# How Much Information Is Incorporated into Financial Asset Prices? Experimental Evidence*

Lionel Page† Christoph Siemroth‡

December 23, 2020

**Abstract**

We investigate the informational content of prices in financial asset markets. To do so, we use a large number of market experiments in which the amount of information held by traders is precisely observed. We derive a new method to estimate how much of this information is incorporated into market prices. We find that public information is almost completely reflected in prices but that surprisingly little private information—less than 50%—is incorporated into prices. Our estimates therefore suggest that, while semistrong informational efficiency is consistent with the data, financial market prices may be very far from strong-form efficiency. (*JEL* C92, D82, D84, G14)

*We thank Yair Antler, Elena Asparouhova, Roy Bailey, Dan Bernhardt, Peter Bossaerts, Brice Corgnet, Cary Deck, Mark DeSantis, Felix Fattinger, Jayant Ganguli, P. J. Healy, Santiago Oliveros, Elise Payzan Lenestour, Talis Putnins, Sander Renes, Yohanes Eko Riyanto, Joel Sobel, Peter Norman Sørensen, Matthias Sutter, and Bertil Tungodden and seminar participants at Essex, Mannheim, NTU, QUT, VU Amsterdam, Zurich, ESA Vienna 2017, APESA Brisbane 2018, RES Brighton 2018, and EEA-ESEM Cologne 2018 for discussion and helpful comments. We are also very grateful to all authors who shared their data with us for this project. Send correspondence to Lionel Page, lionel.page@uts.edu.au.

†University of Technology Sydney, Economics Discipline Group. Email: lionel.page@uts.edu.au.

‡University of Essex, Department of Economics, Wivenhoe Park, CO4 3SQ Colchester, UK. E-Mail: christoph.siemroth@essex.ac.uk.

> Few engineers would ever consider performing a statistical test to determine whether or not a given engine is perfectly efficient – such an engine exists only in the idealized frictionless world of the imagination. But measuring relative efficiency—relative, that is, to the frictionless ideal—is commonplace. —Lo (2008)

The efficient market hypothesis (EMH) is one of the most influential concepts in economics. The hypothesis combines the Hayekian idea that market prices aggregate information held by economic agents (Hayek, 1945) and Samuelson's principles of agents' rationality and market equilibrium (Samuelson, 1965). As stated by Fama, a market is said to be informationally efficient if its prices "fully reflect" the available information (Fama, 1970).

Market efficiency has important consequences for our understanding of how financial markets work and how prices can be used to inform the decisions of economic agents from traders to governments. Of particular interest is how much private information is reflected in prices, since market prices may be the only source to reveal private information about economic fundamentals to decision makers, whereas public information can be gleaned from other sources. And while a theoretical literature on feedback effects investigates the strategic implications of decision makers learning from prices to improve decisions and welfare (e.g., Bond and Goldstein, 2015; Boleslavsky, Kelly, and Taylor, 2017; Siemroth, 2019), its effectiveness in practice depends on how much private information prices contain in the first place. This question, and how this private information amount compares to public information, is the focus of our study.

Whether and to what degree are markets indeed informationally efficient has been the object of intense theoretical and empirical research (Malkiel, 2005; Lim and Brooks, 2011). While the EMH links prices to the available information, the informational content of prices is difficult to assess when trader information is not observed. As a consequence, the bulk of the empirical research has conducted indirect tests and focused on the observable consequences of market efficiency: the fact that prices should follow random walks and be unpredictable thereby preventing the possibility of systematically profitable trading strategies. But if perfect efficiency is rejected, these indirect tests cannot tell us how much of the available information is incorporated into prices.

In the present study, we take a new approach and directly quantify the informational content of prices based on 664 experimental asset markets from five existing studies supplemented by new observations. In these experimental markets with Arrow-Debreu securities, all information sets and distributions are observed, and the theoretical model prices for the information realizations can be computed. We develop a structural model to estimate the share of information (signals)—as a percentage number between 0 and 100—which is reflected in the observed asset prices.

Our results are striking. First, we find evidence supporting semistrong informational efficiency, that is, prices incorporating all public information. Our estimates range from 90% to 100% of public information being reflected in the last transaction prices. Indeed, the hypothesis of all public information being incorporated cannot be rejected, at least when we consider transaction prices at the end of the trading window. Second, our main finding is that experimental asset prices reflect surprisingly little *private* information. Estimates for the share of private information used by the market range between 0% and 30%, depending on the experiment. Thus, the estimates suggest that market prices are far away from the ideal of strong-form informationally efficient markets or fully revealing rational expectations equilibrium prices. Indeed, we can reject the hypothesis that 50% of private information or more is incorporated into asset prices in any of the experiments we use.

Defining a second measure to quantify informational efficiency, by normalizing observed prices between the two benchmarks of "full information prices" and "no information prices", we confirm these findings. This second measure also demonstrates that the mispricing due to imperfect information aggregation by the market is less severe than the low estimates of our first measure might suggest, due to a concavity in the Bayesian posterior probability function. Thus, even though markets incorporated less than 50% of private information, the prices can nevertheless be closer to the "full information price" than to the "no information price." Our approach therefore shows that studies looking at prices, rather than information, may overestimate how informationally efficient markets are. Still, whichever of our two measures is used, our finding remains that private information is inefficiently aggregated by markets.

The recent cursed expectations equilibrium concept (Eyster, Rabin, and Vayanos, 2019) provides one possible explanation for our finding that private information is imperfectly aggregated. It suggests the limited ability of traders to infer information from prices causes private information to be imperfectly aggregated. In contrast, a rational expectations equilibrium, which corresponds to 100% information in our setting, cannot explain the results in private information markets.

Finally, we compare the unexpectedly small share of private information inherent in prices to the average belief on market efficiency among economists. We conducted a survey among all 2017 Econometric Society Meeting participants—overall more than 300 academic economists responded—to see whether the estimates and their beliefs differ about how much information is reflected in asset prices. In particular, we asked academics how much information they believe is reflected in real-world financial markets, and how much in the experimental asset markets we are studying (incentivized). On average, economists believe that real financial markets incorporate 77% of information, and that experimental markets incorporate 71% of information. These beliefs markedly overestimate our results. Indeed, only 4% of respondents had a belief equal to or less than the maximal experimental estimate (30%). These responses suggest that economists overestimate the ability of financial markets and experimental markets to incorporate private information.

A common observation is that strong-form informationally efficient markets (which incorporate all public and private information) imply that traders cannot systematically earn risk-adjusted profits relative to the market. However, the inverse need not be true. Our finding of not strong-form informationally efficient markets does not necessarily imply that systematic profits can be made. Our results state that markets are not perfectly efficient based on the strongest of criteria—the combined information of decentrally distributed private signals—but exploiting this inefficiency for profit might require more information than any single trader has access to. This is in contrast to the weaker criteria of semistrong or weak informational efficiency, where the information needed to exploit an inefficiency is public. Thus, our findings of a small share of private information being integrated into prices are not necessarily in conflict with the many empirical findings reporting that it is difficult to systematically earn risk-adjusted profits.

Our estimates appeal to more than academic interest, because policy proposals regularly call for more reliance on market information and prices, for example, in the case of banking supervision (e.g., Flannery, 1998; Greenspan, 2001; Flannery, Houston, and Partnoy, 2010) or contingent capital with market triggers (e.g., Flannery, 2016). Since the merits of such proposals depend on how much information is incorporated into prices, the results and techniques of this paper might help in these debates. Moreover, our method can help to identify the markets and conditions that reveal more information, which tells regulators and market participants where to look for the most information. Finally, the question of whether prediction markets should be used for decision and policy making (e.g., Wolfers and Zitzewitz, 2004; Cowgill and Zitzewitz, 2015) depends on how much information the prices in these markets reveal. In this vein, our method and estimates can be used to compare prediction markets to alternative information aggregation and forecasting mechanisms. Indeed, the method we derive can easily be applied to quantify information aggregation in those alternatives, such as probability forecasts.

Our paper contributes to the literature by developing a method and estimating how much information is incorporated into asset prices. A vast literature surrounding the efficient market hypothesis reports on tests of informational efficiency by looking at the observable consequences of price efficiency. We will not attempt to review this large body of literature here. Instead, an overview of the papers can be found in, for example, Fama (1998), Malkiel (2003, 2005), Yen and Lee (2008), or Lo (2008). Overall, while this existing empirical literature contributes to the question of "are market prices informationally efficient?" by testing the empirical implications of the efficient market hypothesis, our paper instead contributes to answering the question of "how informationally efficient are financial market prices?" which the initial quote illustrates. While perhaps no market is perfectly efficient, whether prices incorporate 90% or only 10% of information matters greatly to determine their predictive power and usefulness. It is thus important to quantify the extent of informational efficiency.

An advantage of the experimental approach taken here is that, first, all relevant variables, such as private and public information of traders and realized asset values, are observed, unlike in field stock markets. Second, a common prior can be established. And third, causal effects are easily identified via random assignment to treatment, all of which allows for very clean and direct tests, whereas analyses based on field data typically have to rely on indirect tests. The greater amount of information contained in our experimental data allows us to implement structural estimations which deliver new and more precise answers that we would not be able to get otherwise. Hence, our experimental estimates are interesting *because* what we do is typically not feasible with field data and should therefore be seen as a complement to existing field studies.

An often-raised concern is that results from the lab may not generalize to the field. Because this is a central question for the experimental method, it has received a lot of attention. In finance, for example, the common findings of asset bubbles arising in experimental markets (e.g., Smith, Suchanek, and Williams, 1988 and the hundreds of variations of that experiment since) mirror the bubble patterns observed in field markets, and they also occur when finance professionals take part in the same experimental markets (e.g., Weitzel et al., 2019). Hence, qualitative results are similar whether students or professionals participate in the experiments, though magnitudes of behavior can vary slightly.[1] For these reasons, we regard our findings of high informational efficiency for public information and low informational efficiency for private information, which were observed in all of the experiments, as robust.

Our paper also contributes to the literature on market experiments. This literature has played a substantial role in building economists' confidence in market mechanisms by finding that even in small laboratory settings, markets are able to converge to competitive equilibrium prices (Smith, 2007). Early on, several experiments with one-period lived assets tested whether rational expectations equilibrium (REE)—a formalization of strong-form informational efficiency—fits the experimental market data. In these studies, the REE concept was typically able to outperform competing theories, such as Walrasian equilibrium price predictions (e.g., Forsythe, Palfrey, and Plott, 1982; Plott and Sunder, 1982, 1988; Copeland and Friedman, 1987; Forsythe and Lundholm, 1990). However, almost all of these early studies use simple information structures, such as perfect signals for insiders that reveal the realized asset value, while we consider more complex but arguably also more realistic information structures where signals are imperfect and stochastically informative. Indeed, Plott and Sunder (1988) also features imperfect private signals, and while the REE model arguably fits their data best in some conditions, the experimental prices never fully converge to the REE price.

Later studies using similar designs have found more evidence of deviations from REE prices, indicating that the information dispersed in the market was imperfectly incorporated into the experimental prices (e.g., Biais et al., 2005; Hanson, Oprea, and Porter, 2006; Veiga and Vorsatz, 2010; Corgnet, DeSantis, and Porter, 2015). Other experimental papers using different market designs, for example, in the context of decentralized markets, have also found deviations from full informational efficiency (e.g., Huber, Angerer, and Kirchler, 2011; Bossaerts, Frydman, and Ledyard, 2013; Goeree and Zhang, 2015; Asparouhova, Bossaerts, and Yang, 2017). Importantly, these studies test observed data against benchmarks, such as the REE price prediction. But none of them estimates how much information is incorporated into prices nor do they have a metric for it. We fill this gap.

Our approach is related to at least one prior study which in some way quantified information in an experimental market. Bossaerts (2009, section 9) reports results from a lab experiment in which 2 of 16 traders were informed of the asset value, while the remaining traders only knew the prior distribution. The author found that the Walrasian equilibrium prediction had a poor fit to the data. He then generated counterfactual predictions of Walrasian equilibria where more than the two traders are informed of the asset value. The equilibrium prediction with 8 of 16 traders being informed best fits the data; that is, transaction prices behave *as if* 8 of the traders are informed even though only two are informed. This approach can be viewed as quantifying how far perfect information spreads or is transmitted from informed to uninformed traders. The numbers suggest that information spreads to 6 of 14 uninformed traders, or approximately to 43% of uninformed traders. In comparison, our approach consists in directly quantifying how much information is integrated in

[1] See also Cipriani and Guarino (2009); Snowberg and Yariv (2018); Schwaiger et al. (2019), all of whom report findings with similar behavior between students and other subject pools. See also the comprehensive literature review by (Frechette, 2015).

market prices.

