take limit $\tau\to 0$ in $J_\infty(\tau)$, given by (A.40), and show that the limit exists and equals $J_\infty(0)$ when condition (37) is satisfied. Substituting eigenvalues (34) and Fourier coefficients (A.39) into Equation (A.40), and setting $\tau=0$, we obtain:

$$\begin{aligned}
J_\infty(0) &= \eta^2\frac{\widehat{\sigma}+\sigma_\nu}{\sigma_\nu}\sqrt{\frac{\sigma_B^2 a}{1+a^2\sigma_I^2}}\sum_{n=0}^{\infty}\left(\frac{\rho}{2}\right)^n\left(\exp\left(-\frac{0.5\mu^2a^2}{1+a^2\sigma_I^2}\right)\frac{H_n(y)}{\sqrt{2^n n!}} - \frac{H_n(0)}{\sqrt{2^n n!}}\right)^2. \\
&= \eta^2\frac{\widehat{\sigma}+\sigma_\nu}{\sigma_\nu}\sqrt{\frac{\sigma_B^2 a}{1+a^2\sigma_I^2}}\left[\exp\left(-\frac{\mu^2a^2}{1+a^2\sigma_I^2}\right)\sum_{n=0}^{\infty}\left(\frac{\rho}{2}\right)^n\frac{H_n(y)^2}{2^n n!}\right. \\
&\quad \left. -2\exp\left(-\frac{0.5\mu^2a^2}{1+a^2\sigma_I^2}\right)\sum_{n=0}^{\infty}\left(\frac{\rho}{2}\right)^n\frac{H_n(y)H_n(0)}{2^n n!} + \sum_{n=0}^{\infty}\left(\frac{\rho}{2}\right)^n\frac{H_n(0)^2}{2^n n!}\right],
\end{aligned} \tag{A.41}$$

[4]Using the recursive formula for Hermite polynomials, it can be shown that Fourier coefficients f_n satisfy a recursive equation: $(1+a^2\sigma_I^2)\sqrt{n}f_n = a\mu\sqrt{2}f_{n-1} - (1-a^2\sigma_I^2)\sqrt{n-1}f_{n-2}$, where $f_0 = \eta\sqrt{a/(1+a^2\sigma_I^2)}\pi^{-1/4}\exp\{-0.5\mu^2a^2/(1+a^2\sigma_I^2)\}$, and $f_1 = \eta\sqrt{2}\pi^{-1/4}\mu(a/(1+a^2\sigma_I^2)^{3/2}\exp\{-0.5\mu^2a^2/(1+a^2\sigma_I^2)\}$. The latter recursive equation is more convenient for computing Fourier coefficients f_n than the closed-form expression (A.39) because it avoids loss of precision.

where

$$\rho = \frac{\sigma_\nu + \widehat{\sigma}}{\sigma_\nu - \widehat{\sigma}} \cdot \frac{1 - a^2\sigma_I^2}{1 + a^2\sigma_I^2}, \quad y = \frac{a\mu}{\sqrt{1 - a^4\sigma_I^4}}. \tag{A.42}$$

The three sums in (A.41) can be computed using Mehler's formula (A.36) when $|\rho| < 1$. Using some algebra, we verify that the condition $|\rho| < 1$ is equivalent to inequality (37). When the latter inequality is satisfied, the measure of informational efficiency $J_\infty^* = J_\infty(0)$ is finite and is given by (38). Hence, the equilibrium is only partially revealing.

