Speculation, Trading and Bubbles

April 4, 2013

Jose A. Scheinkman

Princeton University

Speculation, Trading and Bubbles
Third Annual Arrow Lecture

José A. Scheinkman [*]

Princeton University and NBER

April 4, 2013

Contents

1 Introduction

The history of financial markets is strewed with periods in which asset prices seem to vastly exceed fundamentals - events commonly called *bubbles*. Nonetheless, there is very little agreement among economists on what are the economic forces that generate such occurrences. Numerous academic papers and books have been written explaining why the prices attained in a particular episode can be justified by economic actors rationally discounting future streams of payoffs. Some proponents of efficient-markets even deny that one can attach any meaning to bubbles.[1]

Part of the difficulty stems from the fact that economists' discussions of bubbles often concentrate solely on the behavior of asset prices. The most common definition of a bubble is "a period in which prices exceed fundamental valuation". Valuation however depends on a view on fundamentals and efficient market advocates correctly point out that valuations are almost always ex post wrong. In addition, bubbles are frequently associated with periods of technological or financial innovations that are of uncertain value at the time of the bubble, making it possible, although often unreasonable, to argue that buyers were paying a price that corresponded to a fair valuation of future dividends, given the information at their disposal.

In this lecture I adopt the alternative approach of starting with a more precise model of asset prices that allows for divergence between asset prices and fundamental valuation and that has additional implications that are easier to test empirically. The model is based on the presence of fluctuating heterogeneous beliefs among investors and the existence of an asymmetry between the cost of acquiring an asset and the cost of *shorting* that same asset. The two basic assumptions of the model - differences in beliefs and higher costs of going short - are far from being standard in the literature on asset pricing. For many types of assets, including stocks, there are good economic reasons why investors should have more difficulty going short than going long, but most economic models assume no asymmetry. The existence

---

[1] e.g. Eugene Fama in Cassidy (2010) "I don't even know what a bubble means. These words have become popular. I don't think they have any meaning."

of differences in beliefs is thought to be obvious for the vast majority of market practitioners, but economists have produced a myriad of results showing that investors cannot agree to disagree. One implication of "cannot agree to disagree" results is that differences in private information *per se* do not generate security transactions, since agents learn from observing security prices that adjust to reflect the information of all parties. Arrow (1986) appropriately calls this implication "[a conclusion] flatly contrary to observation[2]. Because they are not standard, I discuss in Section 4 below some empirical evidence supporting these two central assumptions of the model.

Heterogeneous beliefs make it possible for the coexistence of optimists and pessimists in a market. The cost asymmetry between going long and going short an asset implies that optimists' views are expressed more fully than pessimists' views in the market and thus even when opinions are on average unbiased, prices are biased upwards. Finally, fluctuating beliefs give even the most optimistic the hope that, in the future, an even more optimistic buyer may appear. Thus a buyer would be willing to pay more than the discounted value she attributes to an asset" future payoffs, because the ownership of the asset gives her the option to resell the asset to a future optimist.

The difference between what a buyer is willing to pay and her valuation of the future payoffs of the asset - or equivalently the value of the resale option - is identified as a bubble.[3] An increase in the volatility of beliefs increases the value of the resale option, thus increasing the divergence between asset prices and fundamental valuation, and also increases the volume of trade. Hence, in the model, bubble episodes are associated with increases in trading volume. As we argue in Section 3.1 below, the connection between high trading volume and bubbles is a well established stylized fact. This relationship between bubbles and trading distinguishes models of bubbles based on heterogeneous beliefs and cost asymmetries from "rational bubble" theories.[4] A rational bubble is characterized by a continuous rise in an asset's price. Investors are content to hold the asset at the current price, because they believe that they are compensated for any risk of the bubble bursting by a suitable expected rate of price increase. In contrast to models based on heterogeneous beliefs and costly short selling, rational bubble theories fail to explain the association between bubbles and high trading volume

---

[2] page S398

[3] A related definition of bubbles that has been used in the literature, is of episodes in which buyers purchase an asset not on the basis of payoffs that the asset would generate, but because she intends to resell it at a higher price in the future (Brunnermeier (2008)).

[4] *e.g.* Santos and Woodford (1997)

and cannot be invoked to explain bubbles in assets that have a known value at a certain future date T, such as many credit instruments.[5]

Market prices are determined at each point in time by the amount that the marginal buyer is willing to pay for the asset. When beliefs are not homogeneous, this marginal buyer is the least optimist investor that is still a buyer of the asset. An increase in the capacity of individual investors to buy the asset, perhaps through increased leverage, allows for more extreme optimists to acquire the full supply of the asset at any point, and thus increases the value of the resale option. In a similar manner, when investors have limited capital and restricted access to leverage or limited capacity to bear risk, an increase in the supply of the asset is accompanied by a less optimistic marginal buyer. Thus the valuation that the marginal buyer has of future payoffs declines as supply increases, because the marginal buyer attributes a smaller fundamental value to the asset. However, a buyer also knows today that because of the larger supply that needs to be absorbed, future marginal buyers are likely to be relatively less optimistic and thus the value of the resale option also declines. Hence an increase in the supply of the asset that is unexpected by current holders of the asset diminishes the difference between the price and the fundamental valuation of the marginal buyer - that is it diminishes the size of the bubble. In Section 3.2 below I argue that increases in asset supply helped implode some well known bubbles.

Robert Shiller's influential *Irrational Exuberance*[6] postulates that bubbles result from feedback mechanisms in prices that amplify some initial "precipitating factors". The model in this lecture ignores the effect of this endogenous price dynamics just as it ignores the learning from prices used by rational theorists to dismiss the possibility of disagreement. It does however depend on precipitating factors that would generate optimism at least among some investors. Asset price bubbles often coincide with (over) excitement about a recent real or fake innovation[7] and for the purpose of this lecture one may think of "technological innovations", broadly construed, as the precipitating factors generating bubbles.

This lecture is organized as follows: In Section 2, I summarize some relevant facts concerning the South Sea Bubble, one of the earliest well documented occurrences of a bubble. In Section 3, I present some evidence

---

[5]For if everyone agrees on the value at T, rational investors would refuse to pay at time $T-1$ any price above the discounted value at T. Thus there would be no bubble at time $T-1$. Repeating this reasoning one concludes that a bubble never arises.

[6]Shiller (2006).

[7]See Section 3.3 for some examples.

on the three stylized facts that inspire the model in this lecture - that asset price bubbles coincide with increases in trading volume, that asset price bubble deflation seem to match with increases in an asset's supply and that asset price bubbles often occur in times of financial or technological innovation. In Section 4, I discuss some evidence for the assumption of costly short selling and for the role of overconfidence in generating differences in beliefs. Section 5 presents an informal sketch of the model and a discussion of related issues such as the effect of leverage, the origin of optimism and the role of corporations in sustaining bubbles. I summarize some empirical work that provide evidence for the model in Section 6 and present some concluding thoughts in Section 7. A formal model is exposited in the Appendix.

2 An example: The South Sea bubble

One of the earliest well documented occurrences of a bubble was the extraordinary rise and fall of the prices of shares of the South Sea Company and other similar joint-stock companies in Great Britain in 1720. At its origins in 1710, the South Sea Company had been granted a monopoly to trade with Spain's South American colonies. However, during most of the early 18th century Great Britain was at war against Spain's Phillip V and the South Sea Company never did much goods trading with South America, although it did achieve limited success as a slave trader. The real business of the South Sea Company was to exchange its stock for British government debt. The new equity owners would receive a liquid share with the right to perpetual annual interest payments in exchange for government debt that paid a higher interest rate but was difficult to trade. In the first months of 1720, the Company and its rival the Bank of England engaged on a competition for the right to acquire the debt of the British government. After deliberating for more than two months, the House of Commons passed the bill favoring the South Sea Company, a bill that was then "hurried through all its stages with unexampled rapidity"[8] and received royal assent on the same day, April 7th, 1720, that it passed the House of Lords. The stock of the Company that had traded for £120 in early January was now worth more than £300. However this was just the beginning and share prices approached £1,000 that summer.[9]

In *Famous First Bubbles*, Peter Garber argues that the prices attained by the South Sea Company shares in the summer of 1720 were justified

---

[8]Mackay (1932), page 51.
[9]Neal (1990).

by the belief on "[John] Law's prediction of a commercial expansion associated with the accumulation of a fund of credit."[10] Garber's monograph deals mostly with the Dutch Tulipmania, and Garber presents no original calculations on the South Sea bubble, but cites Scott (1910-12) who wrote "[The] investor who in 1720 bought stock at 300 or even 400, may have been unduly optimistic, but there was still a possibility that his confidence would be rewarded in the future." (pages 313-314). Scott is commenting on prices of shares of the South Sea Company that prevailed until May 18th, before share-prices doubled in a fortnight and continued to go up. In fact, in a passage a few pages later, Scott writes that by August 11 "unless the price of the stock in future issues had been set far above 1,000, the market quotations were unjustifiable...Further, it would have been impossible to have floated the surplus stock at 1,000, much less at an increased issue-price. This must have been apparent to anyone, who considered the position calmly."[11] This seems hardly an endorsement of the view that "[The South Sea] episode is readily understandable as a case of speculators working on the basis of the best economic analysis available and pushing prices along with their changing view of market fundamentals".[12]

The South Sea Bubble involved much more than the company that names it. Other chartered companies holding British government debt such as the Bank of England and the East India Company also experienced rapid share-price appreciation, albeit in a less dramatic form than the South Sea Company. In addition, numerous other joint-stock companies, nicknamed "bubble companies", were founded. Mackay (1932) list of 86 bubble companies that were declared illegal by the "Bubble Act" of July 1720 is often quoted, but Mackay is writing 120 years after the fact. A similar enumeration of bubble companies is in Anderson (1787), pp. 104-112.[13] Anderson's list gives the impression that many, though certainly not all, bubble schemes were fraudulent.