# 1. Market Setting and Experimental Designs

## *1.1. Market setting*

While our method can be applied (with adaptations) to any information structure and asset combination, in this study we investigate experimental markets with Arrow-Debreu state-contingent securities, also called binary options. Such assets have been extensively used in theoretical and experimental research and are also typically used by prediction markets, which have attracted a lot of attention from economists (Arrow et al., 2008; Ottaviani and Sørensen, 2015). The prices of such assets are meant to aggregate the traders' information on the likelihood of event $A$, providing an approximation for the probability of $A$ happening (Pennock et al., 2001; Wolfers and Zitzewitz, 2004).

## *1.2. Experimental designs*

We collect data on laboratory experiments that met the following conditions: markets with Arrow securities, imperfect signals, one-period lived assets (we exclude multiperiod "bubble experiments"), and common asset values. Overall, we collect market data from five experimental studies, all from different sets of authors. We also complemented these existing studies by running a new set of experimental markets. This data collection allows us to build a large data set in terms of number of observations that surpasses most lab market studies. It also demonstrates robustness of our results along various dimensions, such as subject pools (from four different continents) or experimental protocols. Hence, our contribution is not merely an aggregation of existing results and data; rather, we reanalyze existing data that were generated for different purposes in light of a new research question and a different method.

While details and parameter values differed across the different experiments, all share the following features. In each market, the state of the world $\theta$ is binary: $\theta \in \{A, B\}$ with prior probability $\Pr(\theta = A) = 0.5$. A risky asset pays off 1 (if $\theta = A$) or 0 (if $\theta = B$). Thus, the task of the traders was to figure out whether the state of the world was $A$ or $B$, and trade accordingly. Traders (8-12 depending on the experiment) receive an initial endowment of cash and risky assets. They then receive (or buy) private imperfect signals which are independently and identically distributed conditional on the state. All signals $i$ have the following signal quality: $\Pr(s_i = A|\theta = A) = \Pr(s_i = B|\theta = B) = q$, that is, indicate the correct state of the world with a probability $1/2 < q < 1$. Hence, the more $A$ signals a trader receives, the larger the posterior probability that the state is, in fact, $A$.

After obtaining their private signals, traders participate in a continuous double auction where they can submit buy or sell offers for the asset (limit orders), accept the offers of others at any time (market orders), or not trade at all. Thus, traders could buy the assets of others for cash, sell their assets for cash, or not trade at all. At the end of the round, the state of the world as well as trader earnings based on final asset holdings are revealed.

## *1.3. Data*

First, we use all treatments from Page and Siemroth (2017)'s experiment. In each session of the experiment, 9–12 traders participated in 12 markets (rounds), plus one practice round in the beginning. There are 108 markets overall. In the beginning of each round, traders receive two binary signals for free, either private or public, depending on the treatment. Then traders decide whether to buy additional private signals at a fixed cost. They can buy up to 10 additional signals. All signals have a signal quality of $q = 0.6$, that is, indicated the correct state of the world in 60% of the cases. The information acquisition stage is followed by the trading stage, where all traders see their signal profile and trade in a standard continuous double auction for 3 minutes, with the possibility of making limit or market orders.

Second, we use two treatments from the Deck, Lin, and Porter (2013) study (control and liquidity treatments). As in all other studies, they also have a double auction setup with one asset whose return depended on a binary state of the world. The quality of their signals is $q = 2/3$. Moreover, their experiment has eight traders per market and each trader receives exactly one signal. From the two treatments, we use the data of 40 markets in total.

Third, we use all treatments from Fellner and Theissen (2014)'s experiment where the asset takes a value of $L = 100$ or a value of $H = 200$ (which we normalize to $L = 0$ and $H = 1$). In each market, 10 traders receive a private signal with precision $q = 0.6$ or $q = 0.8$, depending on the round and treatment. We use 318 markets from this study.

Fourth, we use two treatments from Halim, Riyanto, and Roy (2019), who use a very similar design to Page and Siemroth (2017). In their control treatment, traders have the possibility to acquire private signals before trading, with a signal precision of $q = 0.6$. In their fully connected network treatment, traders can also acquire signals, but this information is shared publicly with all traders in the network (making it a public information treatment). In all of their treatments, traders receive two public draws for free, but because the vast majority of signals in their control treatment is private, we designate it as a private information treatment. From these two treatments, we use the data of 142 markets in total.

Fifth, we use the real money, no manipulation treatment from Su and Wang (2017). The method replicates the control treatment of Deck, Lin, and Porter (2013) and therefore uses the same parameters, that is, also eight signals for eight traders with a signal precision equal to $q = 2/3$. We use the data of 20 markets from this study.

Sixth, we run additional public information treatments based on the Page and Siemroth (2017) experimental design, with the same signal precision, the same number of traders per market, etc. Table 1 summarizes all of the important experimental parameters. The Internet Appendix gives further summary statistics on trading volume and shows price graphs.

## *1.4. Two important market properties*

Two important properties have been previously documented in markets with Arrow securities that motivate our estimation approach:

**Table 1** Overview of the experimental parameters

| Data source | Signals | Number markets | Traders per market | Num. signals per market | Signal precision |
|---|---|---|---|---|---|
| Deck, Lin, and Porter (2013) | Private | 40 | 8 | 8 | 2/3 |
| Fellner and Theissen (2014) | Private | 318 | 10 | 10 | 0.6 or 0.8 |
| Halim, Riyanto, and Roy (2019) | Private | 70 | 8 | 17.2 | 0.6 |
| Halim, Riyanto, and Roy (2019) | Public | 72 | 8 | 8.2 | 0.6 |
| Page and Siemroth (2017) | Private | 108 | 12 | 40.7 | 0.6 |
| Page & Siemroth (New) | Public | 36 | 12 | 15 | 0.6 |
| Su and Wang (2017) | Private | 20 | 8 | 8 | 2/3 |
| Total | | 664 | | | |

## 1.5. *Calibration*

The empirical research on such markets finds that their prices are well calibrated (e.g., Deck and Porter, 2013; Page and Clemen, 2013 for a short time to maturity; Dreber et al., 2015; Page and Siemroth, 2017). This means that in all the situations in which market prices are, say, 0.6, then in 60% of these cases the asset value is equal to one; when market prices are 0.7, then in 70% of these cases the asset value is equal to one, and so on. Formally, for state of the world $\theta$ and observed prices $p_m$, well calibrated prices require $p_m = \Pr(\theta = A|p_m)$. Thus, market prices of an asset paying 1 if $\theta = A$ on average equal the conditional probability of $\theta = A$ in this binary setting. This result is often interpreted as indicating that these markets provide a good estimate of the probability of event $A$.

The concept of calibration is closely linked with weak-form informational efficiency, which requires that current prices incorporate all information contained in past prices. This is because miscalibrated prices are not weak-form efficient: If the asset value is equal to 1 in 80% of the cases whenever market prices are 0.7, formally $\Pr(\theta = A|p_m = 0.7) = 0.8$, then it would be profitable in expectation to buy whenever $p_m = 0.7$; that is, past price information can be used to make profits. Thus, well calibrated prices are necessary for weak-form informational efficiency. But the fact that prices are well calibrated does not imply that they are semistrong or strong-form efficient and incorporate all private information.[2]

## 1.6. *Underreaction to information*

The second common finding is underreaction of prices to information. Formally, market prices underreact to information and differ from the expected asset value based on the Bayesian posterior taking into account all the information present among traders. Underreaction has been observed in other settings for financial market prices (e.g., Gillette et al.,

[2] For a simple example showing this, suppose a pricing rule determined the price solely based on the prior probability distribution; therefore, it always (independent of the information in the market) yields a price $p = 0.5$. This pricing rule is well calibrated, because the asset pays off 1 (if $\theta = A$) in 50% of the cases, and, thus, $\Pr(\theta = A|p = 0.5) = 0.5$. But, by construction, the pricing rule ignores all information (private or public) and is thus neither semistrong nor strong-form informationally efficient.

1999; Stevens and Williams, 2004; Kirchler, 2009). Page and Siemroth (2017) simply compared the full information price with the observed prices and found significant differences. However, besides rejecting the hypothesis that 100% of information is used, these "reduced-form" analyses cannot tell us how much information is in fact used. In the present study, we address this question.

# 2. Estimation of the Proportion of Information Reflected in Prices

## *2.1. Empirical approach*

We develop a structural model to estimate how much information is incorporated into asset prices. Unlike the conventional approach to structural modeling in microeconomics, we do not directly impose assumptions on traders' behavior or utility functions and instead make assumptions about the market itself. To impose structure, our assumptions are chosen to be consistent with the two prior experimental findings just described. Hence, our approach is deliberately agnostic about the trader level for the following reasons.

First, we do not want to assume a specific trader bias, as our interest is in the informational efficiency of market prices and not in testing or arguing for a specific trader bias. This means that the thrust of our approach does not rely on the plausibility or evidence for a specific bias. Moreover, with our approach we are arguably better able to capture potential inefficiencies that might arise for other reasons than one specific trader bias. This is an important point, since as soon as one deviates from the idea of rational expectations equilibrium prices, there is no consensus on which micro-founded asset pricing model or which bias (overconfidence, cursedness, etc.) should be the right one.

Second, our approach allows us to quantify the informational efficiency of prices, and based on that inform the debate about which asset pricing model fits the data best. We discuss this in Section 4.4.2. Hence, because it is not clear what objective traders are exactly maximizing in practice,[3] our approach is more general to accommodate the outcomes of various different asset pricing models, so our estimates can be used to assess those models, rather than have their assumptions baked in from the outset.

Third, our estimation approach is portable to other domains beyond financial market prices. For example, prices in our context are in the interval $[0, 1]$, and so with almost no changes our model could be used to infer the informational content of probability forecasts (e.g., based on surveys or machine learning models). This is not always possible when making detailed assumptions at the micro level, because assumptions will vary depending on the domain and outcome to be assessed.

Now that we have motivated our approach, the challenge is to derive a method or model that is consistent with the two results from the literature mentioned earlier: market prices are well calibrated but may underreact to the available information. Taking these two properties into account, we posit that it is *as if* the market incorporates only randomly sampled subsets of the available signals into the price of the asset.

[3]We thank an anonymous referee for this phrasing.

The assumption that the market may incorporate only a subset of the available signals into the price is a simple way to model prices underreacting to the available information. And taking well calibrated prices as given implies that prices are not systematically biased toward outcome $A$ or $B$. We model this unbiasedness by assuming that the signals incorporated into prices are randomly selected with equal probability from the set of all available signals without replacement. Note that we do not impose underreaction to information in our model; whether and to what degree prices underreact is a parameter to be estimated ($\lambda$ below).

To illustrate our approach, we consider a simple example. First, a market price of $p = 0.5$ corresponds to the case of 0% information. It is the expected asset value ignoring all information, just based on the prior probability distribution, $P(A) = 0.5$, according to which the asset pays off one half of the time.

Second, the 100% benchmark is the expected asset value based on all information in the market (full informational efficiency). Suppose the entire market holds five signals, which can be written as the realization vector $S = (A, A, A, A, B)$. To illustrate, assume a signal quality of $\Pr(s = \theta) = 0.6$. Then, by Bayes' rule, the posterior probability of $A$ (and thus the asset paying off one) given $S$ is $\Pr(A|S) \approx 0.77$. Based on a simple risk-neutral asset pricing model, 100% information therefore corresponds to an asset price of $p \approx 0.77$. More generally, prices based on more information should be closer to zero or one and farther away from the "0% price" of 0.5, which is the basic idea that we exploit in our estimation.

Going beyond these natural benchmarks, consider now the case if prices are based on a share $\lambda \in (0, 1)$ of information. Given a full information set $S = (A, A, A, A, B)$, if the market only uses a share of $\lambda = 0.4$ of the information, that is, 2 of 5 signals, then the possible information subsets are $S' = (A, A)$ and $S'' = (A, B)$, which correspond to prices $\Pr(A|S') \approx 0.69$ and $\Pr(A|S'') = 0.5$, respectively. If the market uses a larger share of $\lambda = 0.6$ of the information, then the possible subsets are $S^* = (A, A, A)$ or $S^{**} = (A, A, B)$, with corresponding prices of $\Pr(A|S^*) \approx 0.77$ or $\Pr(A|S^{**}) = 0.6$. Clearly, the possible prices tend to get closer to 0 or 1 as the relative size of the information subset $\lambda$ increases. This is how we can link observed prices to the size of the informational subset, that is, how to estimate which share of information fits the observed prices best.