It then remains to verify that $J_\infty(0) = +\infty$ when $|\rho| \geq 1$, and hence, the equilibrium is fully revealing. Using the fact that $H_n(0) = 0$ for odd n (Lebedev, 1972, p. 61, Eq. 4.9.4), from the first equality in (A.41), we obtain:

$$\begin{aligned} J_\infty(0) &= A \sum_{n=0}^{\infty} \left(\frac{\rho}{2}\right)^n \left(\exp\left(-\frac{0.5\mu^2a^2}{1 + a^2\sigma_I^2}\right) \frac{H_n(y)}{\sqrt{2^n n!}} - \frac{H_n(0)}{\sqrt{2^n n!}}\right)^2 \\ &> A \exp\left(-\frac{\mu^2a^2}{1 + a^2\sigma_I^2}\right) \sum_{n=0}^{\infty} \left(\frac{\rho}{2}\right)^{2n+1} \frac{H_{2n+1}(y)^2}{2^{2n+1}(2n+1)!}, \end{aligned} \tag{A.43}$$

where A is a constant. We note that $\rho > 0$ and y is real when $1 - a^2\sigma_I^2 > 0$ and $\rho < 0$ and $y = i\tilde{y}$ is imaginary when $1 - a^2\sigma_I^2 < 0$, where $\tilde{y} > 0$. We need to encompass both of the latter cases. Note that polynomials $H_n(y)$ have only odd powers y^{2k+1} when n is odd and only even powers y^{2k} when n is even. Consider the case $1 - a^2\sigma_I^2 < 0$. The closed-form formula for the $(2n+1)$-th Hermite polynomial (Lebedev, 1972, p. 60, Eq. 4.92)

$$H_{2n+1}(y) = (2n+1)! \sum_{l=0}^{n} \frac{(-1)^l (2y)^{2n+1-2l}}{l!(2n+1-2l)!},$$

implies that for a negative $\rho < 0$ and $y = i\tilde{y}$, we have

$$\rho H_{2n+1}(i\tilde{y})^2 = |\rho| \left((2n+1)! \sum_{l=0}^{n} \frac{(2|y|)^{2n+1-2l}}{l!(2n+1-2l)!}\right)^2 \geq |\rho| H_{2n+1}(|y|)^2.$$

The latter inequality and inequality (A.43) jointly imply that

$$J_\infty(0) > A \exp\left(-\frac{\mu^2a^2}{1 + a^2\sigma_I^2}\right) \sum_{n=0}^{\infty} \left(\frac{|\rho|}{2}\right)^{2n+1} \frac{H_{2n+1}(|y|)^2}{2^{2n+1}(2n+1)!}, \tag{A.44}$$

where $|\rho| \geq 1$. Now, pick arbitrary $\tilde{\rho}$ such that $|\rho| \geq 1 > \tilde{\rho} > 0$. Inequality (A.44) implies:

$$J_\infty(0) > A \exp\left(-\frac{\mu^2a^2}{1 + a^2\sigma_I^2}\right) \sum_{n=0}^{\infty} \left(\frac{\tilde{\rho}}{2}\right)^{2n+1} \frac{H_{2n+1}(|y|)^2}{2^{2n+1}(2n+1)!}. \tag{A.45}$$

Evaluating Mehler's identity (A.36) at $u = \tilde{\rho}$ and $u = -\tilde{\rho}$, setting $x = y = |y|$ and taking the difference, we find that

$$\sum_{n=0}^{\infty} \left(\frac{\tilde{\rho}}{2}\right)^{2n+1} \frac{H_{2n+1}(|y|)^2}{2^{2n+1}(2n+1)!} = \frac{1}{2\sqrt{1-\tilde{\rho}^2}} \left(\exp\left\{ \frac{2\tilde{\rho}}{1+\tilde{\rho}} |y|^2 \right\} - \exp\left\{ -\frac{2\tilde{\rho}}{1-\tilde{\rho}} |y|^2 \right\} \right).$$

The latter identity and inequality (A.45) imply that $J_\infty(0) = +\infty$ because its lower bound converges to infinity when $\tilde{\rho} \to 1 - 0$, and hence, the equilibrium is fully revealing. The proof for the case when $\rho > 1$ is analogous. ■