The speculation mechanism that we propose in this lecture was well understood by contemporary observers of the South Sea Bubble. The pioneering French-Irish economist Richard Cantillon, who was also a successful banker and merchant, wrote to Lady Mary Herbert, on April 29, 1720 when shares of the South Sea Company reached £400 "People are madder than ever to run into the [South Sea Company] stock and don't so much as pretend

---

[10]Garber (1980), pages 121-122.
[11]Scott (1910-12) page 323.
[12]Garber (1980), page 122.
[13]Anderson was a clerk at the South Sea Company during the bubble (Harris (1994), page 615)

to go in to remain in the stock but sell out again to profit".[14] Similarly, in his monumental history of British commerce, Anderson (1787) commented on the initial buyers of bubble-companies stocks: "Yet many of those very subscribers were far from believing those projects feasible: it was enough for their purpose that there would very soon be a premium on the receipts for those subscriptions; when they generally got rid of them in the crowded alley to others more credulous than themselves".[15]

By offering to replace illiquid British national debt by liquid shares, The Lord Treasurer Robert Harley and the other founders of the South Sea Company were pioneers of a "business model" that created value by allowing investors to exercise the option to resell to a future optimist.

3 Three stylized facts

In this Section, I present some evidence on three stylized facts that inspire my modeling choices: (i) asset price bubbles coincide with increases in trading volume (ii) asset price bubble implosions seem to coincide with increases in an asset's supply and (iii) asset price bubbles often coincide with financial or technological innovation. The evidence presented here is not meant to replace systematic empirical analysis - some of which we will discuss later - but simply to motivate the modeling that follows. To bring these stylized facts into focus, I will make references to aspects of four remarkable historical episodes of financial bubbles: The South Sea bubble, the extraordinary rise of stock prices during the roaring twenties, the internet bubble, and the recent credit bubble. I have already provided a short description of the South Sea Bubble and will assume that readers are familiar with a basic outline of the latter three episodes.

3.1 Bubbles and trading volume

Carlos et al. (2006) document that trading on Bank of England stock rose from 2,000 transactions per year in 1717-1719 to 6,846 transactions in the bubble year of 1720. They also estimate that 150% of the outstanding stocks of the East India Company and of the Royal African Company turned over in 1720.

Accounts of the stock-market boom 1928-1929 also emphasize overtrading. In fact, the annual *turnover* (value of shares traded as a percentage of

---

[14]Cited in Murphy (1986), Chapter 9.
[15]Anderson (1787), pages 102-103.

the value of outstanding shares) at NYSE climbed from 100% per annum in 1925-27 to over 140% in 1928 and 1929.[16] Daily share trading volume reached new all time records 10 times in 1928 and 3 times in 1929. No similar trading-volume record was set for nearly forty years, until April 1, 1968, when President Johnson announced he would not seek re-election.[17]

At the peak of the dotcom bubble, internet stocks had 3 times the turnover of similar non-dotcom stocks.[18] Lamont and Thaler (2003) study six cases of spinoffs during that bubble - episodes when publicly traded companies did an equity carve out, by selling a fraction of a subsidiary to the market via an IPO, and announced a plan to spin off the remaining shares of the subsidiary to the parent-company shareholders. A well-known example was Palm and 3Com. Palm, which made hand-held personal organizers, was owned by 3Com, which produced network systems and services. On March 2, 2000, 3Com sold 5% of its stake in Palm via an Initial Public Offering (IPO). 3Com also announced that it would deliver the remaining shares of Palm to 3Com shareholders before the end of that year. Lamont and Thaler (2003) document that prior to the spin off, shares in these six carve-outs, including Palm, sold for substantially more than the value of the shares embedded in the original-company's shares. Since shares of the parent company would necessarily sell for a non-negative price after the spin-off, the observed relationship between the price of carve-outs and original-companies shares indicates a violation of the law of one price, one of the fundamental postulates of textbook finance theory. In addition, the trading volume of the shares in the carve-outs was astonishing - the *daily* turnover in the 6 cases studied by Lamont and Thaler (2003) average 38%,[19] a signal that buyers of the carve-outs, just as the buyers of bubble-companies stocks in 1720, were looking for others more credulous than themselves.

It is frequently argued that excessive trading *causes* asset-prices to exceed fundamental valuations. We will not be making this argument here - in our model excessive trading and prices that exceed fundamentals have a common cause, but the often observed correlation between asset-price bubbles and high trading volume is one of the most intriguing pieces of empirical evidence concerning bubbles and must be accounted in any theoretical attempt to understand these speculative episodes.

---

[16]Davis et al. (2005)

[17]Hong and Stein (2007).

[18]For instance, in February 2000, internet firms represented 6% of the public equity market but 19% of the trading volume (Ofek and Richardson (2003))

[19]This must be compared with an *annual* turnover of 100% for the typical NYSE stock at that time.

3.2 Asset price bubbles implosion and increases in asset supply

The South Sea Bubble lasted less than a year, but in that short period there was a huge increase in the supply of joint-stock companies shares. New issues doubled the amount of shares outstanding of the South Sea Company and more than tripled those of the Royal African Company. Numerous other joint-stock companies were started during that year. The directors of the South Sea Company seem to have understood that the increase in the supply of shares of joint-stock companies threatened their own capacity to sell stock at inflated prices. Harris (1994) thoroughly examined the wording of the Bubble Act of 1720, in which Parliament banned joint-stock companies not authorized by Royal Charter or the extension of corporate charters into new ventures, and the historical evidence on interests and discourses, and concluded that "the [Bubble Act] was a special-interest legislation for the [South-Sea Company], which controlled its framing and its passage". In any case, the South-Sea Company directors used the Bubble Act to sue old chartered companies that had moved into "financial" activities and were competing with the South Sea Company for speculators' capital.

As the dotcom bubble inflated there were numerous IPOs, but in each of these only a fraction of the shares were effectively sold. The remaining shares were assigned to insiders, venture capital funds, institutions, and sophisticated investors, who had agreements to hold their shares for a "lockup" period, often 6 months. An extraordinary number of lock up expirations for dotcom companies occurred during the first half of 2000, vastly increasing the supply of shares.[20] Venture capital firms that had distributed \$3.9 billion to limited partners in the third quarter of 1999, distributed \$21 billion during the first quarter of 2000, either by giving the newly unlocked shares to the limited partners or by selling these shares and distributing cash.[21] The bursting of the bubble in early 2000 coincided with this dramatic increase in the *float* (total number of shares available to the public) of firms in the internet sector.

The recent credit bubble was characterized by an inordinate demand for liquid "safe assets" usually displaying a AAA rating from one or more of the major credit ratings agencies. Financial engineering and rosy assumptions concerning housing-prices growth and correlations of defaults allowed issuers to transform a large fraction of subprime mortgages[22] into AAA credit.

---

[20]Ofek and Richardson (2003)

[21]Janeway (2012)

[22]Including so called Alt-A mortgage loans.

Subprime mortgage loans were pooled to serve as collateral to a Mortgage Backed Security (MBS), a collection of securities (tranches) that may have different priorities on the cash flows generated by the collateral. The senior tranche typically received a AAA rating. Lower rated tranches of MBSs in turn could be pooled as collateral for a credit default obligations (CDO). The senior tranches of the CDO would again have a AAA rating. Lower rated tranches of CDOs could then be combined to serve as collateral for the tranches of a CDO[2] ...

The high prices commanded by the instruments resulting from this securitization process increased the demand by issuers for residential mortgage loans and lowered the cost of taking a mortgage thus facilitating housing purchases. In 2000, issues of private label mortgage backed securities (PLS), that is mortgage backed securities that were not issued by Government Sponsored Enterprises, financed $572 billion in US residential mortgages. By the end of 2006, the volume of outstanding mortgages financed by PLSs had reached $2.6 trillion. Many of these PLSs used less-than-prime mortgage loans and the combined annual subprime and Alt-A origination grew from an estimated $171 billion in 2002 to $877 billion in 2005, an annualized growth rate of 72%.[23]

Several developments added dramatically to the effective supply of securities backed by housing related assets. In the summer of 2005, the International Swaps and Derivatives Association (ISDA) created a standardized credit default swap (CDS), or insurance against default, for mortgage-backed securities. These contracts allowed a pessimist to buy insurance on a subprime MBS he did not own. Early in 2006, Markit launched ABX.HE, subprime mortgage backed credit derivative indexes. Each ABX index was based on 20 MBSs with the same credit rating and issued within a six-month window. The level of the index reflected the price at which a CDS on this set of MBSs was trading. Investors that had optimistic views concerning the risks in subprime MBS could now acquire a short position in a AAA series of the ABX index. If the market became more positive about these securities in the future, the cost of the corresponding CDS would drop and the shorts would make a profit. In the summer of 2006, ISDA went further and created a standard CDS contract on CDO tranches allowing investors who had a pessimistic view of say AAA tranches of subprimes to effectively take short exposures to the subprime market - a market in which, for institutional reasons, it was often difficult to short individual securities. In this way, the supply of AAA tranches of CDOs was effectively increased.

---

[23]Thomas and Van Order (2010)

None of these developments however were fully adequate to satisfy the demand for AAA paper by institutions that, often for regulatory reasons, found it necessary to buy highly rated securities. Synthetic CDOs were a perfect supply response to this demand. These were CDOs that did not contain any actual MBSs but instead consisted of a portfolio of CDSs on MBSs and some high-quality liquid assets. The creation of a standard CDS for MBSs, and the consequent increase in supply of these insurance contracts, allowed Goldman Sachs, Deutsche Bank and other Wall Street powerhouses, but also smaller firms such as Tricardia to create an enormous supply of synthetic CDOs. Wall Street could now satisfy the demands of a German Landesbank for additional US AAA mortgage bonds without any new houses being built in Arizona.[24] The associated increase in the supply of assets carrying housing risk seems to have been enough to satisfy not only optimistic German Landesbanks but also every Lehman trader or Citi SIV portfolio manager that wanted to hold housing risk. In this way, the implosion of the credit bubble parallels the implosion of the South Sea and dotcom bubbles.[25]

3.3 Asset price bubbles and the arrival of "new technologies"

Asset price bubbles tend to appear in periods of excitement about innovations. The stock market bubble of the 1920s was driven primarily by the new technology stocks of the time, namely the automobile, aircraft, motion picture, and radio industries, and the dotcom bubble has an obvious connection to internet technology.