While our estimation approach can be adapted for other information structures and assets, we now derive all expressions specifically for markets with binary signals, which were used in the experiments.

## 2.2. *Full efficiency benchmark: The market uses all available information*

In all of our experimental markets, the signals $s_1, s_2, \ldots, s_N$, which are independent conditional on the state from distribution $\Pr(s_j = A|\theta = A) = \Pr(s_j = B|\theta = B) = q > 1/2$, are dispersed among traders in the market. The total number of signals in the market, which are all contained in set $S_N$, is $N$, and the total number of $s_j = A$ signals in the market is $K$.

With the standard assumption of risk-neutral pricing,[4] financial market prices perfectly incorporate all available trader information if they equal the expected asset value based on

[4] In the Internet Appendix, we allow for different risk attitudes in the model and show that the estimates for the parameter of interest, $\lambda$, remain very close to those derived based on risk neutrality.

the conditional probability of $\theta$ given all signals, that is, if

$$p^* = \Pr(\theta = A | s_1, s_2, \ldots, s_N) = \frac{q^K \cdot (1-q)^{N-K}}{q^K \cdot (1-q)^{N-K} + q^{N-K} \cdot (1-q)^K}. \tag{1}$$

This is the theoretical benchmark and the "ideal" of the efficient market hypothesis mentioned in the introductory quote. It corresponds to a fully revealing rational expectations price with risk-neutral traders, and even a trader who had all (private) information in the market could not make systematic profits. But if market prices deviate from (1), then traders who use all information in the market could potentially make profits in expectation by exploiting the difference in prices to (1). However, exploiting such mispricing would require traders to use the information of other traders to compute (1), which they might not have access to if that information is private.

## *2.3. The market uses a subset of the available information to price the asset*

The informal description of our formal model is that market prices are set *as if* the market incorporates only a randomly drawn (without replacement) subset $S_n \subseteq S_N$, the set of all available signals. Hence, formally, the market incorporates only a number $n \leq N$ of all available signals with a corresponding number of $k \leq K$ $A$ signals.

For a *given* subset of signals $S_n$, the market price is the expected asset value using the conditional probability based on the information subset $S_n$,

$$\hat{p} = \Pr(\theta = A | S_n) = \frac{q^k \cdot (1-q)^{n-k}}{q^k \cdot (1-q)^{n-k} + q^{n-k} \cdot (1-q)^k}. \tag{2}$$

We can easily show that our model accommodates both calibration and underreaction.

First, underreaction is by design allowed by the model: it will be present any time $n < N$. In that case, $\hat{p}$ will usually not be equal to the price $p^*$ reflecting all the information. If $n < N$, then market prices determined by (2) could be potentially exploited for profit by a trader who has access to all $N$ signals in the market, since (2) discards pieces of information which are important for determining the expected value of the asset. Consequently, (2) with $n < N$ corresponds to imperfect trader information aggregation in prices, or informational inefficiency.

Second, market prices respect calibration. Using (2) and applying the conditional expectation with respect to prices on both sides, we obtain

$$\begin{aligned} \mathbb{E}[\hat{p}|\hat{p}] &= \mathbb{E}[\mathbb{E}[\mathbf{1}\{\theta = A\}|S_n]|\hat{p}] \\ \implies \hat{p} &= \mathbb{E}[\mathbf{1}\{\theta = A\}|\hat{p}] = \Pr(\theta = A|\hat{p}), \end{aligned}$$

by the law of iterated expectations. Consequently, our pricing model is consistent with both underreaction to information and the calibration of market prices. The next section explains how to quantify the informational content of prices by estimating the size of the subset of information $S_n$.

## 2.4. Estimation

We can express the subsample size that the market uses as $n = \lceil \lambda N \rceil$, where $\lambda \in [0, 1]$ is the share of all available signals used and $\lceil x \rceil$ is the ceiling function which ensures that every $\lambda$ yields a $n \in \mathbb{N}$. $\lambda$ is the main parameter of interest; it tells us how much information—relative to the available signals $N$ in the entire market—is incorporated into the observed prices. To estimate $\lambda$, we assume that observed prices $p_m$ for markets $m = 1, 2, \ldots, M$ are generated by (2):

$$p_m = \hat{p}_m(\lambda, k) + \varepsilon_m \iff \varepsilon_m = p_m - \frac{q^k \cdot (1-q)^{\lceil \lambda N_m \rceil - k}}{q^k \cdot (1-q)^{\lceil \lambda N_m \rceil - k} + q^{\lceil \lambda N_m \rceil - k} \cdot (1-q)^k}, \tag{3}$$

where $N_m$ indicates the number of signals available in a specific market $m$, and $\varepsilon_m$ a stochastic deviation from the price prediction. As is common in structural estimation, we model the distribution of $\varepsilon_m$ as normal, that is, $\varepsilon_m \sim \mathcal{N}(0, \sigma^2)$, where the standard deviation of the error distribution $\sigma$ is a (nuisance) parameter to be estimated, making the model more flexible.[5]

Now we can assign a probability (technically a density) of observing a specific market price $p_m$ given $(\lambda, k, \sigma)$, which is

$$\Pr(p_m | \lambda, k, \sigma) = \phi\left(\frac{\varepsilon_m}{\sigma}\right) / \sigma = \phi\left(\frac{p_m - \hat{p}_m(\lambda, k)}{\sigma}\right) / \sigma,$$

where $\phi(x)$ is the standard normal density. For a given $\lambda$ (which determines the number of incorporated signals $n$), the observed price $p_m$ can be generated by the model with several different $k$ (number of $A$ signals), yielding varying errors $\varepsilon_m$ and corresponding densities. Indeed, the observed price can be generated by all $k \in [\max\{0, \lceil \lambda N_m \rceil - N_m + K_m\}, \min\{\lceil \lambda N_m \rceil, K_m\}]$. The upper bound ensures that the number of $A$ signals used is not larger than either $n = \lceil \lambda N_m \rceil$ or the $K_m$ $A$ signals available. The lower bound ensures that the number of $B$ signals used is not larger than the $N_m - K_m$ available $B$ signals.

Since the market is assumed to draw a subset with $n = \lceil \lambda N_m \rceil$ of the $N_m$ available signals randomly without replacement, the probability distribution of the samples $(n, k)$ is hypergeometric. Thus, the probability of drawing a specific $k$ given $\lambda$ is

$$\Pr(k | \lambda) = \frac{\binom{K_m}{k} \binom{N_m - K_m}{\lceil \lambda N_m \rceil - k}}{\binom{N_m}{\lceil \lambda N_m \rceil}}. \tag{4}$$

Informally, (4) states that if the full sample has a lot of $A$ signals ($K_m$ large), then a random subsample will likely have a lot of $A$ signals. So random sampling tends to generate similar fractions of $A$ to $B$ signals as in the full information set.

Now we can bring the two probabilities of the $\varepsilon_m$ realizations and of the $k$ realizations together in a single likelihood function. The likelihood of observing market price $p_m$ given

[5]In the Internet Appendix, we show that our estimates do not change much if we use an alternative distribution like the beta distribution. We also confirm via simulations the robustness of our estimator to a wide range of situations in which prices are generated with distributions different than the normal distribution.

the model is the probability of drawing an information subset with $k$ $A$ signals and an error term $\varepsilon_m$ such that $p_m = \hat{p}(\lambda, k) + \varepsilon_m$. Thus,

$$\begin{aligned}
\Pr(p_m|\lambda,\sigma) &= \sum_{k=\max\{0,\lceil\lambda N_m\rceil - N_m + K_m\}}^{\min\{\lceil\lambda N_m\rceil, K_m\}} \Pr(p_m|\lambda,k,\sigma)\cdot\Pr(k|\lambda) \\
&= \sum_{k=\max\{0,\lceil\lambda N_m\rceil - N_m + K_m\}}^{\min\{\lceil\lambda N_m\rceil, K_m\}} \phi\left(\frac{p_m - \frac{q^k\cdot(1-q)^{\lceil\lambda N_m\rceil - k}}{q^k\cdot(1-q)^{\lceil\lambda N_m\rceil-k} + q^{\lceil\lambda N_m\rceil - k}\cdot(1-q)^k}}{\sigma}\right)/\sigma\cdot\frac{\binom{K_m}{k}\binom{N_m-K_m}{\lceil\lambda N_m\rceil - k}}{\binom{N_m}{\lceil\lambda N_m\rceil}}.
\end{aligned}$$

The objective is to find $(\widehat{\lambda}, \widehat{\sigma})$ that maximize the overall log-likelihood of observing the prices $(p_1, p_2, \ldots, p_M)$ in $M$ markets,

$$(\widehat{\lambda}, \widehat{\sigma}) = \arg\max_{\lambda\in[0,1],\sigma>0} \sum_{m=1}^{M} \ln\Pr(p_m|\lambda,\sigma).$$

Since the objective function is not continuous in the parameter $\lambda$ due to the ceiling function in $n = \lceil\lambda N_m\rceil$, standard numerical maximization procedures are not applicable. We therefore use a grid optimization for all $\lambda \in G$ with grid $G = [0,1] \cap \{0 + 0.01 \cdot l\}_{l=0,1,\ldots,100}$, and similarly for $\sigma$ in 0.01 steps. Section 4.4.2 explains which different values and ranges of $\lambda$ correspond to which microfounded asset pricing theories. In particular, if fully revealing rational expectations prices were present in the lab, then the estimates would be $\lambda = 1$, because the full information sample without replacement can only be drawn in one way, so the corresponding price with $\lambda = 1$ would perfectly fit the data. Prices that tend to reflect less information will be more consistent with the estimates of $\lambda < 1$.

In some experiments or treatments, the number of signals is either fixed or relatively small. In these cases, $\widehat{\lambda}$ is typically not unique. For example, if the number of signals is fixed at 10, then any $\lambda \in [0.01, 0.1]$ corresponds to 1 of 10 signals and has the same likelihood. In these cases, we report the $\lambda$ with a †-symbol in the tables and we show the largest estimate of the likelihood-maximizing interval in the results, for example, $\max\ [0.01, 0.1] = 0.1$.

We determine confidence intervals via the bootstrap percentile method by resampling market rounds to ensure the interval can never fall out of the possible interval $\lambda \in [0, 1]$.

# 3. Results

## *3.1. How much private information is incorporated into prices?*

Table 2 displays the maximum likelihood estimates of $\lambda$, which represent the share of the full sample of signals used by the market, based on data from various experiments in the literature. It also displays the standard deviation of the error distribution $\sigma$. We use three measures of the market price, since the double auction typically does not feature uniform prices. In particular, we compute $\lambda$ based on the mean and median transaction price of every market, and based on the last transaction price to allow for the fact that later transaction prices may converge or improve toward the full information value.[6] In this section, we only use treatments with private information,[7] and a comparison with public information treatments follows below.

The most striking result in Table 2 is that the share of signals used, $\lambda$, is small. Most point estimates based on mean or median prices for $\lambda$ are below 20%, except those from the Halim, Riyanto, and Roy (2019) experiment, where the point estimates indicate that almost 30% of information is used to price the asset. For every experiment, we can reject the hypothesis that 50% or more of the information is incorporated into the mean or median prices, as none of the confidence intervals of $\lambda$ includes 0.5. Overall, Su and Wang (2017) is the negative extreme in this comparison with the best estimate indicating that observed prices were no better than prices based on the prior distribution, and Halim, Riyanto, and Roy (2019) is the positive extreme.

[6] In the Internet Appendix, we also run the same estimations for the Page and Siemroth (2017) data on the unaggregated transaction-level data with very similar results. And we investigate the informativeness of bids and asks with very similar results.

[7] In one of the Page and Siemroth (2017) treatments (29 observations) and in the Halim, Riyanto, and Roy (2019) private information treatments, two out of many signals in the market were public, so information was "predominantly private," and we classify these markets as private here. All other treatments/experiments exclusively use private signals.