Proof of Proposition 7. The interest rate is given by Equation (14). Let the state of the economy be $x_m = \ln(D_m/D_0)$. Then, as the number of states converges to infinity, $M \to \infty$, the equation for the riskless rate becomes

$$r_f = \frac{1/\beta}{\int_{-\infty}^{\infty} \bar{\pi}(x;\widehat{s}) e^{-x} dx} - 1, \tag{A.46}$$

where $\bar{\pi}(x;\widehat{s}) = \mu_B \overline{\mathbb{E}}[\pi_B(x;\nu)] + \mu_U \mathbb{E}[\pi_I(x;\varepsilon)|\widehat{s}]$ is the average probability across all investors, $\overline{\mathbb{E}}[\cdot]$ is cross-sectional probability with respect to belief parameter ν, and $\pi_B(x;\nu)$ and $\pi_I(x;\varepsilon)$ are normal probability densities given by Equations (31).[5] Because ε is a binary variable taking values $\varepsilon_H = 1$ and $\varepsilon_L = 0$, it is easy to see that $\pi_I(x;\varepsilon) = \pi_I(x;1)\varepsilon + \pi_I(x;0)(1-\varepsilon)$, which allows us to rewrite the average probability as follows:

$$\bar{\pi}(x;\widehat{s}) = \mu_B \overline{\mathbb{E}}[\pi_B(x;\nu)] + \mu_U [\pi_I(x;0)\widehat{\pi}_L(\widehat{s}) + \pi_I(x;1)(1-\widehat{\pi}_L(\widehat{s}))], \tag{A.47}$$

where $\pi_L(\widehat{s}) = \text{Prob}(\varepsilon = \varepsilon_L|\widehat{s})$ and is given by Equation (A.28). Next, we compute

$$\begin{aligned} \overline{\mathbb{E}}[\pi_B(x;\nu)] &= \frac{1}{2\pi\sigma_B\sigma_\nu} \int_{-\infty}^{\infty} \exp\left\{ -\frac{(x-\nu)^2}{2\sigma_B^2} - \frac{\nu^2}{2\sigma_\nu^2} \right\} d\nu \\ &= \frac{1}{\sqrt{2\pi(\sigma_\nu^2+\sigma_B^2)}} \exp\left\{ -\frac{x^2}{2(\sigma_\nu^2+\sigma_B^2)} \right\}. \end{aligned} \tag{A.48}$$

Consequently, the average probability (A.49) can be rewritten as follows:

$$\begin{aligned} \bar{\pi}(x;\widehat{s}) &= \frac{\mu_B}{\sqrt{2\pi(\sigma_\nu^2+\sigma_B^2)}} \exp\left\{ -\frac{x^2}{2(\sigma_\nu^2+\sigma_B^2)} \right\} + \frac{\mu_U}{\sqrt{2\pi\sigma_I^2}} \exp\left\{ -\frac{x^2}{2\sigma_I^2} \right\} \widehat{\pi}_L(\widehat{s}) \\ &\quad + \frac{\mu_U}{\sqrt{2\pi\sigma_I^2}} \exp\left\{ -\frac{(x-\mu)^2}{2\sigma_I^2} \right\} (1-\widehat{\pi}_L(\widehat{s})). \end{aligned} \tag{A.49}$$

[5] The mass of informed investors $\mu_I = \widehat{\mu}_I/\sqrt{N}$ does not enter the equation for the average probability because μ_I is asymptotically zero given that the number of biased investors N is large.