In the US there has been notable attention to the recent housing bubble. However the housing bubble was simply one manifestation of an enormous credit bubble that took place in the early part of this century. In April 2006, while the Case-Schiller housing index reached its peak, you could buy a 5 year CDS on Greek debt for less than 15 bp (.15%) per year.[26] Similarly, in

---

[24]Synthetic CDOs have been blamed for the inordinate damage created by the subprime implosion (See *e.g.* Nocera (2010) for a non-technical indictment of synthetics), because it allowed optimistic financial institutions to take even more subprime risks. However it is not obvious what would have happened if synthetics had not existed. First, the price of "safe" subprime based securities would have been higher, causing bigger losses per security albeit on a smaller number of securities. Second, and more scary, is that we would have ended up with an even larger number of unfinished houses in the Southwest.

[25]Glaeser et al. (2006) argues that real estate bubbles are also deflated by increases in housing supply.

[26]The lowest spreads for a 5 year CDS on Greek debt, in the single digits, were reached later, in January 2007.

April 2006, the average spread for a CDS on debt from Argentina, a country that had defaulted repeatedly and as recently as 2002, was less than 3% per year.

This credit bubble coincided with advances in *financial engineering* - the introduction of new financial instruments and hedging techniques and advances in risk measurement that promised better risk management and "justified" lower risk-premia.

4   Evidence for costly short selling and overconfidence

Economists typically treat short sales of an asset as the purchase of a negative amount of that asset, and assume that short-sales generate just as much transaction costs as purchases. Although there are exceptions - such as future markets - legal and institutional constraints make this assumption problematic in almost all cases. To short an asset requires finding a lender for that asset and, because often there are no organized market for borrowing an asset, finding a lender can be difficult. In addition, securities are often loaned *on call* and borrowers face the risk of replacing the borrowed securities or being forced to cover their short position.[27] Securities loans are often collateralized with cash. The security lender pays interest on the collateral, but the lender pays the borrower of the security a *rebate rate* that is less than the market rate for cash funds. Rebate rates may be negative and thus the fee effectively paid by the borrower of the security can exceed market interest rates. Among other factors, the rebate rate reflects the supply and demand for a particular securities' loan and the likelihood that the lender recalls the security. D'avolio (2002) documents that rebate rates are negatively while recalls are positively correlated with measures of divergence of opinions. The possibility of recall makes shorting securities with a small float and/or little liquidity especially risky. Individual MBS securities or certain tranches of CDOs had relatively small face values.

Diether et al. (2002) provide evidence that stocks with higher dispersion in analysts earnings forecasts earn lower future returns than otherwise similar stocks. It is reasonable to take the dispersion in analysts forecasts as a proxy for differences in opinion about a stock, and the observation of lower returns for stocks with more difference in opinions is consistent with the hypothesis that prices will reflect a relatively optimistic view whenever

---

[27]In fact, U.S. regulations often require many institutional lenders to maintain the right to terminate a stock loan at any time. (D'avolio (2002))

going long is cheaper than going short. In contrast, the evidence reported by Diether et al. (2002) is inconsistent with a view that dispersion in analysts forecasts proxies for risk, since in this case stocks with higher dispersion should not exhibit lower returns.

There are of course many possible ways in which differences in beliefs may arise. In this lecture I will assume that differences in beliefs are related to *overconfidence* - the tendency of individuals to exaggerate the precision of their knowledge. The original paper documenting overconfidence is Alpert and Raiffa (1982). Since then overconfidence has been documented in a variety of groups of decision makers, including engineers (Kidd (1970)) and entrepreneurs (Cooper et al. (1988)). Tetlock (2005) discusses overconfidence in a group of professional experts who earn a living commenting or advising on political and economic trends, such as journalists, foreign policy specialists, economists and intelligence analysts. The vast majority of these pundits' predictions seem to be no better than random chance.

Even more directly relevant to the topic of this lecture is the paper by Ben-David et al. (2010). Between June 2001 and September 2010, Duke University collected quarterly surveys of senior finance executives, the majority of whom were CFOs and financial vice-presidents. Among other questions, the respondents were asked to give two numbers: The first was a number that they believed there was only a one-in-ten chance that the actual S&P annual return over the next year would be below that number. The second was a number that they believed there was only a one-in-ten chance that the actual S&P annual return over the next year would be above that number. These two numbers form the 10-90 interval, that is the interval of numbers for which a respondent believes there is a 10% chance that the actual S&P returns would fall to the left of that interval and a 10% chance that the actual returns would fall to the right of that interval. The 10-90 interval should cover 80% of the realizations. In total, the surveys collected over 12,500 of these intervals and the realized returns in the S&P over each year following a survey fell within the executives 10-90 intervals only 33% of the time. Evidently, these senior finance executives grossly overestimated the precision of their knowledge concerning future stock returns.

5   Sketch of a model

The appendix contains a model connecting difference of opinions and costly shorting to speculation and trading. The model in the appendix is a simplified version of an already stylized model developed in Scheinkman and

Xiong (2003), who in turn were inspired by a pioneering paper by Harrison and Kreps (1978).

In the model in the appendix, there are two types of investors that for simplicity are assumed to be risk neutral - that is, they are willing to pay for an asset that they are forced to hold to maturity that asset's (discounted) expected payoff. Differences of opinions arise because investors estimate future payoffs of a risky asset using signals they believe are useful to predict payoffs. Some investors are "rational" and use signals in an optimal fashion. Others attribute value to information they should ignore - perhaps a cable-TV host named JC recommending a "buy" or a "sell." In the model "irrational" investors are on average right, but depending on the particular value of the useless information that they observe, they can be excessively optimistic or excessively pessimistic.[28] Thus, on average, opinions of investors are unbiased. I also assume for simplicity that short sales are not allowed, although it would suffice to assume costly short-selling.

Suppose that an asset will have a payoff two years from now which may be high or low with equal probability. Suppose further that one year from now, JC may voice an opinion on which of the two payoffs is likely to occur. The TV host's opinion is totally unfounded, but there is a large group of investors that believe that JC's views are valuable. Since there are no short sales allowed, if each group of agents has more than enough capital to acquire the whole float of the asset at their own valuation, then once JC's opinion is known, members of the most optimistic group would acquire the whole supply and, because they compete with others of the same group, buyers would end up paying their expected payoff. If JC claims the higher payoff is likely to obtain, the irrational agents would pay a price that reflects an optimistic view of the asset payoff. If JC claims the lower payoff is likely to occur, then the irrational agents would be pessimists, but rational agents would still be willing to buy the asset paying a price equal to the rational-agents expected payoff. And if JC is silent, both agents agree that the asset is worth the rational-agents' expected payoff. Now suppose a market where the asset is traded opens today. A rational investor knows that if a year from now JC screams "high dividend", she would have the option to sell the asset at that moment to an irrational investor at a price higher than her own valuation would be at that point. Otherwise, if JC stays silent or utters a pessimistic opinion, the investor would be happy to hold the asset. Thus a rational buyer would be willing to pay today in excess of her own valuation of future payoffs, because she acquires an option to resell the asset

---

[28]That is, if JC emits an opinion, it is equally likely to be a "buy" or a "sell".

one year from now if JC screams "high dividend". The more likely it is that in one year from now JC would claim that a high dividend will obtain, the larger would be the amount that a rational investor would pay for the asset today. Because of the symmetry we assumed between the probability that JC claims that a high payoff will occur and the probability that JC claims that a low payoff will occur, the rational investor would pay more for the resale option when there is a higher probability that JC would emit any opinion. Similarly, an irrational investor would pay more than his own valuation for the asset today, because he knows that if JC claims next year that a low payoff will occur, he would be able to sell the asset to someone that he would judge to be overoptimistic.

In the context of the model, I define a bubble as the value that a buyer pays for this option to resell. Thus a bubble occurs when a buyer pays in excess of her valuation of future dividends, because she values the opportunity to resell to a more optimistic buyer in the future. Since buyers would tend to be among the most optimistic agents, it would be natural to call the difference between buyer's valuation and a "rational" valuation also a bubble. Here, I do not include buyers' excessive optimism as part of the bubble, and thus the definition of a bubble used in this lecture is somewhat conservative. Although bubbles certainly coincide with periods in which excessive optimism prevails among many investors, our definition emphasizes the role of the *existence* of divergent opinions as opposed to the actual differences in opinions that occur during these episodes.

If the ownership of the asset is equally distributed initially between rational and irrational agents, trades will occur whenever JC emits an opinion. On average we would get a higher volume of trade whenever there is a larger probability that JC would give an opinion. Thus the same cause - the frequency of JC opinions - creates differences in opinion, a bubble and trading. In the appendix we show that this difference in opinions can be identified with overconfidence.

The value of the resale option is naturally a function of the costs of funds. The higher the interest rate faced by investors, the less they are willing to pay for the resale option. The model in the appendix thus gives a simple theoretical justification for the argument that lower interest rates are conducive to bubbles. In a similar manner, shorter horizons yield fewer opportunities to resell, making the resale option less valuable.

The model in this lecture ignores two forces that have been invoked to dismiss the importance of differences in beliefs. The first is learning - the irrational agents should eventually learn that the signal they are using is useless. Learning no doubt plays a role in diminishing differences in be-

liefs over long horizons, but bubbles last for a relatively short period when learning must have a limited effect. The second argument brought against the importance of irrational beliefs is survivorship. As argued by Friedman (1966), irrational agents should loose wealth on average and thus have a vanishing influence on market outcomes. However,Yan (2008) performed calibration exercises on Friedman's argument and concluded that for reasonable parameter values, it may take *hundreds of years* for irrational investors to loose even half their wealth.Because bubbles are relatively short-lived, I will ignore learning and survivorship and emphasize other forces that create and deflate bubbles.

5.1   Limited capital

If irrational investors have limited access to capital and the supply of the asset increases, perhaps as a result of sales by insiders, then even when JC emits an optimistic opinion, irrational investors may not be able to buy the full asset float while paying their own valuation. When the capital constraint of irrational investors is severe enough, even when irrational optimism occurs, the marginal buyer may be a rational investor who has a lower valuation of the asset. Hence when irrational agents have limited capital, the size of the bubble depends on the asset supply. For the same reason, if the asset's float is large enough, some of the asset supply may end up in the hands of rational investors even though irrational investors are optimists and have a higher valuation for the asset. As a consequence, the turnover (volume traded as a fraction of the float) of an asset is smaller for assets with a larger float.