**Table 2** Maximum likelihood estimates of $\lambda$ based on data from the literature

| Data source | Mean price | | | Median price | | | Last price | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\lambda$ | $\sigma$ | LL | $\lambda$ | $\sigma$ | LL | $\lambda$ | $\sigma$ | LL |
| Deck, Lin, and Porter (2013) | $0.12^{\dagger}$ | 0.13 | 9.15 | $0.12^{\dagger}$ | 0.13 | 7.93 | $0.12^{\dagger}$ | 0.18 | 2.01 |
| (40 obs.) | [0.00, 0.25] | [0.09, 0.20] | | [0.00, 0.25] | [0.07, 0.21] | | [0.00, 0.25] | [0.08, 0.23] | |
| Fellner and Theissen (2014) | $0.10^{\dagger}$ | 0.14 | 100.71 | $0.20^{\dagger}$ | 0.16 | 91.62 | $0.30^{\dagger}$ | 0.15 | 35.46 |
| (318 obs.) | [0.10, 0.20] | [0.11, 0.15] | | [0.10.20] | [0.09, 0.15] | | [0.20, 0.30] | [0.11, 0.16] | |
| Page and Siemroth (2017) | 0.01 | 0.09 | 73.09 | 0.04 | 0.08 | 63.18 | 0.10 | 0.08 | 40.09 |
| (108 obs.) | [0.01, 0.05] | [0.06, 0.09] | | [0.01, 0.07] | [0.06, 0.1] | | [0.05, 0.12] | [0.06,0.12] | |
| Su and Wang (2017) | 0 | 0.12 | 14.27 | 0 | 0.12 | 13.91 | 0 | 0.15 | 9.15 |
| (20 obs.) | [0.00, 0.12] | [0.08, 0.13] | | [0.00, 0.12] | [0.08, 0.14] | | [0.00, 0.25] | [0.08, 0.16] | |
| Halim, Riyanto, and Roy (2019) | 0.29 | 0.08 | 27.16 | 0.29 | 0.08 | 25.65 | 0.30 | 0.13 | 13.93 |
| (70 obs.) | [0.12, 0.43] | [0.06, 0.13] | | [0.15, 0.46] | [0.05, 0.13] | | [0.19, 0.46] | [0.06, 0.17] | |

*Note:* The table displays the maximum likelihood estimates for the share of signals in the market used for pricing the asset ($\lambda$) and the standard deviation of the error distribution ($\sigma$). Estimates are based on the observed mean, median, or last transaction price. LL is the log-likelihood. The 95% confidence intervals below the estimates were determined by non-parametric percentile bootstrap, i.e., the confidence intervals correspond to the inner 95% of bootstrapped parameter estimates.

$^{\dagger}$ Estimate is the largest in the likelihood maximizing interval (see 2.2.4 for explanation)

The estimates of $\lambda$ tend to be larger for the last transaction price, increasing, for example, from 4% to 10% in the Page and Siemroth (2017) experiment or from 20% to 30% in the Fellner and Theissen (2014) experiment. This improvement means that later transaction prices are closer to the full information price (1) and incorporate more of the available information in the market.[8] Convergence or improvement over time is a common finding in the experimental asset market literature, and very early studies already found it for double auctions (e.g., Plott and Sunder, 1982, 1988).[9] More information in late transaction prices is consistent with informational theories where informed traders withhold their trades until the last moment in order not to reveal their private information to others (e.g., Ottaviani and Sørensen, 2006, 2009 in the context of parimutuel betting), but it is likely not the entire explanation, as all traders had private information in most of the experiments. This result also suggests that the double auction trading process can reveal some private information over time, so that later trades are based on more information.

While framing the results in terms of "percentage of information" $\lambda$ makes the experiments comparable, we can also look at the underlying number of signals to make the estimates more tangible. In the experiment with the most markets (observations), Fellner and Theissen (2014), ten traders each got one signal, so the estimates $\lambda = 0.1$, $\lambda = 0.2$, and $\lambda = 0.3$ indicate that 1, 2, or 3 of the 10 available signals are used by the market to price the asset. The difference between the Deck, Lin, and Porter (2013) and Su and Wang (2017) experiments are 12%, even though the latter is a replication of the former. However, in terms of signals, there is just a difference of one increment, that is, 1 of 8 signals, so these experiments are actually quite close.

The estimates overall imply that the market does not incorporate all *private* information into prices; that is, it is not strong-form informationally efficient. However, the estimates also suggest that markets use some information and improve over the prior price prediction of $p = 0.5$, which does not use any information, since most (but not all) 95% confidence intervals do not include $\lambda = 0$.

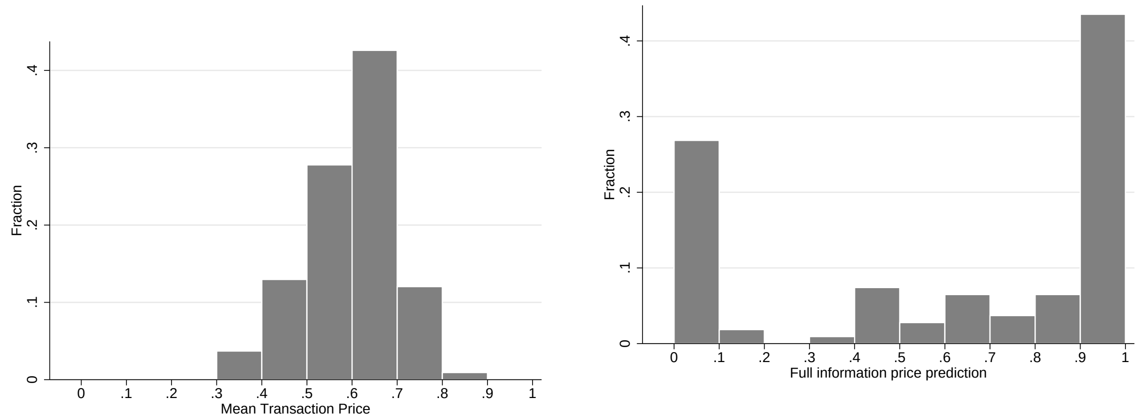
Result 1: *In none of the experiments is more than 50% of private information incorporated into asset prices. Hence, prices are not strong-form informationally efficient. In all but one experiment, the mean and median prices incorporate no more than 20% of information.*

## *3.2. Understanding the low estimates of $\lambda$ (Result 1)*

Let us consider the Page and Siemroth (2017) experiment to gain an understanding of the (perhaps surprising) finding that only a small part of the available information is reflected in prices. The results follow inevitably from the observed prices, which are largely between 0.4 and 0.7 and are displayed in Figure 1a (a histogram of the observed mean prices), whereas the market price should be close to 0 or 1 if all information had been used (see Figure 1b for

[8]This result is not specific to the very last transaction price. We computed the average prices of the last three transactions, and the $\lambda$ estimates (not reported) tend to be larger than those for the average prices but smaller than those for the last transaction price, also indicating that the later transaction prices carry more information.

[9]This improvement over time is also visible when plotting the absolute differences between the transaction prices and the full information prices over the 180 seconds of trading time. See Figure 2 in the Internet Appendix.




**(a)** Mean transaction prices

**(b)** Full information price prediction, Equation (1)

**Figure 1.** Mean transaction prices compared with the full information price predictions

a histogram of the full information prices). Comparing graphs 1a and 1b, the observed prices do not follow the full information price prediction and they are considerably less extreme, that is, closer to the (no information) prior price prediction of 0.5. This pattern reflects an underreaction to information.

According to our estimates, 1% to 4% of information (mean and median price estimates from Table 2) corresponds to about 1 or 2 signals out of on average 41 signals to price the asset. Consider some simple calculations to understand the low magnitude of the estimated $\lambda$. If the price incorporates only one signal, then the posterior probability of $A$ is $\Pr(\theta = A|s = A) = 0.6$ or $\Pr(\theta = A|s = B) = 0.4$ depending on the signal realization. This is already remarkably close to the market prices between 0.4 and 0.6 that we observe (Figure 1a). If the price incorporates 2 signals, then the posterior probabilities are $\Pr(\theta = A|s_1 = A, s_2 = A) \approx 0.69$, $\Pr(\theta = A|s_1 = B, s_2 = B) \approx 0.31$, or $\Pr(\theta = A|s_1 = A, s_2 = B) = 0.5$, depending on the signal realizations. As more signals are used, the posteriors can become more extreme, but more extreme posteriors imply a larger deviation from the observed prices and therefore do not fit the data as well.

## *3.3. How much public information is incorporated into prices?*

To understand better why only a small share of information is incorporated into asset prices, we investigate the informational reasons and compare experimental treatments where all of the signals are private and treatments where all of the signals are public. The difference $\lambda_{\text{public}} - \lambda_{\text{private}}$ represents the difficulty of aggregating and incorporating private information. There may be other reasons apart from information asymmetries to explain why not all information is incorporated, such as mistakes in using Bayes' rule or short sale constraints, in which case we would obtain $\lambda_{\text{public}} \neq 1$.

Halim, Riyanto, and Roy (2019) runs both private and public information treatments. Table 3 displays the $\lambda$ estimates for both treatments, as well as the difference. The share of incorporated information increases from 29% to 100% when information is public rather than private, a difference of 71 percentage points (which is significant at any conventional level). Thus, we cannot reject the hypothesis that the public information treatment of Halim, Riyanto, and Roy (2019) produces semistrong informationally efficient prices.

In addition to the private information treatments from Page and Siemroth (2017), we run additional public information treatments for this study, where (instead of an information acquisition stage) the computer randomly chose 10 or 20 signals to be shown to all traders prior to the double auction trading. The results in Table 3 show that while the share of information incorporated increases relative to the private information treatments, the number is still very low for mean and median transaction prices at 15%.

**Table 3** The effect of private vs public information on $\lambda$

| Data source | Mean price | | | Median price | | | Last price | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\lambda$ | $\sigma$ | LL | $\lambda$ | $\sigma$ | LL | $\lambda$ | $\sigma$ | LL |
| Halim, Riyanto, and Roy (2019),<br>Public info (72 obs.) | $1.00^\dagger$<br>[0.70, 1.00] | 0.12<br>[0.08, 0.13] | 51.74 | $1.00^\dagger$<br>[0.79, 1.00] | 0.12<br>[0.09, 0.13] | 50.86 | $1.00^\dagger$<br>[0.83, 1.00] | 0.12<br>[0.09, 0.14] | 49.84 |
| New treatment<br>Public info (36 obs.) | $0.15^\dagger$<br>[0.05, 0.40] | 0.09<br>[0.04, 0.13] | 19.05 | $0.15^\dagger$<br>[0.05, 0.60] | 0.10<br>[0.05, 0.13] | 15.62 | $0.90^\dagger$<br>[0.20, 1.00] | 0.17<br>[0.06, 0.21] | 8.96 |
| Halim, Riyanto, and Roy (2019)<br>Private info (70 obs.) | 0.29<br>[0.12, 0.43] | 0.08<br>[0.06, 0.13] | 27.16 | 0.29<br>[0.15, 0.46] | 0.08<br>[0.05, 0.13] | 25.65 | 0.30<br>[0.19, 0.46] | 0.13<br>[0.06, 0.17] | 13.93 |
| Page and Siemroth (2017)<br>Private info (108 obs.) | 0.01<br>[0.01, 0.05] | 0.09<br>[0.06, 0.09] | 73.09 | 0.04<br>[0.01, 0.07] | 0.08<br>[0.06, 0.1] | 63.18 | 0.10<br>[0.05, 0.12] | 0.08<br>[0.06,0.12] | 40.09 |
| Mean difference<br>public vs private | 0.40<br>[0.01, 0.81] | | | 0.42<br>[0.01,0.81] | | | 0.63<br>[0.14,0.92] | | |

*Note:* The table displays the maximum likelihood estimates for the share of signals in the market used for pricing the asset ($\lambda$) and the standard deviation of the error distribution ($\sigma$). Estimates are based either on the observed mean, median, or last transaction prices. LL is the log-likelihood. The 95% confidence intervals below the estimates were determined by non-parametric percentile bootstrap. The mean difference is determined by bootstrapping the differences in $\lambda$ between the public and private information treatments, giving equal weight to all experiments, and then taking the mean of these differences.