Next, we compute the following integral:

$$\int_{-\infty}^{\infty} \bar{\pi}(x;\widehat{s})e^{-x}dx = \mu_B \exp\{0.5(\sigma_\nu^2+\sigma_B^2)\}+\mu_U\Big[\exp\{0.5\sigma_I^2\}\widehat{\pi}_L(\widehat{s})+\exp\{-\mu+0.5\sigma_I^2\}(1-\widehat{\pi}_L(\widehat{s}))\Big].$$

The latter integral is computed noting that probability (A.49) is a weighted sum of normal distributions and that for normally distributed variables x we have $\mathbb{E}[\exp\{-x\}] = \exp\{-\mathbb{E}[x] + 0.5\,\mathrm{var}[x]\}$. Substituting the latter integral into (A.46), we obtain the riskless rate (40). ■

Lemma A.5. *Consider a matrix $\mathcal{K} \in \mathbb{R}^{M\times M}$ given by*

$$\mathcal{K} = \sum_{m=1}^{M} \tilde{\pi}_m \Pi_m \Pi_m^\top, \tag{A.50}$$

where vectors $\Pi_m \in \mathbb{R}^M$ are linearly independent and $\tilde{\pi}_m \neq 0$. Then, equation $\mathcal{K}v_i = \Pi_i$, where $i \in \{1,\dots,M\}$, has unique solution v_i, such that $\Pi_m^\top v_i = 0$ if $m \neq i$ and $\Pi_i^\top v_i = 1/\tilde{\pi}_i$.

Proof of Lemma A.5. Fist, we show that matrix $\mathcal{K}$ is invertible. Suppose, the matrix is singular. Then, there exists vector v such that $\mathcal{K}v = 0$. Substituting matrix $\mathcal{K}$ from (A.50) into the latter equation, we find that

$$\sum_{m=1}^{M} (\tilde{\pi}_m \Pi_m^\top v)\Pi_m = 0.$$

Linear independence of vectors Π_m then implies that $\Pi_m^\top v = 0$ for all m. The latter identities imply that $v = 0$ because vectors Π_m form a basis in $\mathbb{R}^M$, and hence, $\mathcal{K}$ is invertible. Consequently, equations $\mathcal{K}v_i = \Pi_i$ have unique solutions for all i. Substituting matrix $\mathcal{K}$ from (A.50) into the latter equation, we find that

$$\sum_{m=1}^{M} (\tilde{\pi}_m \Pi_m^\top v_i)\Pi_m = \Pi_i.$$

Linear independence of Π_m then implies that $\Pi_m^\top v_i = 0$ for all $m \neq i$ and $\Pi_i^\top v_i = 1/\tilde{\pi}_i$. ■

Proof of Proposition 8. As shown in Lemma 1, the measure of informational efficiency is given by $J_M^* = f^\top \mathcal{K}^{-1} f$, where f has elements $f_m = \eta(\pi_{I,m}(\varepsilon_H) - \pi_{I,m}(\varepsilon_L))$. Consequently, $f = \eta(\Pi_I(\varepsilon_H) - \Pi_I(\varepsilon_L))$, where $\Pi_I(\varepsilon)$ is a vector of informed investor's probabilities when shock is ε. Moreover, matrix $\mathcal{K}$ in Lemma 1 can be represented as $\mathcal{K} = \overline{\mathbb{E}}[\Pi\Pi^\top]$, which can be further expanded as follows:

$$\mathcal{K} = \sum_{m=1}^{M} \tilde{\pi}_m \Pi_m \Pi_m^\top. \tag{A.51}$$

Finding J_M^* then reduces to solving equation $\mathcal{K}v = f$. Using expansion (41), we rewrite the latter equation as follows:

$$\mathcal{K}v = \eta[(d_1(\varepsilon_H) - d_1(\varepsilon_L))\Pi_1 + \cdots + (d_M(\varepsilon_H) - d_M(\varepsilon_L))\Pi_M]. \tag{A.52}$$

Let v_i be the solution of equation $\mathcal{K}v_i = \Pi_i$. Then, the solution of (A.52) is given by

$$v = \eta[(d_1(\varepsilon_H) - d_1(\varepsilon_L))v_1 + \cdots + (d_M(\varepsilon_H) - d_M(\varepsilon_L))v_M].$$