In the model developed in the appendix, in the presence of capital constraints, an increase in supply that is not fully expected leads to a deflation of the bubble. This was one of the main insights in Hong et al. (2006). In the internet bubble, increases in supply were often the result of sales by insiders. Hong et al. (2006) observed that it is reasonable to assume that unexpected sales by insiders lead to a revision of forecasts by current investors and potential buyers and that this revision in beliefs must reinforce the tendency of supply increases to produce bubble's implosion.

5.2   Leverage

The model in the appendix does not explicitly treat leverage, but the observations on limited capital also provide insights on the role of leverage. Investors can often access capital using their purchases of assets as collat-

eral for loans. The amount loaned to finance the purchase of one unit of a risky asset would typically be less than the price of the asset. The difference between the price of the asset and the value of the loan is the *margin* and its reciprocal the *leverage*. A homeowner that acquired a house in 2004 with a 5% downpayment would have thus a leverage of 20. Higher leverage would increase the access to capital by optimists and thus help to augment and sustain bubbles.

In the presence of belief disagreements pessimists should be willing to make loans collateralized by the risky-asset to optimists. Market conditions determine the leverage and interest rates charged on these loans, but one should expect that pessimists would demand relatively low leverage and/or high interest rates. Pioneered by Geanokoplos,[29] a literature which studies the equilibrium determination of leverage and interest rates for loans from pessimists to optimists has developed.

In reality however, because of tax or regulatory reasons, not all optimists are adequate holders of certain risky assets. For instance, homeowners benefit from the absence of taxation on imputed rent and the unique treatment of capital gains in owner-occupied homes. Although they were not the most appropriate direct investors in houses, optimistic banks had another way to benefit from the housing price increases that they anticipated in 2004. They could make loans charging more than prime rates to subprime buyers that the banks believed would be capable of repaying their loans in the very likely event that house prices continued to behave as they had in the previous ten years. Contemporary analysts' reports from major financial institutions recognized the potentially negative impact of house-prices decline on the value of mortgage related securities, but underestimated the probability of occurrence of these adverse events and thus as argued by Foote et al. (2012), it is reasonable to conclude that some investors in mortgage related securities were simply excessively optimistic about the possibility of house-price declines.[30] Money market funds, which by their nature must invest in short-term "safe" securities, participate heavily in repo markets - essentially loans collateralized by securities. A money market fund that was willing to finance 98.4% of the purchase-price of a AAA mortgage security to an investor in 2006 probably thought that these securities were actually nearly risk-free, warranting a leverage of 60. In this way, a chain of opti-

---

[29]See Geanakoplos (2010) for a recent summary.

[30]"the average coupon on subprime adjustable-rate mortgages was several hundred basis points above the comparable prime loans. And yet, if investors think that house prices can rise 11 percent per year, expected losses are minimal." (Foote et al. (2012) pp 32,33)

mists provided leverage to optimistic investors in the housing market.[31] This chain was reinforced by US regulation that placed low capital-risk weights on securities deemed AAA by "Nationally Recognized Statistical Rating Organizations" and by similar regulations in other countries, and was amplified by innovations in finance, such as the MBS based CDO. It is ironic that this same process of innovation in financial engineering eventually allowed pessimists to express their negative views on these markets and speeded up the implosion of the bubble.

Compared to pessimists, optimists without direct access to a risky asset are bound to accept terms more favorable to the borrower on loans backed by that asset. Thus it is reasonable to argue that during the credit bubble, leverage from optimists was a more important source of capital for mortgage-securities investors than leverage from pessimists.

5.3   Origins of optimism

The formal model exposited in the appendix is silent concerning the precipitating factors that generate optimism among investors. In practice, investors rely on advice from friends, acquaintances and "experts." Some advice is without doubt biased because of the financial incentives faced by experts, but it has been documented that during speculative episodes, apparently unbiased advisors also issue overoptimistic forecasts.[32]

Motivated by the coincidence of bubbles and periods of excitement about new technologies Hong et al. (2008) proposed a theory based on the role of formal or informal advisors. In the model in Hong et al. (2008), there are two types of advisors. Tech-savvy advisors understand the new technologies - think of a finance quant during the credit bubble - while old-fogies are uniformly pessimistic concerning the new technologies. Tech-savvies may be well intentioned advisors, but worry about being confused with old-foggies. As the art critic in Tom Wolfe's *The Painted Word*, tech-savvies worry that "To be against what is new is not to be modern. Not to be modern is to write yourself out of the scene. Not to be in the scene is to be nowhere".[33]

To insure that their advisees do not confuse them with old-foggies, tech-savvies issue over-optimistic forecasts concerning assets related to the new

---

[31]Optimistic investors also obtained leverage from financial market participants that understood the risks involved, but benefitted from skewed incentives.

[32]During the dotcom period, so-called objective research firms with no investment banking business, such as Sanford and Bernstein, issued recommendations every bit as optimistic as investment banks (*e.g.* Cowen et al. (2006)).

[33]Wolfe (1975), page 84.

technologies. Rational investors understand the advisors motivations and "debias" the advice, but naive investors take advisors recommendations at face value. Although the presence of old-foggies tends to depress prices of the assets related to the new technology, when there is a sufficient number of naive investors guided by tech-savvies, the biased advice overcomes the effect of old-foggies and induces over-optimism among investors.

5.4 Executive compensation, risk-taking and speculation

Although our discussion until now has been mainly concerned with the behavior of individual investors, corporations have played a central role in recent bubbles. The implosion of the credit bubble and consequent Wall Street bailout brought deserved attention to the risk-taking behavior of financial firms during that episode and lead to calls for compensation reforms that would eliminate excessive incentives for managers to take risks.

The standard economists' approach to compensation uses the "Principal-Agent" framework which emphasizes how managerial contracts are set by boards as shareholders' representatives to solve the misalignment of interests between managers and stockholders. Bebchuk and Fried (2006) and other critics of this approach contend that CEOs have been able to essentially set their own contracts through captured boards and remuneration committees and that major reforms in corporate governance to increase shareholder power are necessary to remediate the current state of affairs.

The critics of the standard approach to compensation are no doubt correct in pointing out important ways in which the selection of board members and executive pay negotiations depart from the idealized "arms-length" bargaining of the Principal-Agent paradigm. However the critics have more difficulties explaining how the relatively recent phenomenon of rise in pay and stock-based compensation has coincided with the dramatic rise in shareholder influence that began in the 1980s.[34] In fact, we have observed a tendency towards greater board independence, a higher proportion of externally recruited CEOs, a decrease in the average tenure of CEOs, and higher forced CEO turnover during this period.

Bolton et al. (2006) point out that a speculative market creates a divergence between the interests of short-term versus long-term stockholders and between the interests of current versus future stockholders. Short-term stockholders would like managers to take actions that increase the speculative value of shares, even if at the cost of the fundamental value of the firm.

---

[34]See Holmström (2006) and Holmström and Kaplan (2001).

If stockholders with a short-term horizon dominate a board, they would select contracts for managers that emphasize stock-price based compensation that vests early, to align the interests of managers with their own interests.[35]

Examining a panel of US financial firms during the period of 1992-2008, Cheng et al. (2010) found substantial cross-firm differences in total executive compensation even after controlling for firm size. Top management level of pay is positively correlated with price-based risk-taking measures including firm beta, return volatility, the sensitivity of firm stock price to the ABX subprime index, and tail cumulative return performance. Managers' compensation and firm risk-taking are not related to governance variables but covary with ownership by institutional investors who tend to have short-termist preferences and the power to influence a firm's management policy.[36]

The empirical results in Cheng et al. (2010) indicate that governance reforms are hardly the solution for excessive risk-taking by financial firms.

6   Some additional evidence

Two data sets, both coincidently from China, provide additional evidence to support the mechanisms I proposed in this lecture. These data sets have been used in research that was motivated by the bubble models discussed in this lecture.

Between 1993 and 2000, 73 Chinese firms offered two classes of shares, A and B, with identical rights. Until 2001, domestic Chinese investors could only buy A shares while foreign investors could only hold B shares. Mei et al. (2009) used these data to test implications of models of heterogenous beliefs and short-sale constraints. This is particularly appropriate, because at that time Chinese buyers of A shares faced very stringent short-sale constraints, and IPOs and SEO's (Seasoned Equity Offerings) were tightly controlled by the central government.

Despite their identical rights to dividends and voting rights, A shares traded on average at a premium of 420% relative to B shares. The annual turnover of B shares, around 100%, was similar to the turnover of NYSE

---

[35]Or as Lewis (2004) wrote: "The investor cares about short-term gains in stock prices a lot more than he does about the long-term viability of a company. .... The investor, of course, likes to think of himself as a force for honesty and transparency, but he has proved, in recent years, that he prefers a lucrative lie to an expensive truth. And he's very good at letting corporate management know it".

[36]Froot et al. (1991) point out that the horizon of many institutional investors is around 1 year.

shares at the time, while A shares traded at 500% a year. The relatively large panel of 73 stocks allows Mei et al. (2009) to control for cross-sectional differences in risk and liquidity and time variations in China's risk and risk-premium. They find that A-share turnover is significantly positively correlated with the A-B share premium, and in fact 20% of that premium can be "explained" by turnover variation. On the other hand, B-share turnover had a positive association with the A-B premium, albeit not statistically significant. This supports the hypothesis that A-share prices (but not B-share prices) were driven by speculation. Mei et al. (2009) also show that the A-B share premium and A-share turnover increase with a firm's idiosyncratic return volatility, a proxy for uncertainty. This is consistent with the bubble theory based on heterogeneous beliefs if one believes that more fundamental uncertainty would increase fluctuations in heterogenous beliefs.

Furthermore, Mei et al. (2009) show that controlling for the proportion of days in which a share did not display a price change - a proxy for liquidity that has been used in the finance literature - does not significantly change the association between A-share turnover and the A-B premium. To determine whether trading in A and B shares was driven by speculation or liquidity, they examined the cross-sectional correlation between share turnover and asset float of A and B shares. Liquidity typically increases with asset float, since as float increases, it is easier for buyers to match up with sellers. On the other hand, as we argued above, in the presence of speculation and limited capital, a larger float is associated with a smaller turnover. Mei et al. (2009) finds a significant negative relationship between share turnover and float in A shares markets in 1993-2000, suggesting that the large trading volume in A shares was not a result of liquidity. However for B shares, which were held by more sophisticated foreign investors, Mei et al. (2009) found that turnover was positively associated with float - suggesting that liquidity played a role in attracting trading in B shares.