$^\dagger$ Estimate is the largest in the likelihood maximizing interval (see 2.2.4 for explanation)

However, the estimate based on last transaction prices is 90%, a large improvement which corresponds to about 18 of 20 signals. Moreover, the last transaction prices are not significantly different from fully informationally efficient prices ($\lambda = 1$). Thus, in our experiment, prices in the public information treatments needed time to move in the right direction.[10] Such an improvement is often found in private information treatments, but it is somewhat surprising here, since the trades of others should not reveal any relevant information unlike in private information treatments, as all information was made public before trading. A recent theory explaining underreaction to public information (Ottaviani and Sørensen, 2015) does not seem to be at play here, since it relies on heterogeneous priors, while all experiments were designed with a common prior probability distribution.[11]

Result 2: *In both experiments, late transaction prices are semistrong-form informationally efficient, that is, $\lambda$ is not significantly different from one.*

The mean differences of $\lambda_{\text{public}}$ and $\lambda_{\text{private}}$ over both experiments are large and significant at about 40 percentage points for the mean and median transaction prices and at about 63 percentage points for the last transaction prices. Hence, we have strong evidence that the aggregation of private information is imperfect, as the share of signals that are incorporated into prices is larger in the public information treatments. Since the $\lambda$ estimates are based on experimental data, these differences have a causal interpretation.

Result 3: *The share of information incorporated into market prices ($\lambda$) is larger if information is public rather than private.*

Overall, the Halim, Riyanto, and Roy (2019) experiment suggests that information asymmetries are the only obstacle to achieving full informational efficiency, as their public information treatment prices incorporate all information ($\lambda = 1$). Our data also suggest that informational asymmetries play a role, but separate noninformational reasons prevent full informational efficiency, at least for early transaction prices.

## *3.4. What else affects informational efficiency?*

The previous section established that the share of information incorporated into prices, $\lambda$, is significantly larger if information is public compared to private information. In this section, we conduct further tests to identify or rule out potential causes that affect how much private information is incorporated into prices. This illustrates the strength of our approach, because existing methods can typically only reject full efficiency, but are unable to distinguish between degrees of inefficiency, which this section makes use of.

First, we test whether $\lambda$ differs in the early three rounds of the Page and Siemroth (2017) private information experiment compared to the last three rounds.[12] Subjects likely have more experience in the last three rounds, having completed at least nine rounds before, which

[10] Figure 3 in the Internet Appendix visualizes this improvement over time, displaying the absolute difference between the transaction prices and the full information price over the 180 seconds of the double auction trading window.

[11] But we cannot entirely dismiss this explanation, as one could still argue that subjects had heterogeneous subjective priors despite the common prior probability distribution communicated in the instructions.

[12] We use the Page and Siemroth (2017) data for the $\lambda$ difference tests, since traders had the most signals in this experiment, so if there are differences in $\lambda$ to be seen, it is there.

**Table 4** Difference tests of $\lambda$

| | Mean price $\Delta\lambda$ | Median price $\Delta\lambda$ | Last price $\Delta\lambda$ |
|---|---|---|---|
| Last 3 rounds vs. first 3 rounds | 0.02 [−0.09, 0.06] | 0 [−0.10, 0.07] | 0.06 [−0.18, 0.25] |
| Many informed traders vs. few informed traders | −0.02 [−0.05, 0.02] | −0.03 [−0.08, 0.02] | 0.03 [−0.11, 0.19] |
| High trade volume vs. low trade volume | −0.03 [−0.07, 0.01] | −0.06** [−0.11, −0.01] | 0.01 [−0.14, 0.06] |
| Shortsales allowed vs. shortsale constraints | 0 [−0.1, 0.2] | 0 [0, 0.1] | 0 [−0.1, 0.2] |
| Last transaction price vs. first transaction price | | 0.07*** [0.02, 0.11] | |

The table displays the average difference of $\lambda$ estimated from two subsamples, based on 1000 bootstrap draws. The short-sale test is based on the Fellner and Theissen (2014) data, the other tests are based on the Page and Siemroth (2017) data. The 95% confidence intervals of the differences were determined by non-parametric percentile bootstrap. *$p < .1$; **$p < .05$; ***$p < .01$.

might translate into more efficient prices. However, the estimates in Table 4 show that this is not the case: while the point estimate of the difference is up to six percentage points, these differences are not statistically different from zero.

Next, we test whether the share of signals $\lambda$ incorporated into prices changes by how many traders were informed; that is, they acquired extra signals in the Page and Siemroth (2017) experiment. We split the experimental data at the median to divide the sample in markets where many traders were informed in this sense and in markets in which few traders were informed. We find that the share $\lambda$ does not significantly differ by how many traders are informed.

Third, we test whether information incorporated into prices changes by trade volume, that is, the number of transactions. Again, we split the experimental data at the median in markets with above and below median trade volume. If anything, the significantly negative estimate on median prices is suggestive of lower trade volume being associated with more information incorporated into prices, but this result is not very robust with respect to the price measure. In addition to this median split analysis, we also computed $\lambda$ for every single market and determined its correlation with the number of transactions in those markets. The correlation coefficients are $r = \{-0.03, -0.11, -0.18\}$ for mean, median and last transaction prices, respectively. None of these correlation coefficients is significantly different from zero at the 5% level. Hence, overall, the evidence for an association of trade volume and private information incorporated into prices is weak.

Fourth, the Fellner and Theissen (2014) experiment varied whether short sales were allowed or not. The original paper found that short-sale constraints are associated with higher price levels. However, Table 4 shows that the presence of short-sale constraints does

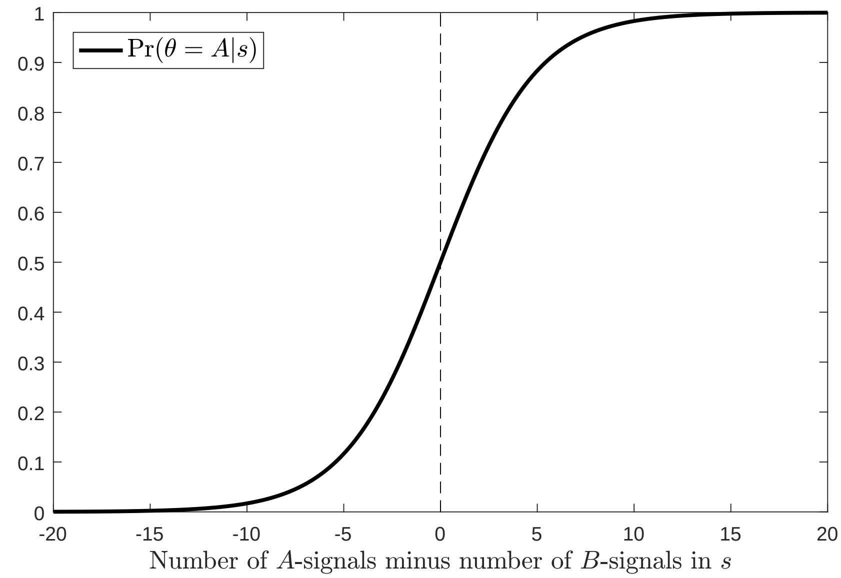



**Figure 2.** Posterior probability, $\Pr(A|s)$, depending on the number of $A$ minus $B$ signals in information set $s$

not translate into a difference in information incorporated into prices.

Finally, we test whether the information incorporated by the last transaction price of every round differs from the information incorporated into the first transaction price. We already found that $\lambda$ based on the last transaction prices tends to be larger than the estimates based on the mean or median transaction prices, so here we specifically test whether more information is aggregated over the course of the trading window. Table 4 confirms this hypothesis bears out. For the Page and Siemroth (2017) experiment, the difference is seven percentage points, which is significantly different from zero at the 1% level.

Result 4: *The share of information incorporated into market prices ($\lambda$) is larger in the last transaction price compared to the first transaction price, and does not differ early versus late rounds of the experiment, nor by how many traders are informed or whether short-sale constraints are present.*

## *3.5. From imperfect information aggregation to mispricing*

The posterior probabilities $\Pr(A|s)$ in (1), and thus prices based upon them, are concave in the number of $A$ signals (if $s$ contains more $A$ than $B$ signals) and in the number of $B$ signals (if $s$ contains more $B$ than $A$ signals), see Figure 2 for a plot. For this reason, the mispricing due to imperfect information aggregation is not as pronounced as the imperfect information aggregation itself, because the first signals move the posterior a lot more than the last signals.

To quantify the effect of imperfect information aggregation on prices, we define and estimate a second measure to quantify informational efficiency, which we call "price accuracy" (absence of mispricing) $\psi$, and which captures how close the observed prices are to full information prices. This measure differs from $\lambda$, which captures how close the information set consistent with observed market prices is to the full information set. In other words, $\psi$ is a measure in "price space," whereas $\lambda$ is a measure in "information space." So if we ask how much information prices in the market reflect, which refers to the underlying information, then $\lambda$ is the more appropriate measure. If we ask about the effects of information aggregation on prices, for example, how much money can be made due to imperfect information aggregation, then $\psi$ is the more appropriate measure, since differences in prices (and not differences in information sets) determine the scope for earning profits. Hence, both measures have their uses, and while they have some conceptual differences, both yield the result that private information is not completely reflected in prices.

The straightforward measure of price accuracy $\psi$ is the distance between the observed price $p_m$ and the prior (no) information price 0.5, relative to the distance between the full information price and the prior information price:

$$\psi_m = \min\left\{1, \frac{p_m - 0.5}{\text{Full Information Price}_m - 0.5}\right\}. \tag{5}$$

The minimum operator ensures that $\psi$ cannot exceed 1 for prices that overreact to information.[13] A larger $\psi$ indicates more price accuracy and less mispricing. A value of $\psi = 1$ corresponds to no underreaction. The price accuracy measure $\psi$ can be negative if the observed price and the full information price are on opposite sides of the prior price of 0.5. But empirically $\psi$ is almost always between zero and one, so we can compare the magnitudes of price accuracy $\psi$ and the share of incorporated information $\lambda$.[14]

While $\psi_m$ is the price accuracy for market $m$, we can estimate the average accuracy $\psi$ in an experiment—a linear statistic—via an ordinary least squares (OLS) regression as follows:

$$Y_m = \psi(\text{Full Information Price}_m - 0.5) + \varepsilon_m,$$

with

$$Y_m = \begin{cases} \min\{p_m - 0.5, \text{Full Information Price}_m - 0.5\} & \text{if Full Information Price}_m > 0.5, \\ \max\{p_m - 0.5, \text{Full Information Price}_m - 0.5\} & \text{if Full Information Price}_m < 0.5. \end{cases}$$

Table 5 displays the estimates for $\psi$. Clearly, for all private information treatments $\psi > \lambda$, that is, observed prices are closer to the ideal of full information prices than the information sets underlying the observed prices are to the ideal full information sets, which reflects the concavity of the posterior discussed above. However, all private information treatments have a price accuracy $\psi$ that is significantly smaller than one, or in other words, observed prices significantly differ from the full information prices.

[13]The minimum operator ensures that markets that overreact to information do not compensate for underreaction to information in other markets that might otherwise yield a perfect score of $\psi = 1$ on average, even if observed prices never actually equal full information prices.

[14]Several other mispricing measures have been used in the literature, especially when analyzing price bubbles. For example, the relative absolute deviation measure from Stöckl, Huber, and Kirchler (2010) relates observed prices to fundamental values. While these measure mispricing from a single fundamental value very well, we instead want to relate deviations of observed prices from full information prices relative to prior information prices, yielding an index between zero and one.

**Table 5** OLS estimates of price accuracy $\psi$ based on data from the literature

| Data source | Mean price | | Median price | | Last price | |
|---|---|---|---|---|---|---|
| | $\psi$ | LL | $\psi$ | LL | $\psi$ | LL |
| **Private information** | | | | | | |
| Deck, Lin, and Porter (2013) | 0.13 | 11.92 | 0.12 | 10.92 | 0.24 | 8.83 |
| (40 obs.) | $[-0.05, 0.31]$ | | $[-0.05, 0.30]$ | | $[0.04, 0.44]$ | |
| Fellner and Theissen (2014) | 0.45 | 86.38 | 0.49 | 54.03 | 0.54 | 9.06 |
| (318 obs.) | $[0.38, 0.51]$ | | $[0.42, 0.56]$ | | $[0.46, 0.61]$ | |
| Page and Siemroth (2017) | 0.17 | 78.13 | 0.17 | 64.77 | 0.30 | 48.18 |
| (108 obs.) | $[0.11, 0.22]$ | | $[0.10, 0.23]$ | | $[0.22, 0.36]$ | |
| Su and Wang (2017) | 0.20 | 17.21 | 0.18 | 16.01 | 0.20 | 12.86 |
| (20 obs.) | $[0.04, 0.36]$ | | $[0.01, 0.34]$ | | $[0.01, 0.38]$ | |
| Halim, Riyanto, and Roy (2019) | 0.38 | 41.25 | 0.40 | 40.72 | 0.44 | 27.94 |
| (70 obs.) | $[0.28, 0.49]$ | | $[0.29, 0.50]$ | | $[0.29, 0.56]$ | |
| **Public information** | | | | | | |
| Halim, Riyanto, and Roy (2019) | 0.89 | 105.13 | 0.90 | 104.84 | 0.91 | 114.75 |
| (72 obs.) | $[0.83, 0.93]$ | | $[0.84, 0.94]$ | | $[0.85, 0.96]$ | |
| New (Page & Siemroth) | 0.34 | 33.84 | 0.37 | 28.64 | 0.58 | 25.92 |
| (36 obs.) | $[0.23, 0.47]$ | | $[0.22, 0.52]$ | | $[0.42, 0.74]$ | |

The table displays the OLS estimates for the price accuracy measure $\psi$, see (5). Estimates are based on the observed mean, median, or last transaction price. LL is the log-likelihood. The 95% confidence intervals below the estimates were determined by nonparametric percentile bootstrap.