The measure of informational efficiency is then given by $J_M^* = \eta(\Pi_I(\varepsilon_H) - \Pi_I(\varepsilon_L))^\top v$, which we expand as follows:

$$J_M^* = \eta^2 \left[\sum_{m=1}^{M}\Big(d_m(\varepsilon_H) - d_m(\varepsilon_L)\Big)\Pi_m\right]^\top \left[\sum_{m=1}^{M}\Big(d_m(\varepsilon_H) - d_m(\varepsilon_L)\Big)v_m\right]. \tag{A.53}$$

By Lemma A.5, $\Pi_m^\top v_i = 0$ if $m \neq i$ and $\Pi_i^\top v_i = 1/\tilde{\pi}_i$. Using the latter fact, we expand brackets in Equation (A.53) and simplify it to the expression given by Equation (42). ■

References

Admati, A. R., 1985, "A noisy rational expectations equilibrium for multi-asset securities markets," *Econometrica*, 53, pp. 629–658. 1, 5

Albagli, E., C. Hellwig, and A. Tsyvinski, 2023a, "Imperfect financial markets and investment inefficiencies," *American Economic Review*, 113, 2323–2354. 6

Albagli, E., C. Hellwig, and A. Tsyvinski, 2023b, "Information aggregation with asymmetric asset payoffs," *The Journal of Finance*, Forthcoming. 6

Allen, B., 1981, "Generic existence of completely revealing equilibria for economies with uncertainty when prices convey information," *Econometrica*, 49, 1173–1199. 6

Allen, B., 1982, "Strict rational expectations equilibria with diffuseness," *Journal of Economic Theory*, 27, 20–46. 6

Ang, A., M. Piazzesi, and M. Wei, 2006, "What does the yield curve tell us about GDP growth?," *Journal of Econometrics*, 131, 359–403. 2, 26

Atmaz, A., and S. Basak, 2018, "Belief dispersion in the stock market," *Journal of Finance*, 73, 1225–1279. 7, 8

Avdis, E., and S. Glebkin, 2023, "CHILE," Working Paper. 6

Bai, J., T. Philippon, and A. Savov, 2016, "Have financial markets become more informative?," *Journal of Financial Economics*, 122, 625–654. 25

Banerjee, S., and B. Green, 2015, "Signal or noise? Uncertainty and learning about whether other traders and informed," *Journal of Financial Economics*, 117, 398–423. 9, 12

Basak, S., 2000, "A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk," *Journal of Economic Dynamics & Control*, 24, pp. 63–95. 7, 8

Basak, S., 2005, "Asset pricing with heterogeneous beliefs," *Journal of Banking & Finance*, 29, pp. 2849–2881. 7, 8

Basak, S., and D. Cuoco, 1998, "An equilibrium model with restricted stock market participation," *Review of Financial Studies*, 11, pp. 309–341. 24

Begenau, J., M. Farboodi, and L. Veldkamp, 2018, "Big data in finance and the growth of large firms," *Journal of Monetary Economics*, 97, 71–87. 7

Blume, L., D. Easley, and M. O'Hara, 1994, "Market statistics and technical analysis: the role of volume," *The Journal of Finance*, 49, 153–181. 9

Brennan, M. J., and H. H. Cao, 1996, "Information, trade, and derivative securities," *Review of Financial Studies*, 9, pp. 163–208. 1, 5

Breon-Drish, B., 2010, "Asymmetric information in financial markets: anything goes," *Working Paper*. 9, 12

Breon-Drish, B., 2015, "On existence and uniqueness of equilibrium in a class of noisy rational expectations models," *Review of Economic Studies*, 82, pp. 868–921. 3, 5, 13, 16

Breugem, M., A. Buss, and J. Peress, 2020, "Learning from interest rates: implications for stock-market efficiency," Working Paper. 6

Brunnermeier, M. K., A. Simsek, and W. Xiong, 2014, "A welfare criterion for models with distorted beliefs," *Quarterly Journal of Economics*, 129, 1753–1797. 7, 8