On February 28, 2001, the Chinese government allowed domestic investors to buy B shares provided they used foreign currency. The A-B premium decreased but almost exclusively because B share prices went up. Monthly B-share turnover in the six months following this event averaged 44%, almost 4 times the monthly turnover of these shares in the 6 months preceding the liberalization. Moreover, the coefficient of the A-B premium on B-share turnover becomes significantly negative after the liberalization. This contrasts with the results for the earlier period (positive and insignificant) and indicates that speculation became a relevant component of B-share price formation. In addition, after Chinese investors were allowed to buy B shares using foreign currency, the coefficient of a regression of turnover of B

shares on B-shares float turned from positive to negative, suggesting again that trading in B shares may have become more driven by speculation.

The Chinese warrant bubble of 2005-2008 was used by Xiong and Yu (2011) to test predictions of heterogeneous beliefs *cum* short-sale constraints theories of bubbles. In 2005-2008, 18 Chinese companies issued put warrants on their stock with maturities ranging from six months to two years. These warrants gave the holder the right to sell the issuing companies' stocks at predetermined prices.

The extraordinary rise of prices in Chinese stocks between 2005 and 2007 made it almost certain that these warrants would expire without being exercised. In fact, using the familiar Black-Scholes option-pricing formula Xiong and Yu (2011) calculated that close to their expiration date, these warrants often were worth less than .05 hundredth of a yuan. However prices of these virtually worthless warrants varied substantially and averaged .948 yuan during the days in which their Black-Scholes values fell below .05 hundredth of a yuan. Xiong and Yu (2011) also found that these warrants sometimes traded for more than the strike price, what could only be justified if investors were counting on a *negative* stock price in the future. Xiong and Yu (2011) provide other bounds on the value of warrants that are violated in this sample of warrants.

One security Xiong and Yu (2011) describe in detail is the put warrant on the stock of WuLiangYe Corporation, a liquor producer.[37] The put warrant was issued on April 3, 2006 in-the-money with an exercise price of 7.95 yuan while WuLiangYe's stock traded at 7.11 yuan. Initially, the warrant was valued close to 1 yuan, but in two weeks, WuLiangYe's stock price exceeded the strike and the warrant never returned in-the-money. On October 15th 2007, the stock reached a peak of 71.56 yuan and then drifted down to close at 26 yuan at the expiration of the warrant on April 2, 2008. The calculation by Xiong and Yu (2011) is that from July 07, the Black-Scholes price of this put was below .05 hundredth of a yuan, but the warrant traded for a few yuans, and only dropped below its initial price of .99 yuan in the last few trading days. Xiong and Yu (2011) also notes that on at least two trading days, June 13 and 14, 2007, the put warrant traded for more than the strike price.

As in other episodes discussed in this lecture, these unjustifiably high prices were accompanied by trading frenzies. The warrants with a Black-Scholes value of less than .05 hundredth of a yuan had an average daily turnover rate of 328 percent. On their last trading day, when they were

---

[37] *c.f.* Figure 1 in Xiong and Yu (2011)

all virtually worthless, these 18 warrants, on average, turned over 100% of their float every 20 minutes! The trading volume on the warrant on the stock of WuLiangYe Corporation reached 1,841% of that warrant's float in the last trading day. Xiong and Yu (2011) show that as predicted by the models discussed here, the size of the price bubble on a warrant was positively correlated with the trading volume of that warrant or the time to expiration, and negatively correlated with the warrant's float.

7 Some final observations

One question left unanswered in this lecture is whether one could use the signals associated with bubbles such as inordinate trading volume or high leverage, to detect and perhaps stop bubbles. One of the difficulties in using these signals is that we know next to nothing about false positives. For instance, the typical empirical paper studying the association between volume of trade and bubbles examines data during a bubble episode.

Even if we could effectively detect bubbles, it is not obvious that we should try to stop all bubbles. Although credit bubbles have proven to have devastating consequences, the relationship between bubbles and technological innovation suggests that some of these episodes may play a positive role in economic growth. The increase in the price of assets during a bubble makes it easier to finance investments related to the new technologies.

The most straightforward policy recommendations that arise rom the arguments advanced in this lecture is that to avoid bubbles, policy makers should consider limiting leverage and facilitating, instead of impeding, short-selling. In the panic that followed the implosion of the credit bubble, the SEC banned short-sales of financial stocks. In August 2011, as the markets questioned the health and funding needs of European financial institutions, France, Italy, Spain and Belgium imposed bans on short-sales of financial stocks. Each of these interventions may have have given a temporary respite to the markets for these assets, but caused losses to investors that were short these assets and had to cover their positions. Investors learned one more time that it is dangerous to bet against overvalued assets - a lesson that they will surely keep in mind in the next bubble.

A Formal model

A.1 A Simple model

I first exposit a simple model to illustrate the role of costly shorting and differences in beliefs in generating bubbles and the association between bubbles and trading. Consider 4 periods $t = 1, 2, 3, 4$, a single good, and a single risky asset in finite supply S. In addition to the risky-asset there exits also a risk-free technology. An investment of δ units of the good in the risk-free technology at t yields one unit in period $t+1$. Assume there is a large number of risk-neutral investors, that only value consumption in the final period $t = 4$. Each investor is endowed with an amount W_0 of the good.

The risky asset produces dividends at times $t = 2, 3, 4$. At each $t = 2, 3, 4$ each unit of the risky asset pays a dividend $\theta_t \in \{\theta_\ell, \theta_h\}$ with $\theta_h > \theta_\ell$. In what follows, I will refer to θ_h (θ_ℓ) as the high (*resp.* low) dividend. Dividends at any t are independent of past and future dividends.[38] The probability that $\theta_t = \theta_\ell$ is .5, and we write

$$\bar{\theta} = \mathrm{E}(\theta_t) = .5\theta_\ell + .5\theta_h.$$

Assets are traded at $t = 1, 2, 3, 4$. If a dividend is paid in period t, trading occurs after the dividend is distributed - that is the asset trades *ex-dividend* and the buyer of the asset in period t has the rights to all dividends from time $t + 1$ on. Thus in the final period $t = 4$ the price of the asset $p_4 = 0$, since there are no dividends paid after period 4. The price at time $t = 1, 2, 3$ depends on the expectations of investors regarding the dividends to be paid in the future. We first calculate the willingness to pay of a "rational" risk-neutral investor that is not allowed to resell the asset after she buys it. Since the investor is risk-neutral, at time $t = 3$ she is willing to pay $\delta\bar{\theta}$ for a unit of the asset. In the absence of resale opportunities, at $t = 2$, the rational investor is willing to pay $\delta[\bar{\theta} + \delta\bar{\theta}] = (\delta + \delta^2)\bar{\theta}$. Finally at $t = 1$ that same rational investor with no resale opportunities would be willing to pay

$$(\delta + \delta^2 + \delta^3)\bar{\theta} \tag{1}$$

In addition, we suppose that at each $t = 1, 2, 3$ a signal s_t is observed *after* the dividend at t (if $t > 1$) is observed but *before* trading occurs at t. Each signal s_t assumes one of three values $\{0, 1, 2\}$, is independent of past realizations of the signal and of the dividends and has no predictive power

---

[38]It would be easy to accommodate a non-stationary dividend

for future dividends. Thus the signal s_t is pure noise. There are however two sets of investors, A and B. Each set has many investors. Agents in group A are rational and understand that the distribution of future dividends is independent of s_t. Agents in B actually believe that s_t predicts θ_{t+1} and that the probability that $\theta_{t+1} = \theta_h$ given s_t is:

$$\text{Probability}[\theta_{t+1} = \theta_h | s_t] = .5 + .25(s_t - 1).$$

Thus agents in group B believe that the probability of a high dividend at $t+1$ increases with the observed s_t and that when $s_t = 0$ ($s_t = 2$), $\theta_{t+1} = \theta_\ell$ (resp. $\theta_{t+1} = \theta_h$) is more probable. All agents agree that s_t does not help predict θ_{t+j} for $j \geq 2$, and thus the only disagreement among investors is whether s_t can predict θ_{t+1}. To make agents in set B correct on average, and thus assure that *ex-ante* there are no optimistic or pessimistic investors, assume that the probability that $s_t = 0$ equals the probability that $s_t = 2$. Write $q < .5$ for this common probability, and observe that the probability that $s_t = 1$ is $1 - 2q > 0$.

In the case of binary random variables, forecasts have minimal precision[39] when the probability of each realization is $1/2$. This is exactly the forecast of rational agents here. Agents in group B after observing $s_t = 1$ also have the same minimal forecast precision. However if they observe $s_t = 0$ or $s_t = 2$ they employ forecasts that have higher precision, since they (mistakenly) believe that one of the two possible events has a probability of $3/4$. In this sense agents in group B have an exaggerated view of the precision of their beliefs.