Result 5: *In all experiments, observed prices underreact to private information ($\psi < 1$).*

Considering only private information treatments, the average $\psi$ exceeds the average $\lambda$ by a factor of about 2.6 for mean transaction prices. So while the information incorporated into prices is about 10% over all experiments, the price accuracy is about 26%. Moreover, for median transaction prices $\psi$ exceeds $\lambda$ by a factor of 2.1 and for last transaction prices $\psi$ exceeds $\lambda$ by a factor of 2.1 over all private information treatments. Consequently, the mispricing due to imperfect information aggregation is smaller than the low $\lambda$ estimates may suggest. However, our estimates also show that other studies that only investigate prices will considerably overestimate the informational content of prices.

Result 6: *Over all private information treatments, price accuracy is larger than the share of incorporated information by a factor of more than two, that is, $\psi > 2\lambda$.*

Finally, and unsurprisingly, there is generally less mispricing/more price accuracy $\psi$ with public information compared to private information, in line with the results based on $\lambda$.

Result 7: *Prices are closer to full information prices if information is public rather than private.*

Overall, the $\psi$ estimates in the private information treatments—all significantly below 62%—suggest that the mispricing is substantial.

# 4. Discussion

## *4.1. Experiments: Why is private information aggregation in prices inefficient?*

Since our private information estimates indicate a surprisingly inefficient information aggregation in prices, a natural question is why the information aggregation does not work as predicted by fully revealing REE. The experiments we used here were not designed to test possible explanations, so future work is called for where potential explanations are rigorously tested. Nevertheless, we will discuss what our estimates and other experimental work tell us about the possible causes of inefficient information aggregation in prices.

First, is the problem that subjects have trouble using Bayes' rule appropriately when evaluating their information? Bayes' rule is necessary to compute the correct probability distribution and expected asset value conditional on any given information set. Two types of bias could explain our low $\lambda$ estimates. One possibility could be that subjects underweight the information they have, so that they trade at nonextreme prices closer to the "0% information" price. Another possibility could be that traders avoid trading, because they can't figure out the probabilities of the different outcomes and are ambiguity averse Asparouhova et al. (2015). The comparison of our public and private information market estimates suggest that such failures to do Bayesian updating are not the cause of the low $\lambda$ estimates. Bayes' rule needs to be correctly used in both of these settings, whether information is public or private, and our finding is that it works very well with public information, but not nearly as well with private information.

Second, the comparison of public versus private information markets also suggest that the market trading rules (i.e., rules of the double auction) are not the cause, as the same trading rules applied in both public and private information markets, and informational efficiency is very high in the public information markets. This conclusion is echoed by the analysis of an exogenous variation of short-sale constraints in Section 3.3.4, which found no effect on $\lambda$.

Third, our findings from Section 3.3.4 suggest that subject experience with the experimental market does not seem to play a major role, as the $\lambda$ estimates did not significantly improve over the many rounds of the experiment. This might suggest a fundamental difficulty in attaining full private information revelation, which cannot be overcome with more experience. Section 4.4.3 discusses a conceptual problem with fully revealing REE that might be a cause for this.

Fourth, a comparison with other experiments from the literature suggests that the complexity of the trader information structure plays a role. More specifically, if the information structure requires traders to infer other traders' information from prices to combine it with their own, as in our case, then the informational content of prices is low. If the information structure is simpler and does not require this inference from prices, then it can work better under some conditions.

In our experiments traders receive noisy signals of a binary state, which provide probabilistic information about the two possible states, but never reveal or rule out a state. In other asymmetric information experiments with different information structures, which rule out some states, information aggregation appears to work better (e.g., Plott and Sunder, 1988; Choo, Kaplan, and Zultan, 2019). Consider the design of Plott and Sunder (1988), series B. There are three states of the world $\{A, B, C\}$, and for each there is a market with an Arrow security that pays one if that state realized, and zero otherwise. If the realization of the state is $C$, then half of the traders in the market get the private signal "not $A$" and the other half of the traders get the private signal "not $B$." Thus, no single trader perfectly knows the state, but the pooled private information perfectly reveals the state, and perhaps more importantly, a single private signal *rules out one state*. Consequently, this signal structure is simpler than the one we consider, as it provides some certainty.

It is easy to see why information aggregation might be more successful in these cases, as these information structures allow for riskless arbitrage based on private information: If trader $i$ has a signal that the state is not $A$ (but does not know whether the state is $B$ or $C$), then $i$ can (short-)sell the $A$-Arrow security whenever there is a positive price, and will make a certain profit doing so. Other traders can arbitrage away positive prices for the remaining "incorrect" Arrow security. Hence, positive prices would remain only for the asset corresponding to the actual state. The crucial difference to the information structure used here is that it does not require traders to infer other trader information from prices, in order to combine it with one's own, which is however required with our information structure to reach REE prices.

But both the original Plott and Sunder (1988) series A and follow-up papers (e.g., Corgnet, DeSantis, and Porter, 2015) show that such simpler information structures are *not sufficient* for fully revealing REE prices. If the market trades an asset that maps the three states into different values of that one asset, instead of trading a full set of Arrow securities as described above, then information aggregation fails to reach REE levels. Thus, in accordance with the experimental evidence, both a specific set of assets *and* a specific information structure appears to be necessary for full informational efficiency, which is therefore the exception rather than the rule in the lab. Nevertheless, the comparison of the different information structures suggests that the inference from prices appears to be a cause of the informational inefficiency, although more work is needed to confirm this.

## *4.2. Implications for asymmetric information asset pricing theory*

The standard asset pricing model in our context is the fully revealing rational expectations equilibrium (REE, e.g., Grossman, 1976),[15] which replaced the earlier Walrasian equilibrium. Let the realization of the asset value be $\theta$, the vector of all (public and private) signals in the market about $\theta$ be $\boldsymbol{s} \in \boldsymbol{S}$, and index traders by $i \in I$. In a REE, risk-neutral trader net demand $x_i$ with private signals $s_i$ and price $p$ maximizes

$$x_i = \arg\max_x \mathbb{E}[(\theta - p)x | s_i, p = R(\boldsymbol{s})] \text{ s.t. budget constraints } \forall i \in I,$$

$$\text{with } R(\boldsymbol{s}) \text{ such that } \sum_{i \in I} x_i(s_i, p = R(\boldsymbol{s})) = 0 \; \forall \boldsymbol{s} \in \boldsymbol{S},$$

[15] Since the experiments feature neither a random supply of assets nor explicit noise trader demand, the noisy rational expectations equilibrium concept does not apply in our setting by design.

that is, traders take into account the information contained in prices $p$, and the REE price function $R(\boldsymbol{s})$ features $\lambda = 1$ (e.g., Radner, 1979). Our estimates clearly reject this equilibrium notion. Below, in Section 4.4.3, we also discuss a conceptual problem of this equilibrium concept.

In the older Walrasian equilibrium concept (which Plott and Sunder (1988) call "prior information equilibrium"), trader net demand $x_i$ maximizes

$$x_i = \arg\max_x \mathbb{E}[(\theta - p)x|s_i] \text{ s.t. budget constraints } \forall i \in I,$$
$$\text{with } W(\boldsymbol{s}) \text{ such that } \sum_{i \in I} x_i(s_i, p = W(\boldsymbol{s})) = 0 \ \forall \boldsymbol{s} \in \boldsymbol{S},$$

that is, traders' evaluations of the asset only depends on their own private signals $s_i$, and information contained in prices $p = W(\boldsymbol{s})$ is ignored. As a consequence, the Walrasian equilibrium price function $W(\boldsymbol{s})$ cannot be more extreme than the most extreme private evaluation among all traders, that is, cannot be larger than $\max_i \mathbb{E}[\theta|s_i]$ nor smaller than $\min_i \mathbb{E}[\theta|s_i]$, as such extreme prices would not clear the market. Thus, Walrasian equilibrium predicts a lower share of information in prices $\lambda$ than REE.

More precisely, if all traders have the same number of i.i.d. private signals (as in three of the five experiments we use), and we have $T$ traders in the market, then the Walrasian equilibrium predicts at most $\lambda = 1/T$. That is, its price is based on the amount of information that a single trader has, because there is no information aggregation via the price. Interestingly, considering Table 2, this is remarkably close to some of the estimates we have. The Deck, Lin, and Porter (2013) experiment had eight traders, and our estimates for the mean, median, and last transaction prices are $\lambda = 1/8$. For the Fellner and Theissen (2014) experiment with 10 traders, the estimate for mean prices is $\lambda = 0.1$, although the estimates for median and late transaction prices are higher.

So while the Walrasian equilibrium concept does not explain the information contained in late transaction prices well, it is very close for the mean prices of some of the experiments, and certainly closer than the REE prediction. Nevertheless, we can reject the Walrasian equilibrium concept on the basis that late transaction prices have significantly more information than early transaction prices (Section 3.3.4), while there should be no belief updating in this equilibrium concept via prices.

A recently proposed behavioral equilibrium concept is the cursed expectations equilibrium (CEE, Eyster, Rabin, and Vayanos, 2019). The concept is motivated by ample experimental evidence that suggests that people do not infer all private information that is revealed by the observable actions of others, and prices in financial markets are a function of the (trading) actions of others. In CEE, risk-neutral trader net demand maximizes[16]

$$x_i = \arg\max_x \mathbb{E}[(\theta - p)x|s_i]^{\chi_i} \mathbb{E}[(\theta - p)x|s_i, p = C(\boldsymbol{s})]^{1-\chi_i} \text{ s.t. budget constraints } \forall i \in I,$$
$$\text{with } C(\boldsymbol{s}) \text{ such that } \sum_{i \in I} x_i(s_i, p = C(\boldsymbol{s})) = 0 \ \forall \boldsymbol{s} \in \boldsymbol{S},$$

for "cursedness parameters" $\chi_i$. That is, depending on $\chi_i$, trader $i$ fully infers all information from prices as in REE if $\chi_i = 0$, fully ignores all information in prices as in Walrasian

[16] For simplicity, we omit the additional random endowment shocks present in the model of Eyster, Rabin, and Vayanos (2019). These shocks are not studied or featured in our experiments.

equilibrium if $\chi_i = 1$, or something in between ($\chi_i \in (0, 1)$).[17] Consequently, CEE can explain everything from the Walrasian outcome $\lambda = 1/T$ to the REE outcome $\lambda = 1$, depending on the values of $\chi_i$ for each $i$. Thus, CEE is more flexible than the older equilibrium concepts and can explain the range of our estimates. Importantly, both CEE and Walrasian equilibrium are consistent with our estimate of $\lambda = 1$ with *public* information, and consistent with our estimates of $\lambda < 1$ with *private* information, while REE predicts $\lambda = 1$ in both cases.

Informally, for $\chi_i \in (0, 1)$, CEE says that information gets into prices because traders put more weight on their own private information, compared to REE, when trading: if they have positive signals they are more willing to buy, if they have more negative signals they are more willing to sell. But this implies underreaction of CEE prices to information, because a few traders with a negative signal, while all other traders have positive signals, will drag the price down compared to REE prices. That is, traders hang too much onto their own information and they underappreciate the information of others contained in market prices. And because traders have different beliefs about the asset value even after making inferences from market prices, CEE also tends to imply more trading volume than REE.