Campbell, J. Y., 2003, "Consumption-based asset pricing," in G. Constantinides, M. Harris, and R. Stulz (ed.), *Handbook of Economics and Finance* . pp. 801–885, Elsevier, Amsterdam. 24

Chabakauri, G., 2015, "Asset pricing with heterogeneous preferences, beliefs, and portfolio constraints," *Journal of Monetary Economics*, 75, 21–34. 7, 8

Chabakauri, G., V. Fos, and W. Jiang, 2023, "Trading ahead of barbarians' arrival at the gate: insider trading on non-inside information," Working Paper. 9

Chabakauri, G., and B. Y. Han, 2020, "Collateral constraints and asset prices," *Journal of Financial Economics*, 138, 754–776. 7, 8

Chabakauri, G., K. Yuan, and K. E. Zachariadis, 2022, "Multi-asset noisy rational expectations equilibrium with contingent claims," *Review of Economic Studies*, 89, 2445–2490. 5, 10, 16

Chakravarty, S., H. Gulen, and S. Mayhew, 2004, "Informed trading in stock and option markets," *Journal of Finance*, 59, pp. 1235–1257. 1, 25

Cvitanić, J., and S. Malamud, 2011, "Price impact and portfolio impact," *Journal of Financial Economics*, 100, 201–225. 7, 8

Daniel, K., D. Hirshleifer, and A. Subrahmanyam, 1998, "Investor psychology and security market under- and overreactions," *Journal of Finance*, 53, 1839–1885. 25

Davila, E., and C. Parlatore, 2021, "Trading costs and informational efficiency," *Journal of Finance*, 76, pp. 1471–1539. 5, 8

Detemple, J., and S. Murthy, 1997, "Equilibrium asset prices and no-arbitrage with portfolio constraints," *The Review of Financial Studies*, 10, 1133–1174. 7, 8

Detemple, J., M. Rindisbacher, and S. Robertson, 2020, "Dynamic noisy rational expectations equilibrium with insider information," *Econometrica*, 88, 2697–2737. 5, 8, 9, 12

Detemple, J., M. Rindisbacher, and S. Robertson, 2022, "Dynamic noisy rational expectations equilibrium with insider information: welfare and regulation," *Journal of Economic Dynamics and Control*, 141, 1–31. 5

Diamond, D. W., and R. E. Verrecchia, 1981, "Information aggregation in a noisy rational expectations economy," *Journal of Financial Economics*, 9, pp. 221–235. 1, 5

Duffie, D., 2001, *Dynamic Asset Pricing Theory, 3rd Edition*, Princeton University Press. 11

Easley, D., M. O'Hara, and P. S. Srinivas, 1998, "Option volume and stock prices: evidence on where informed raders trade," *Journal of Finance*, 53, pp. 431–465. 1, 25

Farboodi, M., and L. Veldkamp, 2020, "Long run growth of financial data technology," *American Economic Review*, 110, 2485–2523. 7

Glebkin, S., N. Gondhi, and J. C.-F. Kuong, 2021, "Funding constraints and informational efficiency," *The Review of Financial Studies*, 34, 4269–4322. 5

Glebkin, S., S. Malamud, and A. Teguia, 2023a, "Illiquidity and higher cumulants," *Review of Financial Studies*, 36, 2131–2173. 6

Glebkin, S., S. Malamud, and A. Teguia, 2023b, "Strategic trading with wealth effects," Working paper. 6

Gradshteyn, I. S., and I. M. Ryzhik, 2007, *Table of Integrals, Series, and Products*, Elsevier. 40, 41

Grossman, S. J., 1976, "On the efficiency of competitive stock markets where trades have diverse information," *Journal of Finance*, 31, pp. 573–585. 1, 5

Grossman, S. J., and J. E. Stiglitz, 1980, "On the impossibility of informationally efficient markets," *American Economic Review*, 70, pp. 393–408. 1, 5