At time $t = 3$, rational agents in group A are willing to bid up to $\delta\bar{\theta}$ for a unit of the asset. However, if $s_3 = 2$, agents in set B believe that the probability of a high dividend in period 4 is .75. Since these agents are risk-neutral and the risk-free technology transforms δ units at 2 into 1 unit at time 3, when $s_3 = 2$ members of group B are willing to pay up to

$$\delta(.75\theta_h + .25\theta_\ell) = \delta[\bar{\theta} + .25(\theta_h - \theta_\ell)] > \delta\bar{\theta}.$$

I assume that there are no short sales. Section A.2 below treats the case when there is limited capital in the hands of a group of investors, but I will initially assume that W_0 is large enough so that each group of investors has sufficient aggregate wealth to acquire the full supply of the asset at their own valuation. Suppose $s_3 = 2$. Since there are many agents in group B, no short sales, and agents in group B have sufficient wealth to acquire the

---

[39]That is, maximal variance.

full supply, all risky assets end up in the hand of *some* agents in group B and competition among agents of group B guarantees that when $s_3 = 2$ the price

$$p_3(2) = \delta[\bar{\theta} + .25(\theta_h - \theta_\ell)].$$

If $s_3 = 1$, then all agents value the asset at $\delta\bar{\theta}$, and thus

$$p_3(1) = \delta\bar{\theta}.$$

If $s_3 = 0$, agents in set B are unduly pessimistic and the asset ends up in the hands of some rational agents that pay

$$p_3(0) = \delta\bar{\theta}.$$

Since the probability that $s_3 = 2$ is q, the price p_3 equals $\delta\bar{\theta}$ with probability $1 - q$ and equals $\delta[\bar{\theta} + .25(\theta_h - \theta_\ell)] > \delta\bar{\theta}$ with probability q. Hence, before s_3 is observed, all agents anticipate that the price of the asset in period 3 will on average equal

$$\mathrm{E}p_3 = \delta[\bar{\theta} + .25q(\theta_h - \theta_\ell)]. \tag{2}$$

Every agent in group A *and* in group B expects the payoff of the asset in period 4 to be on average exactly $\bar{\theta}$, but the average price in period 3 exceeds the discounted value of this expected payoff, $\delta\bar{\theta}$, by

$$\delta[.25q(\theta_h - \theta_\ell)],$$

reflecting the fact that for each realization of the signal s_3 a member of the most optimistic group would acquire the asset. When $s_3 = 0$ ($s_3 = 2$) agents in group A (*resp. B*) are the most optimistic and end up holding the total supply of the asset. When $s_3 = 1$ both groups are equally optimistic and any distribution of asset holdings across the two groups is compatible with equilibrium.

Although the price paid in period t sometimes reflects the excessively optimistic views of group B agents, our definition of a bubble - that is a bubble occurs when buyers pay more than *they* think the future dividends are worth - implies that there is no bubble in period 3. As I will show next, this is not the case for any $t < 3$.

At time $t = 2$ if $s_2 = 2$, agents in set B assume that the probability of a high dividend θ_h in period 3 is .75. Since s_3 has not yet been observed, agents in set B forecast the price in period 3 to be on average $\mathrm{E}p_3$. Thus, when $s_2 = 2$ agents in group B are willing to pay,

$$p_2(2) = \delta(\bar{\theta} + .25(\theta_h - \theta_l) + \mathrm{E}p_3). \tag{3}$$

Similarly, if $s_2 = 1$ ($s_2 = 0$) agents of both types (*resp.* agents of type A) are willing to pay

$$p_2(0) = p_2(1) = \delta(\bar{\theta} + \mathrm{E}p_3).$$

Hence, before s_2 is observed, agents anticipate an average price for the asset:

$$\mathrm{E}p_2 = \delta(\bar{\theta} + .25q(\theta_h - \theta_\ell) + \mathrm{E}p_3) = (\delta + \delta^2)[\bar{\theta} + .25q(\theta_h - \theta_\ell)]. \quad (4)$$

A buyer of the asset in period 2 acquires the right to dividends in period 3 and 4. Before s_2 is observed, agents of both types agree that these dividends are worth (in period 2) exactly $(\delta + \delta^2)\bar{\theta}$. However the average price in period 2 exceeds this fundamental value by

$$(\delta + \delta^2) \times .25q(\theta_h - \theta_\ell).$$

This difference is a consequence of (i) the fact that in period 2 for each realization of the signal s_2, the asset will be sold to the highest bidder and (ii) any buyer of the asset in period 2 acquires the right to *resell* it in period 3 to a buyer that is more optimistic than she is. (i) is worth $\delta \times .25q(\theta_h - \theta_\ell)$, while the option to resell in period 3 (ii) is worth $\delta^2 \times .25q(\theta_h - \theta_\ell)$. Furthermore even in the event $s_2 = 0$ when "rational" (group A) agents acquire the full supply of the risky asset, these agents pay

$$\delta(\bar{\theta} + \mathrm{E}p_3) = \delta\bar{\theta} + \delta^2[\bar{\theta} + .25q(\theta_h - \theta_\ell)] \quad (5)$$

This value exceeds the fundamental value, since buyers of the asset in period 2 will benefit from the resale option in period 3. When $s_2 = 1$, holders of the asset (in group A or B) are also willing to pay this same amount. Whereas in period 3, prices exceed fundamentals only if group B agents are optimistic, in period 2, prices exceed fundamentals even when group B agents are unduly pessimistic.

More importantly, for every realization of the signal s_2, the buyer of the asset is willing to pay in excess of her estimate of the value of future dividends, an amount that represents the option to resell in period 3 and that equals

$$b_2 := .25\delta^2 q(\theta_h - \theta_\ell). \quad (6)$$

This difference can be naturally called a (period 2) bubble and is a result of fluctuating differences of opinion and future opportunities to trade.

In order to model a bubble resulting from differences in beliefs and restrictions to short selling, it suffices to consider three periods. However, to examine bubble implosions as in Section A.2, we need two periods in which

an asset pricing bubble can potentially occur. For this reason, I consider four periods, but as I hope it becomes transparent, the reasoning in period $t = 1$ duplicates exactly the argument we used for period 2 above.

In fact, buyers at $t = 1$ are willing to bid for the asset an amount that reflects the sum of the dividends they expect the asset to pay and the value of the option of reselling the asset in future periods. For instance, the valuation of future dividends at $t = 1$ for a rational buyer is given by expression (1). However the rational buyer is willing to pay at $t = 1$

$$\delta(\bar{\theta} + \mathrm{E}p_2) \tag{7}$$

The difference between (7) and (1) is

$$b_1 := .25(\delta^2 + \delta^3)q(\theta_h - \theta_\ell). \tag{8}$$

It is easy to check that (8) also expresses the difference between the reservation value of a B agent and her valuation of future dividends for every value of the signal s_1. Thus for every realization of the signal s_1, b_1 represents the amount that the buyer of the asset is willing to pay in excess of her estimate of the value of future dividends.

Since there is no bubble in period 3, we may set $b_3 = 0$ and a comparison of (6) and (8) establishes that:

$$b_1 > b_2 > b_3.$$

Bubbles decline over time because there are fewer opportunities to resell.

In this very simple model, we may think of the parameter q as a measure of differences in beliefs. After all, in periods $t = 1, 2, 3$, with probability $2q$ there are differences in beliefs once the signals s_t are observed. In addition, as argued above, agents in group B exaggerate the precision of their beliefs whenever $s_t \neq 1$. The event $s_t \neq 1$ has probability $2q$. Thus an increase in q also corresponds to an increase in the probability that agents in group B exhibit overconfidence. The size of the bubble in periods 1 or 2 is increasing as a function of q. Further, if we write

$$r := \frac{1}{\delta} - 1$$

for the (risk-free) interest rate implicit in the risk-free technology, then in periods 1 and 2 the bubble decreases with the risk-free interest rate.

The model discussed here has predictions on the effect of the difference in beliefs q on the volume of trade. For symmetry, suppose both groups are

28

of the same size and at time 0 each group owns half of the supply of the risky asset. In period 1, with probability $2q$, all assets are bought by agents in one of the two groups, and with probability $1 - 2q$, $s_1 = 1$ and all agents agree on the distribution of dividends in the future, and thus there is no reason to trade. Hence, since S is the total supply of the asset, the average volume of trade in period 1 is

$$\text{EV}_1 = \frac{1}{2} \times 2q \times S = qS. \tag{9}$$

If $s_1 = 2$ and $s_2 = 0$ or if $s_1 = 0$ and $s_2 = 2$ the group holding the risky-asset in period 1 would sell it in period 2. The probability that $s_1 = 2$ and $s_2 = 0$ equals q^2, which is also the probability that $s_1 = 0$ and $s_2 = 2$. Also, trade will occur if $s_1 = 1$, and $s_2 = 0$ or $s_2 = 2$, but in this case, only half the assets would change hands. The probability of this event is $(1 - 2q)2q$. Finally, in all other cases no trade would occur in period 2. Thus the expected volume of trade in period 2 is,

$$\text{EV}_2 = 2q^2 S + (1 - 2q)2q \times \frac{S}{2} = qS. \tag{10}$$

All shares will change hand in period 3, if $s_3 = 2$ and $s_2 = 0$ or if $s_3 = 0$ and $s_2 = 2$. In addition, all shares will change hand at $t = 3$ if the history of signals is $(0, 1, 2)$ or $(2, 1, 0)$. These four events have an aggregate probability $2q^2 + 2q^2(1 - 2q)$. In addition, trade will occur if the histories of the signals is $(1, 1, 2)$ or $(1, 1, 0)$, and each of these events have probability $q(1 - 2q)^2$, but if histories $(1, 1, 2)$ or $(1, 1, 0)$ obtain, only half the shares will exchange hands at $t = 3$. Thus the average volume in period 3 is given by

$$\text{EV}_3 = 2q^2 S + 2q^2(1 - 2q)S + q(1 - 2q)^2 S = qS. \tag{11}$$

An examination of expressions (9) - (11) makes it evident that an increase in the parameter q increases the expected volume of trade in every period, just as it increases the value of the bubble. An increase in q raises the probability that the option to resell will be exercised (*volume*) and the value of that resale option (*size of bubble*).

I summarize these results in a Proposition:

Proposition 1. *In the presence of fluctuating differences in beliefs and short-sale constraints, bubbles exist - investors are willing to pay for an asset in excess of their own valuation of future dividends. In addition,*

   (i) *The size of the bubble increases when the probability of disagreement increases.*

(ii) The volume of trade also increases with the probability of disagreement.

(iii) The size of the bubble decreases with the risk-free interest rate.

(iv) The bubble declines as the time of maturity of the asset approaches, because there are fewer opportunities to trade.

A.2 Limited capital

It is straightforward to incorporate more periods into the model or treat a stationary model with infinite horizon. The introduction of risk-aversion complicates substantially the computations. However as Hong et al. (2006) showed, in a model of heterogeneous beliefs and costly short sale with risk-averse agents, the bubble and asset turnover rate decrease as the supply of the asset increases and bubbles may implode when float unexpectedly increases. The intuition is that as their holdings of the asset increases, risk averse agents have a smaller marginal valuation for the asset and, as a consequence, it takes a bigger difference in opinions for the whole asset supply to change hands. An alternative that displays the same relationships between float, bubbles and turnover as when agents are risk averse, albeit in a less continuous manner, is to adopt the short-cut proposed by Allen and Gale (2002) of limited capital and *cash-in-the-market-pricing*. Each group of agents will have the same preferences and views on signals as in appendix A.1, but here I no longer assume that the wealth of agents in both groups is so large that the they can always acquire the full supply of the risky asset at their own valuation. The demand for the risky asset by an agent will be limited by the liquidity in her hands.