Overall, our estimates suggest that the inference of information from prices does not work as well as predicted by REE, although there is some learning from prices. CEE is one model that can account for these findings. But more work is needed to identify the exact mechanisms and individual level inferences that are made by traders, for example, by eliciting beliefs during or after trading. Moreover, given its flexibility, it is difficult to reject CEE based on $\lambda$ estimates alone, so more work is needed to test the additional predictions of CEE that are independent of its free parameters $\{\chi_i\}_{i \in I}$. Still, CEE suggests one possible explanation for our estimates: Traders do not infer as much information from prices as they would in REE, possibly because of cognitive limitations or "cursedness."

## *4.3. Are fully revealing REE prices realistic in practice? A conjecture*

Why do prices not incorporate all private information as prescribed by REE, even in the most controlled lab environment where all distributions are common knowledge, the number of traders is perfectly known, the attention is focused on one task only, etc.? Instead of seeking a cause, one might question the benchmark: Is it realistic to expect prices to reach this level of informational efficiency in practice? Conceptual problems with the fully revealing REE model itself, which were raised early by game theorists, suggest this benchmark might be unreachable in practice. While the Grossman and Stiglitz paradox is probably the most well-known problem with fully revealing REE (FRREE), there are others. In particular, Dubey, Geanakoplos, and Shubik (1987) criticize FRREE for not actually explaining how information in the static model is incorporated into prices and the "circularity of reasoning regarding cause and effect and their timing." As they write, "[traders] are supposed to know the public information revealed by the prices *before* they act to form those prices." (Vives, 2008, p. 81) describes similar concerns regarding FRREE:

[17]The demand functions generated by maximizing the CEE objective are similar to noisy REE demand functions, even though there is no exogenous noise, in the sense that both depend on the trader's private information and on the market price. In noisy REE, the noise variance relative to the precision of the private signals scales how close prices get to fully revealing. In CEE, the cursedness parameters scale how close prices get to fully revealing.

> The concept of competitive FRREE is not without problems. At an FRREE, agents, by looking at the market price, know all they need to know about the uncertain state of the world to take action. This means that they will disregard their private signals. How does price reflect all the information agents have?

In fact, in such a FRREE "the price is fully revealing even though a trader's demand [...] is independent of the signals received" (Vives, 2014, p. 1208), so we cannot use the FRREE equilibrium demand functions to learn anything about what sort of trades should happen in practice to reach fully revealing prices. Vives further writes that some FRREE outcomes are not implementable; that is, they are not an equilibrium in any well-specified game that models the trading process in order to break the circularity mentioned by Dubey, Geanakoplos, and Shubik (1987).

Indeed, attempts to provide a game theoretical foundation for FRREE typically need additional assumptions for full revelation of private information to occur via prices, for example, private value components in the evaluation of assets and taking a limit (Reny and Perry, 2006; Vives, 2014) or dynamic trading with specific assets (Ostrovsky, 2012). The latter model in particular would suggest that we should expect $\lambda = 1$ at most at the end of the trading window, but not in the first trades. While we do find that $\lambda$ improves during trading in the experiments, even the last transactions are far from reflecting all private information.

Thus, these game theoretical critiques might suggest that the question should not be why we do not observe the fully revealing outcome, but rather whether it is possible if FRREE does not convincingly explain the *process* of how information is incorporated into prices.[18] A cause for our inefficiency results with private information therefore might be that perfect efficiency is not possible beginning with the first trade, even theoretically, when using stringent game theoretical criteria. In more practical terms, the reason fully revealing prices do not arise might be that individual traders do not have the information necessary to correct the mispricing, which, after all, is only mispricing based on the combined private information of all traders.

# 5. Survey of Beliefs among Academics

In his famous review of the efficient market hypothesis, Lo (2004) described EMH as "one of the most enduring ideas" in modern finance. After having been widely accepted in the 1970s, the efficient market hypothesis has had its validity heavily debated by academics, thanks to the rise of behavioral finance. However, economists tend to share the view that financial markets are quite efficient (Doran, Peterson, and Wright, 2010). This view possibly influences economists' trust in market prices across a wide range of situations, for instance, when assessing the risk of market bubbles (Krugman, 2009).

Our methodology allows us to measure how much information market experiments incorporate. The first results of our analyses surprised us. While we did not think that markets,

[18] Note that these critiques relate to the fully revealing REE. That they are largely resolved by the noisy REE, as in Grossman and Stiglitz (1980), explains very well how information is incorporated into prices (namely, by the aggressive positions taken by insiders to push the price in the right direction). However, as noted before, by design, noisy REE is not applicable to our setting.

in general, and experimental markets, in particular, are fully efficient, we nevertheless expected the markets we investigate to be fairly close to efficiency. Our surprise leads us to investigate the views of academic economists about informational efficiency.

To do so, we designed a short survey, asking economists to tell us how much information they think is incorporated into financial markets in general (question 1) and in experimental asset markets in particular (question 2, incentivized). We invited all participants from the Econometric Society meetings in 2017 in Australia/Asia, North America, and Europe, which are some of the largest general interest economics conferences on these continents, along with other economists to fill out the survey. Overall, 336 academic economists answered our survey (detailed description of the sample is in the Internet Appendix). We interpret the results of this survey as suggestive of economists' views about market efficiency.

To incentivize participants, we randomly select 10 respondents for payment (known to respondents beforehand), which was computed from respondent $i$'s answer to question 2 (denoted by $r_i$) with the quadratic scoring rule

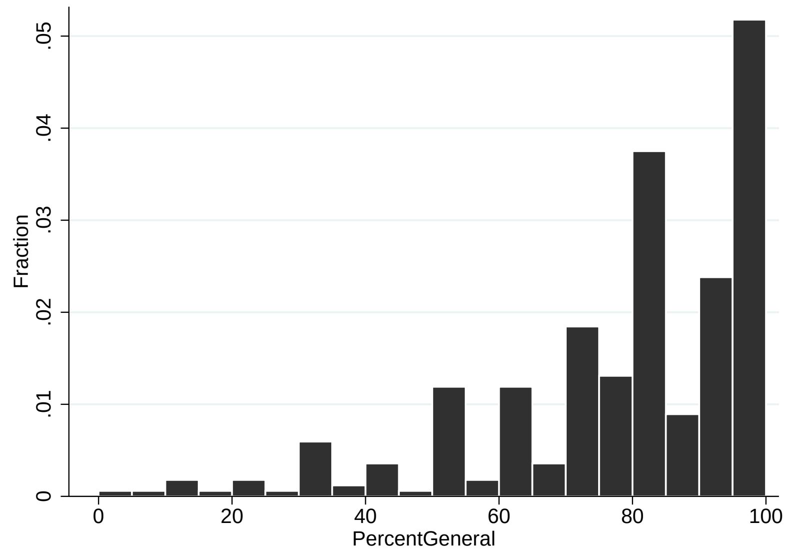
$$\text{Payment}(r_i) = 100 - 50(r_i/100 - \lambda)^2 \text{ (in \$)},$$

with payments ranging from a possible minimum of \$53.92[19] to a possible maximum of \$100.

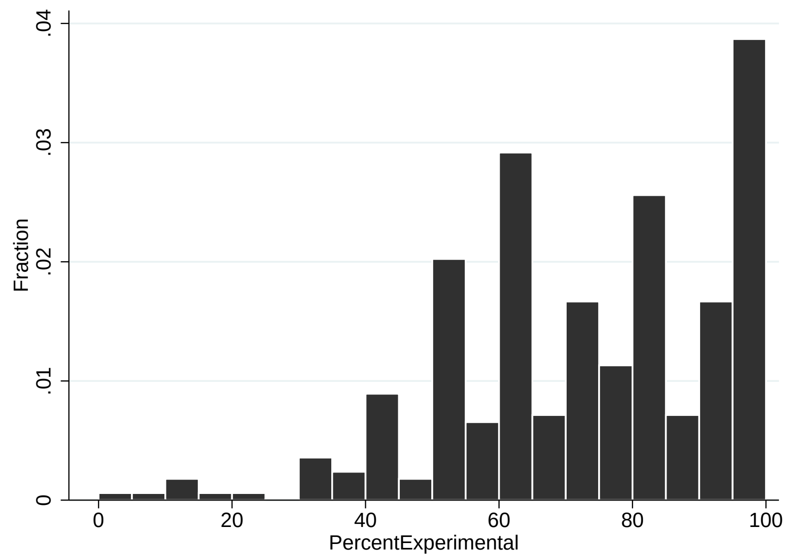
The survey comprises the following two main questions (secondary questions about demographics, etc., and explanations are omitted here but can be found in the complete survey in the Internet Appendix).

1. Consider financial markets with risky assets whose payoffs can have different values. Traders in the economy have information, public and private, about the value of these assets. How much of this information would you say is reflected in the price of the assets? Please specify a percentage number between 0 and 100. [variable: PercentGeneral]
2. Now consider a small laboratory financial market with 10–12 traders (see details). How much information do you think is reflected in these asset prices? Please specify a percentage number between 0 and 100. [variable: PercentExperimental]

[19]This minimum is computed based on our estimate $\lambda = 0.04$ from the Page and Siemroth (2017) median prices data, which was our main estimate at the time of the survey.




**(a)** Distribution of beliefs how much information is incorporated into financial market prices in general




**(b)** Distribution of beliefs how much information is incorporated into experimental financial market prices

**Figure 3.** Distribution of responses to questions 1 and 2

Academics on average believe that 77% of information is incorporated into financial market prices in general (PercentGeneral) and that 71% of information is incorporated into experimental asset prices (PercentExperimental). Figures 3a and 3b displays histograms of the beliefs.

We observe that the average belief is significantly larger for real-world markets compared to experimental markets ($t = 4.288$, $df = 335$, $p < .0001$). However, about 32% of academics

**Table 6** OLS regression explaining the survey responses.

| | PercentGeneral | PercentExperimental |
|---|---|---|
| AcademicEconomics | -2.421 | 5.917 |
| | (3.253) | (3.810) |
| MethodExperimental | -0.520 | -7.814** |
| | (3.385) | (3.203) |
| MethodTheory | -0.441 | -4.452 |
| | (2.723) | (2.751) |
| Constant | 82.80 | 63.78 |
| | (6.204) | (7.343) |
| Demographic controls | Yes | Yes |
| *N* | 317 | 317 |

Standard errors are indicated in brackets.

believe that experimental markets incorporate more information than real-world markets.

We also observe that the beliefs of academics are much more optimistic about market efficiency than our results suggest. The average belief is considerably larger than the largest estimate of *private* information incorporated from the data ($\lambda = 0.3$), and this difference is significant at any level ($t = 36.28$, $df = 335$, $p < .0001$).[20] Moreover, only 4% of academics believe that experimental markets incorporate 30% or less of private information.

Thus, these results suggest that economists overestimate the share of private information that is incorporated into experimental asset markets. This might have important implications for the interpretation of academic policy advice, as financial market prices may not be as informed as often believed. Consequently, regulations and policies that depend on market information, such as contingent capital with market triggers (e.g., Flannery, 2016), may be less effective than previously thought.

Finally, we run a multiple regression using the demographic responses to explain the responses PercentGeneral and PercentExperimental. The reference category in Table 6 is finance researcher and empirical method, and all the coefficients are relative to this group. Interestingly, experimentalists are significantly more pessimistic about the ability of experimental markets to incorporate information than empirical researchers. Economists on average do not hold significantly different beliefs from finance researchers, either about financial markets in general or about experimental markets.

# 6. Concluding Remarks

In this paper, we present a new method to estimate how much information is incorporated into experimental asset market prices. While we derive all expressions for the specific case of binary options whose data we use in the estimations, our method of pricing based on random

[20]The average belief is also significantly different from the maximal price accuracy estimate (based on private information treatments), $\psi = 0.54$, at any level ($t = 15.28$, $df = 335$, $p < .0001$).

signal-subsets can be applied to other assets and information structures with appropriate adaptations. Thus, on a methodological level, our contribution is to introduce an approach to quantifying the information contained in prices.

We use our method in a meta-analysis of 664 markets from five different experiments from the literature. Our most important finding is that surprisingly little private information is reflected in prices—for most experiments less than 30%. Our private information estimates about how much information is incorporated are considerably smaller than the beliefs of most economists that participated in our incentivized survey, which indicates that economists might overestimate the ability of financial markets to aggregate private information. However, we also show that markets are very successful in incorporating public information into prices, and late transaction prices are statistically not distinguishable from full information prices. Thus, our estimates indicate that experimental markets may be semistrong informationally efficient, but not even close to strong-form efficient.