Han, B. Y., 2018, "Dynamic information acquisition and asset prices," *Working Paper*. 5

Harvey, C. R., 1988, "The real term structure and consumption growth," *Journal of Financial Economics*, 22, 305–333. 2, 26

Hellwig, M. R., 1980, "On the aggregation of information in competitive markets," *Journal of Economic Theory*, 22, 477–498. 1, 5

Huang, S., B. Z. Yueshen, and C. Zhang, 2023, "Derivatives and Market (Il)liquidity," *Journal of Financial and Quantitative Analysis*, Forthcoming. 5

Jameson, G. J. O., 2006, "Counting zeros of generalised polynomials: Descartes' rule of signs and Laguerre's extensions," *Mathematical Gazette*, 90, 223–234. 32

Johnston, W., 2014, "The weighted Hermite polynomials form a basis for $L_2(R)$," *American Mathematical Monthly*, 121, 249–253. 40

Kacperczyk, M., and E. S. Pagnotta, 2019, "Chasing Private Information," *Review of Financial Studies*, 32, 4997–5047. 1, 25

Kolmogorov, A. N., and S. V. Fomin, 1999, *Elements of the Theory of Functions and Functional Analysis*, Dover Publications. 40

Kreps, D., 1977, "A note on "fulfilled expectations" equilibria," *Journal of Economic Theory*, 14, 32–43. 9, 12

Kurlat, P., and L. Veldkamp, 2015, "Should we regulate financial information?," *Journal of Economic Theory*, 158, pp. 697–720. 5

Lebedev, N. N., 1972, *Special Functions and Their Applications*, Dover Publications. 39, 40, 42

Lucas, R. E., 1978, "Asset prices in an exchange economy," *Econometrica*, 46, 1429–1445. 8

Malamud, S., 2015, "Noisy Arrow-Debreu equilibria," *Working Paper*. 5, 6, 9, 12

Martin, I. W. R., and D. Papadimitriou, 2022, "Sentiment and speculation in a market with heterogeneous beliefs," *American Economic Review*, 112, 2465–2517. 7, 8

Pálvölgyi, D., and G. Venter, 2015, "On equilibrium uniqueness in multi-asset noisy rational expectations economies," *Working Paper. Available at SSRN: http://ssrn.com/abstract=2631627*. 5

Pan, J., and A. M. Poteshman, 2006, "The information in option volume for future stock prices," *Review of Financial Studies*, 19, pp. 871–908. 1, 25

Peress, J., 2004, "Wealth, information acquisition, and portfolio choice," *Review of Financial Studies*, 17, 879–914. 6

Radner, R., 1979, "Rational expectations equilibrium: generic existence and information revealed by prices," *Econometrica*, 47, 655–678. 6

Rostek, M., and M. Weretka, 2012, "Price inference in small markets," *Econometrica*, 80, 687–711. 7

Schneider, J., 2009, "A rational expectations equilibrium with informative trading volume," *The Journal of Finance*, 64, 2825–2859. 9

Shervashidze, T. L., 1971, "On a uniform estimate of the rate of convergence in the multidimensional local limit theorem for densities," *Probability Theory and Applications*, 16, 741–743. 35

Smith, K., 2019, "Financial markets with trade on risk and return," *The Review of Financial Studies*, 32, 4041–4078. 5

Tetlock, P., 2009, *Expert Political Judgement: How Good Is It? How Can We Know?*, Princeton University Press, Princeton. 3, 17

Tikhonov, A. N., and V. Arsenin, 1977, *Solutions of Ill-posed Problems*, John Wiley & Sons. 20

Vanden, J. M., 2008, "Information quality and options," *Review of Financial Studies*, 21, pp. 2635–2676. 2, 5

Vives, X., 2008, *Information and Learning in Markets*, Princeton University Press. 5

Zorich, V. A., 2004, *Mathematical Analysis: Volume II*, Springer. 37, 38