Before trading in period t, but after receiving period t dividends, agents in group $C \in \{A, B\}$ will have an aggregate portfolio (S_{t-1}^C, K_t^C). Here S_{t-1}^C is the amount of the risky-asset they chose to hold in period $t-1$ and $K_t^C > 0$ is the cash available to them. K_t^C is a result of investments in the risk-free technology in period $t-1$, period t dividends on risky-asset holdings and any (net) risk-free borrowing the agents in group C may access. Below, I will deduce each group's portfolio holdings before trading in period t, as a function of their initial holdings (S_0^C, K_0^C), portfolio decisions, and the realizations of signals and dividends, but at this point I treat these portfolios as exogenous.

Write v_t^C for the (marginal) valuation that agents in group $C \in \{A, B\}$ have for the risky asset at time t. More specifically v_t^C equals the maximum price per-unit that an agent in group C that has a positive income is willing to pay for an infinitesimal amount of the asset. This amount is a function

of the expectations that the agent has concerning dividends and prices in period $t+1$. The limitation imposed by liquidity constraints on demand may yield lower equilibrium prices than those derived in appendix A.1 and thus cause the marginal valuation of some agents to be smaller than the marginal valuations deduced in appendix A.1.

Write S_t for the total supply of the asset at t. To simplify matters, I will assume that there is enough capital in the hands of group A agents so that at time t they can acquire the whole supply not yet in the hands of group A agents, $S_t - S_{t-1}^A$, at their valuation v_t^A. That is, I assume that:

$$\frac{K_t^A}{S_t - S_{t-1}^A} > v_t^A \tag{12}$$

Notice that inequality (12) involves the portfolios held by group A agents and the valuation of A agents, v_t^A. However one may replace v_t^A in the r.h.s. of (12) by the (larger) valuation of the risky-asset by A agents that was already calculated in appendix A.1, and the denominator $S_t - S_{t-1}^A$ by S_t. This stronger assumption would imply inequality (12).

I will assume however that agents in group B have more limited capital. Suppose in period t group B agents have a valuation v_t^B for the risky asset, while the valuation of A agents is v_t^A. If $v_t^A \geq v_t^B$ then the price of the asset $p_t = v_t^A$ since inequality (12) insures that group A agents have sufficient capital to acquire the total supply S_t at their valuation v_t^A. If $\frac{K_t^B}{(S_t - S_{t-1}^B)} \geq v_t^B > v_t^A$ then B agents would bid up the price, until $p_t = v_t^B$, since in this case group B agents have sufficient capital to acquire the risky assets not in their hands at their own valuation. If $v_t^A \leq \frac{K_t^B}{(S_t - S_{t-1}^B)} < v_t^B$, then the marginal buyer would belong to group B, but because of the liquidity constraints on group B agents, the price is then given by

$$p_t = \frac{K_t^B}{(S_t - S_{t-1}^B)}.$$

In the language of Allen and Gale (2002), cash-in-the-market-pricing obtains; the price of the asset is determined by the liquidity available to group B agents. Finally, if $v_t^A > \frac{K_t^B}{(S_t - S_{t-1}^B)}$ then the marginal buyer is necessarily in group A and $p_t = v_t^A$. Thus:

$$p_t = \max\left\{ v_t^A, \min\left\{ v_t^B, \frac{K_t^B}{(S_t - S_{t-1}^B)} \right\} \right\}$$

Other things equal, this price decreases as S_t increases. For instance, the larger is the initial float S_1 relative to the initial holding of the B group, the lower would be the initial price p_1. If the initial float is large enough, then even when B agents are optimists some of the asset supply ends up in the hands of A agents, because of the limited capital of group B agents. In fact, when B agents are optimists, but the marginal buyer is an A agent, the amount that ends in the hands of B agents, S_1^B, solves:

$$v_1^A = \frac{K_1^B}{(S_1^B - S_0^B)}.$$

In this way, a larger float lowers the price and the turnover of the asset. If the supply of the risky asset remains constant through time, an asset with a larger float will also have a smaller turnover in subsequent periods.

To study the effect of an increase in the supply of the risky asset in period 2, we will first consider the price that would prevail in period 3. When $s_3 = 2$ (group B agents are optimists) $v_3^B > v_3^A$. If $K_3^B \geq (S_3 - S_2^B)v_3^B$, then group B agents will bid up the price of the asset to their valuation v_3^B and acquire the full supply. However, if $s_3 = 2$ but $K_3^B < (S_3 - S_2^B)v_3^B$, group B agents do not have enough wealth to buy the full supply of the risky-asset at their valuation v_3^B. In this case, if they buy the full supply, the maximum group B agents can pay for an unit of the asset is $\frac{K_3^B}{S_3 - S_2^B}$. If the price group B agents can afford when they acquire the full supply in the hands of group A exceeds the valuation of group A agents, that is, $\frac{K_3^B}{S_3 - S_2^B} > v_3^A = \delta\bar{\theta}$, then the equilibrium price when $s_3 = 2$ is $\frac{K_3^B}{S_3 - S_2^B}$. Finally if $\frac{K_3^B}{S_3 - S_2^B} < \delta\bar{\theta} = v_3^A$, some of the asset is held by group A agents even though $s_3 = 2$, and thus the price of the risky asset must equal $\delta\bar{\theta}$. The price p_3 when $s_3 = 2$ is a (weakly) decreasing function of S_3, and equals $\delta\bar{\theta}$ if the supply of the asset in period 3 is large enough. Since $p_3 = \delta\bar{\theta}$ whenever $s_3 \neq 2$ the average price in period 3, $\mathrm{E}p_3$, is also a (weakly) decreasing function of S_3 and equals $\delta\bar{\theta}$, if the supply of the asset in period 3 is large enough, when compared to the holdings of B agents in period 2. This contrasts with the case without limited capital when $\mathrm{E}p_3$ would always exceed $\delta\bar{\theta}$ (c.f. equation (2) above.)

Buyer's of the stock of the South Sea Company or buyers of Internet stocks in the late 90's did not know with certainty the future supply of these assets. One can model this uncertainty in a simple way by assuming that supply may increase in period 2. Suppose $S_0 = S_1 = S$, and that $S_3 = S_2 = S$ with probability π and equals $S + \Delta S > S$ with probability $1 - \pi$. Suppose further that the realization of the supply of the risky asset is

independent of the realization of $(\theta_1, \theta_2, s_1, s_2)$, and that this realization is observed in period 2 *before* the signal s_2 is observed. Since trading in period 2 occurs after s_2 is observed, investors know the actual supply when they trade in period 2. While the amount S is held from the beginning by agents in groups A and B any increase in supply will come from sales of the asset by "insiders", as in Hong et al. (2006). To simplify matters, I will assume that insiders only wish to sell.[40]

I will show that for certain parameter values, when the supply does not increase ($S_2 = S$) the bubble will persist but when the supply increases the bubble will deflate. The deflation of the bubble in period 2 will be a consequence of the realization by all investors that, because of the supply increase and the limited aggregate wealth of group B agents, the marginal buyer in period 3 will necessarily be rational even when $s_3 = 2$. The occurrence of this scenario will depend on the wealth of group B agents in period 3 and the supply they would have to absorb in that same period to guarantee that the price exceeds the valuation of the rational buyers. I will start by imposing some bounds on the aggregate portfolio of group B agents that will insure that the bubble persists when the supply of the asset is unchanged and the bubble deflates when the supply increases, and will later show that these bounds would hold when the initial supply of the risky asset, S, is sufficiently small and the potential increase in the supply in period 2, ΔS, is sufficiently large.

Suppose that the capital constraints of group B only bind in period 3 when the supply increases:

$$\frac{K_3^B}{S - S_2^B} \geq \delta[\bar{\theta} + .25(\theta_h - \theta_\ell)], \tag{13}$$

and that when the capital constraints bind, they are strong enough that is:

$$\frac{K_3^B}{S + \Delta S - S_2^B} < \delta\bar{\theta}. \tag{14}$$

Equation (13) guarantees that if $S_2 = S$ and $s_3 = 2$, the demand by group B agents will drive the price of the asset in period 2 to v_2^B and thus exceed the expected "rational" payoff in period 3. Inequality (14) insures that if S_2 increases to $S + \Delta S$ the price of the asset in period 3 will equal the (discounted) average payoff expected by rational agents, since a member of group A would necessarily be the marginal buyer of the asset.[41]

---

[40]However see footnote 41 below, for an assumption that implies that insiders only sell.
[41]Equivalently we could have assumed instead that the reservation price of insiders

33

Thus when $S_2 = S$ the price in period 3 matches exactly the price when there are no capital constraints whereas when $S_2 = S + \Delta S$ the price in period 3 equals the valuation of the rational agents, independent of the signal. In particular, if the supply of the asset increases in period 2 because of sales by insiders, that is $S_2 = S + \Delta S$, there is no period 2 bubble, since it is known that period 3 prices are independent of the signal s_3 that is observed. On the other hand, if supply of the asset does not increase in period 2, then, as in the case of no capital constraints,

$$\mathrm{E}p_3 = \delta[\bar{\theta} + .25q(\theta_h - \theta_\ell)],$$

reflecting the fact that B agents would acquire the whole float if $s_3 = 2$.