The implications of that latter finding for field markets will depend on the informational environment. In situations in which a lot of independent pieces of private information are dispersed among a lot of traders, then even an inefficient aggregation of that information in prices would suggest that prices are nevertheless very informative about the underlying state. However, the lab is a very clean environment, where the number of traders, the structure of the information, and the prior distribution are all common knowledge, so traders in the lab know many things about the market and trading environment that they might not know in field markets. The lack of aggregation of private information in the lab raises questions about the ability of field markets to do so in less transparent environments.

In addition, field markets may not necessarily feature a lot of private independent pieces of information. Traders may have private information which is correlated or even shared in groups or networks of traders. In that case, even if a field market features a lot of traders, there may actually not be a similarly large number of independent pieces of private information to aggregate in the price. In such situations, our findings would suggest that prices are only very noisy signals of the underlying private information. Hence, suboptimal policy or investment decisions might be made based on these prices.

Besides the surprisingly small amount of private information being aggregated in prices, the main takeaways from our experimental findings on informational efficiency for field markets can be summarized as follows. First, it takes time for information to be reflected in prices, so the inefficiency is worse immediately after private information release. Consequently, early movers with private information stand to make larger profits than late movers. Second, short-sale constraints do not significantly affect the informational content of prices (Section 3.3.4), which one might have expected as short sales in theory help to integrate negative signals. This does not mean, however, that short-sale constraints do not affect prices; previous research has established they can affect price levels (Fellner and Theissen, 2014). Third, trading volume does not correlate significantly with the share of private information in prices, suggesting that it is not trading constraints that prevent some traders from "getting their information in." These results illustrate how our method to quantify the informational content of prices can be used to identify market characteristics that improve informational efficiency.

But we have only started to investigate what drives informational inefficiency, and more research is called for on the causes of the inefficient information aggregation, and on interventions that could help to lessen inefficiency. A next step will be to investigate in more detail the strategic and behavioral roots of our inefficiency result. While we argue that rational expectations equilibrium cannot explain the observed outcomes, cursed expectations equilibrium can, but the latter cannot be falsified just based on the informational content of prices that we estimate. Hence, further tests of this theory and related ones are required before we can be confident that "cursedness" is a cause of the failure of strong-form informational efficiency. Furthermore, future research is needed to identify how different market structures and information structures may make information aggregation easier or harder. Such steps hold promise in greatly improving our understanding of how financial markets aggregate private information.

# REFERENCES

Arrow, K., R. Forsythe, M. Gorham, R. Hahn, R. Hanson, J. Ledyard, S. Levmore, R. Litan, P. Milgrom, F. Nelson, et al. 2008. The promise of prediction markets. *Science* 320:877–8.

Asparouhova, E., P. Bossaerts, J. Eguia, and W. Zame. 2015. Asset pricing and asymmetric reasoning. *Journal of Political Economy* 123:66–122.

Asparouhova, E., P. Bossaerts, and W. Yang. 2017. Costly information acquisition in decentralized markets: An experiment Working paper.

Biais, B., D. Hilton, K. Mazurier, and S. Pouget. 2005. Judgemental overconfidence, self-monitoring, and trading performance in an experimental financial market. *The Review of Economic Studies* 72:287–312.

Boleslavsky, R., D. L. Kelly, and C. R. Taylor. 2017. Selloffs, bailouts, and feedback: Can asset markets inform policy? *Journal of Economic Theory* 169:294–343.

Bond, P., and I. Goldstein. 2015. Government intervention and information aggregation by prices. *The Journal of Finance* 70:2777–812.

Bossaerts, P. 2009. The experimental study of asset pricing theory. *Foundations and Trends in Finance* 3:289–361.

Bossaerts, P., C. Frydman, and J. Ledyard. 2013. The speed of information revelation and eventual price quality in markets with insiders: comparing two theories. *Review of Finance* 18:1–22.

Choo, L., T. R. Kaplan, and R. Zultan. 2019. Information aggregation in arrow–debreu markets: an experiment. *Experimental Economics* 22:625–52.

Cipriani, M., and A. Guarino. 2009. Herd behavior in financial markets: an experiment with financial market professionals. *Journal of the European Economic Association* 7:206–33.

Copeland, T. E., and D. Friedman. 1987. The effect of sequential information arrival on asset prices: An experimental study. *The Journal of Finance* 42:763–97.

Corgnet, B., M. DeSantis, and D. Porter. 2015. Revisiting information aggregation in asset markets: Reflective learning & market efficiency. Working Paper, Chapman University, Economic Science Institute.

Cowgill, B., and E. Zitzewitz. 2015. Corporate prediction markets: Evidence from google, ford, and firm x. *The Review of Economic Studies* 82:1309–41.

Deck, C., S. Lin, and D. Porter. 2013. Affecting policy by manipulating prediction markets: Experimental evidence. *Journal of Economic Behavior & Organization* 85:48–62.

Deck, C., and D. Porter. 2013. Prediction markets in the laboratory. *Journal of Economic Surveys* 27:589–603.

Doran, J. S., D. R. Peterson, and C. Wright. 2010. Confidence, opinions of market efficiency, and investment behavior of finance professors. *Journal of Financial Markets* 13:174–95.

Dreber, A., T. Pfeiffer, J. Almenberg, S. Isaksson, B. Wilson, Y. Chen, B. A. Nosek, and M. Johannesson. 2015. Using prediction markets to estimate the reproducibility of scientific research. *Proceedings of the National Academy of Sciences* 112:15343–7.

Dubey, P., J. Geanakoplos, and M. Shubik. 1987. The revelation of information in strategic market games: A critique of rational expectations equilibrium. *Journal of Mathematical Economics* 16:105–37.

Eyster, E., M. Rabin, and D. Vayanos. 2019. Financial markets where traders neglect the informational content of prices. *Journal of Finance* forthcoming.

Fama, E. F. 1970. Efficient capital markets: A review of theory and empirical work. *The Journal of Finance* 25:383–417.

———. 1998. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics* 49:283–306.

Fellner, G., and E. Theissen. 2014. Short sale constraints, divergence of opinion and asset prices: Evidence from the laboratory. *Journal of Economic Behavior & Organization* 101:113–27.

Flannery, M. J. 1998. Using market information in prudential bank supervision: A review of the us empirical evidence. *Journal of Money, Credit and Banking* 273–305.

———. 2016. Stabilizing large financial institutions with contingent capital certificates. *Quarterly Journal of Finance* 6:1650006–.

Flannery, M. J., J. F. Houston, and F. Partnoy. 2010. Credit default swap spreads as viable substitutes for credit ratings. *University of Pennsylvania Law Review* 2085–123.

Forsythe, R., and R. Lundholm. 1990. Information aggregation in an experimental market. *Econometrica* 309–47.

Forsythe, R., T. R. Palfrey, and C. R. Plott. 1982. Asset valuation in an experimental market. *Econometrica* 537–67.

Frechette, G. R. 2015. Laboratory experiments: Professionals versus students. In G. R. Frechette and A. Schotter, eds., *Handbook of Experimental Economic Methodology*, 360–90. Oxford University Press.

Gillette, A. B., D. E. Stevens, S. G. Watts, and A. W. Williams. 1999. Price and volume reactions to public information releases: An experimental approach incorporating traders' subjective beliefs. *Contemporary Accounting Research* 16:437–79.

Goeree, J. K., and J. Zhang. 2015. Inefficient markets Working paper.

Greenspan, A. 2001. Harnessing market discipline. *The Region* 6–7.

Grossman, S. 1976. On the efficiency of competitive stock markets where trades have diverse information. *Journal of Finance* 31:573–85.

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

Halim, E., Y. Riyanto, and N. Roy. 2019. Costly information acquisition, social networks and asset prices: Experimental evidence. *Journal of Finance* forthcoming.

Hanson, R., R. Oprea, and D. Porter. 2006. Information aggregation and manipulation in an experimental market. *Journal of Economic Behavior & Organization* 60:449–59.

Hayek, F. A. 1945. The use of knowledge in society. *The American Economic Review* 519–30.

Huber, J., M. Angerer, and M. Kirchler. 2011. Experimental asset markets with endogenous choice of costly asymmetric information. *Experimental Economics* 14:223–40.

Kirchler, M. 2009. Underreaction to fundamental information and asymmetry in mispricing between bullish and bearish markets. an experimental study. *Journal of Economic Dynamics and Control* 33:491–506.

Krugman, P. 2009. How did economists get it so wrong? *New York Times* 2:2009–.

Lim, K.-P., and R. Brooks. 2011. The evolution of stock market efficiency over time: a survey of the empirical literature. *Journal of Economic Surveys* 25:69–108.

Lo, A. W. 2004. The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management* 30:15–29.

———. 2008. Efficient markets hypothesis. *The New Pelgrave Dictionary of Economics* .

Malkiel, B. G. 2003. The efficient market hypothesis and its critics. *The Journal of Economic Perspectives* 17:59–82.

———. 2005. Reflections on the efficient market hypothesis: 30 years later. *Financial Review* 40:1–9.

Ostrovsky, M. 2012. Information aggregation in dynamic markets with strategic traders. *Econometrica* 80:2595–647.

Ottaviani, M., and P. N. Sørensen. 2006. The timing of parimutuel bets Working paper.

———. 2009. Surprised by the parimutuel odds? *American Economic Review* 99:2129–34.

———. 2015. Price reaction to information with heterogeneous beliefs and wealth effects: Underreaction, momentum, and reversal. *American Economic Review* 105:1–34.

Page, L., and R. Clemen. 2013. Do prediction markets produce well calibrated probability forecasts? *Economic Journal* 123:491–513.

Page, L., and C. Siemroth. 2017. An experimental analysis of information acquisition in prediction markets. *Games and Economic Behavior* 101:354–78.

Pennock, D. M., S. Lawrence, C. L. Giles, F. A. Nielsen, et al. 2001. The real power of artificial markets. *Science* 291:987–8.

Plott, C. R., and S. Sunder. 1982. Efficiency of experimental security markets with insider information: An application of rational-expectations models. *Journal of Political Economy* 90:663–98.

———. 1988. Rational expectations and the aggregation of diverse information in laboratory security markets. *Econometrica* 1085–118.

Radner, R. 1979. Rational expectations equilibrium: Generic existence and the information revealed by prices. *Econometrica* 655–78.

Reny, P. J., and M. Perry. 2006. Toward a strategic foundation for rational expectations equilibrium. *Econometrica* 74:1231–69.

Samuelson, P. A. 1965. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review* 6:41–.

Schwaiger, R., M. Kirchler, F. Lindner, and U. Weitzel. 2019. Determinants of investor expectations and satisfaction. a study with financial professionals. *Journal of Economic Dynamics and Control* .

Siemroth, C. 2019. The informational content of prices when policy makers react to financial markets. *Journal of Economic Theory* 179:240–74.

Smith, V. L. 2007. *Rationality in economics: Constructivist and ecological forms*. Cambridge University Press.

Smith, V. L., G. L. Suchanek, and A. W. Williams. 1988. Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica: Journal of the Econometric Society* 1119–51.

Snowberg, E., and L. Yariv. 2018. Testing the waters: Behavior across participant pools. Working Paper, National Bureau of Economic Research.

Stevens, D. E., and A. W. Williams. 2004. Inefficiency in earnings forecasts: Experimental evidence of reactions to positive vs. negative information. *Experimental Economics* 7:75–92.

Stöckl, T., J. Huber, and M. Kirchler. 2010. Bubble measures in experimental asset markets. *Experimental Economics* 13:284–98.

Su, M.-C., and J. T.-y. Wang. 2017. Just run sweepstakes - designing prediction markets to counter manipulation Working Paper.

Veiga, H., and M. Vorsatz. 2010. Information aggregation in experimental asset markets in the presence of a manipulator. *Experimental Economics* 13:379–98.

Vives, X. 2008. *Information and learning in markets: The impact of market microstructure.* Princeton University Press.

———. 2014. On the possibility of informationally efficient markets. *Journal of the European Economic Association* 12:1200–39.

Weitzel, U., C. Huber, J. Huber, M. Kirchler, F. Lindner, and J. Rose. 2019. Bubbles and financial professionals. *Review of Financial Studies* .

Wolfers, J., and E. Zitzewitz. 2004. Prediction markets. *Journal of Economic Perspectives* 18:107–26.

Yen, G., and C.-f. Lee. 2008. Efficient market hypothesis (emh): past, present and future. *Review of Pacific Basin Financial Markets and Policies* 11:305–29.