In order to guarantee that if $S_2 = S$ prices in period 2 would be exactly as in the case of no capital constraints we need to insure that

$$\frac{K_2^B}{S - S_1^B} \geq (\delta + \delta^2)[\bar{\theta} + .25q(\theta_h - \theta_\ell)]. \tag{15}$$

Inequality (15) guarantees that when $s_2 = 2$ and $S_2 = S$, B agents will acquire the full float of the asset. To insure that when $S_2 = S + \Delta S$ the marginal buyer of the asset always belongs to group A, it suffices that

$$\frac{K_2^B}{S + \Delta S - S_1^B} \leq (\delta + \delta^2)\bar{\theta}. \tag{16}$$

If inequalities (13) and (15) hold then before S_2 (and hence s_2) is observed

$$\mathrm{E}p_2 = (\delta + \delta^2)[\bar{\theta} + .25q\pi(\theta_h - \theta_\ell)]$$

reflecting the fact that a bubble would occur in period 2 if and only if the supply stays constant. In addition when inequalities (13), (14) and (15) hold, rational buyers in period 1 will always be willing to pay in excess of their own valuations of future dividends of the risky asset, because if supply does not increase they may have an opportunity to sell the asset to over-optimistic B agents in the future. Since we have assumed that group A agents have sufficient capital to buy the total supply of the asset at these higher prices, the price of the asset in period 1 exceeds the expected

is greater that v_2^B, with probability π, and less than $\delta\bar{\theta}$, with probability $1 - \pi$. For in this case, given that equation (13) and inequality (14) hold, insiders sell no shares with probability π and sell all their shares with probability $1 - \pi$.

discounted dividends independently of the realized signal in period 1 (s_1) and the capital constraints of group B agents.

When $s_1 = 2$, group B valuation at time 1 is

$$v_1^B(2) = \delta \left(\bar{\theta} + .25q(\theta_h - \theta_\ell) + (\delta + \delta^2)[\bar{\theta} + .25q\pi(\theta_h - \theta_\ell)]\right),$$

which again exceeds the present value of dividends expected by group B agents by $(\delta + \delta^2)[.25q\pi(\theta_h - \theta_\ell)]$. To guarantee that the price in period 1 when $s_1 = 2$ is observed equals $v_1^B(2)$ one must assume that:

$$\frac{K_1^B}{S - S_0^B} \geq v_1^B(2) \tag{17}$$

In this case, there will be a bubble in period 1, independently of the signal s_1 observed and $b_1 = .25(\delta + \delta^2)q\pi(\theta_h - \theta_\ell)$, which is smaller than the bubble that obtains when liquidity constraints are not binding - equation (8) - reflecting the possibility that future supplies may increase. However as $\pi \to 1$ the value of the bubbles under the two scenarios converge to each other.

Until now I took the portfolio held by group B agents after the payment of dividends in periods 1 to 3 as given, and assumed that inequalities (13)-(17) hold. Agents in group B start with a non-negative initial endowment of $S_0^B < S$ units of the risky asset and K_0^B units of the good and the values $S_{t-1}^B \leq S$ and K_t^B for $t = 2, 3$ are consequences of their actions, realizations of the random shocks and equilibrium prices in period 1 and 2. Furthermore, for a given K_0^B it is straightforward to show that inequalities (13), (15) and (17) hold, whenever S is small enough. The intuition underlying this result is that if S is small enough, even if B agents acquire the full supply of the risky asset and have no possibility of borrowing, they would have a minimum amount left over to invest in the risk-free technology. This delivers a lower bound on K_1^B, K_2^B and K_3^B, the amounts available to agents in group B to acquire additional shares in periods 1 to 3 . By assuming an even smaller value for S, if necessary, we can thus guarantee that inequalities (13), (15) and (17) hold.

I now show that inequalities (14) and (16) hold provided ΔS is large enough and agents in group B have no access to borrowing in addition to the amount that is already reflected in their initial liquidity K_0^B. To accomplish this, we have to study the dynamics of the evolution of the aggregate wealth of B agents. Given K_0^B and S_0^B, the evolution of the aggregate wealth of group B in equilibrium depends on the realizations of the dividends, signals and supply and on the way assets are allocated between the two groups when

their valuations are identical. I assume that when the two groups have identical valuations for the risky asset, group B agents get all the shares they want.[42] Write $W_t^B(K_0^B, S_0^B)$ for the maximum wealth that agents in group B can have after dividend payments in period t, where the maximum is taken over all possible realizations of signals, dividends and all portfolio choices. Notice that that an increase in ΔS cannot increase W_1^B or W_2^B because the price of the risky asset can only decrease with an increase in ΔS. The price of the asset in period 2 is not less than $(\delta + \delta^2)\bar{\theta}$, the expected (by A agents) discounted dividends of the asset. If ΔS is large enough,

$$W_2^B < \left(\delta + \delta^2\right)\bar{\theta}\frac{S + \Delta S}{2} \leq p_2\frac{S + \Delta S}{2}.$$

Even if B agents use all their wealth in period 2 to buy the asset they cannot acquire more then half the total (larger) supply, and thus inequality (16) holds and,

$$S_2^B < \frac{S + \Delta S}{2}. \tag{18}$$

Hence the marginal buyer in period 2 when supply increases always belongs to group A. Further, the *ex-post* rate of return of the risky asset between periods 2 and 3 will depend on the price of the asset that prevails in periods 2 and 3 and the realization of the dividend θ_3. Since the price of the risky asset in period 3 is at most the expected (by B agents) discounted dividends of the asset when $s_3 = 2$, and the dividend paid is at most θ_h, the rate of return is at most

$$\bar{R} := \frac{\theta_h + \delta\left(\bar{\theta} + .25(\theta_h - \theta_\ell)\right)}{(\delta + \delta^2)\bar{\theta}}$$

Notice that this bound exceeds $1 + r$ and thus is also a bound for the growth in wealth. Hence,

$$W_3^B \leq \bar{R}W_2^B$$

and by choosing if necessary a larger ΔS we can insure that

$$K_3^B \leq W_3^B \leq \bar{R}W_2^B \leq \delta\bar{\theta}\frac{\Delta S}{2} < \delta\bar{\theta}\left(S + \Delta S - S_2^B\right),$$

where the last inequality follows from equation (18). Hence inequality (14) holds.

Thus in the model in this section, a bubble arises in period 1 provided that "irrational" agents have enough initial wealth relative to the initial

---

[42]A different rule, such as giving priority to agents in group A, would change the details of the computations that follow, but would not alter the result of interest

supply of the risky asset and the bubble implodes in period 2 if and only if the supply of the risky asset increases by a sufficiently large amount in period 2.

References

Allen, F. and D. Gale. 2002. Optimal financial crises. *The Journal of Finance* 53 (4):1245–1284.

Alpert, M. and H. Raiffa. 1982. A progress report on the training of probability assessors. *Judgment under uncertainty: Heuristics and biases* 294–305.

Anderson, A. 1787. *An historical; and chronological deduction of the origin of commerce*, vol. 3. London, UK.

Arrow, K.J. 1986. Rationality of self and others in an economic system. *Journal of Business* 385–399.

Bebchuk, L.A. and J.M. Fried. 2006. *Pay without performance: The unfulfilled promise of executive compensation.* Harvard University Press.

Ben-David, I., J.R. Graham, and C.R. Harvey. 2010. Managerial miscalibration. Tech. rep., National Bureau of Economic Research.

Bolton, P., J. Scheinkman, and W. Xiong. 2006. Executive compensation and short-termist behaviour in speculative markets. *The Review of Economic Studies* 73 (3):577–610.

Brunnermeier, M. K. 2008. bubbles. In *The New Palgrave Dictionary of Economics*, edited by S. Durlauf and L. Blume. Basingstoke: Palgrave Macmillan.

Carlos, A., L. Neal, and K. Wandschneider. 2006. Dissecting the anatomy of exchange alley: The dealings of stockjobbers during and after the South Sea bubble. *Unpublished paper, University of Illinois* .

Cassidy, J. 2010. Rational irrationality: An interview with Eugene Fama. *The New Yorker* November 1.

Cheng, I.H., H. Hong, and J. Scheinkman. 2010. Yesterday's Heroes: Compensation and Creative Risk-Taking. *ECGI-Finance Working Paper* (285).

Cooper, A.C., C.Y. Woo, and W.C. Dunkelberg. 1988. Entrepreneurs' perceived chances for success. *Journal of business venturing* 3 (2):97–108.

Cowen, A., B. Groysberg, and P. Healy. 2006. Which types of analyst firms are more optimistic? *Journal of Accounting and Economics* 41 (1):119–146.

Davis, L.E., L.D. Neal, and E.N. White. 2005. The Highest price ever: The great NYSE seat sale of 1928-1929 and capacity constraints. Tech. rep., National Bureau of Economic Research Cambridge, Mass., USA.

D'avolio, G. 2002. The market for borrowing stock. *Journal of Financial Economics* 66 (2):271–306.

Diether, K.B., C.J. Malloy, and A. Scherbina. 2002. Differences of opinion and the cross section of stock returns. *The Journal of Finance* 57 (5):2113–2141.

Foote, C.L., K.S. Gerardi, and P.S. Willen. 2012. Why did so many people make so many ex post bad decisions? the causes of the foreclosure crisis. Tech. rep., National Bureau of Economic Research.

Friedman, M. 1966. *Essays in positive economics*, vol. 231. University of Chicago Press.

Froot, K.A., A.F. Perold, and J.C. Stein. 1991. Shareholder trading practices and corporate investment horizons. Tech. rep., National Bureau of Economic Research.

Garber, P. 1980. *Famous first bubbles: The fundamentals of early mania.* Cambridge, Ma: MIT Press, 1st ed.

Geanakoplos, J. 2010. The leverage cycle. *NBER Macroeconomics Annual* 24 (1):1–66.

Glaeser, Edward L, Joseph Gyourko, and Raven E Saks. 2006. Urban growth and housing supply. *Journal of Economic Geography* 6 (1):71–89.

Harris, Ron. 1994. The Bubble Act: Its passage and its effects on business organization. *The Journal of Economic History* 54 (3):pp. 610–627.

Harrison, J.M. and D.M. Kreps. 1978. Speculative investor behavior in a stock market with heterogeneous expectations. *The Quarterly Journal of Economics* 92 (2):323–336.

Holmström, B. 2006. Pay without performance and the managerial power hypothesis: A comment. *Available at SSRN 899096* .

Holmström, B. and S.N. Kaplan. 2001. Corporate Governance and Merger Activity in the US: Making Sense of the 1980s and 1990s. Tech. rep., National Bureau of Economic Research.

Hong, H. and J.C. Stein. 2007. Disagreement and the stock market. *The Journal of Economic Perspectives* 21 (2):109–128.

Hong, H., J. Scheinkman, and W. Xiong. 2006. Asset float and speculative bubbles. *The Journal of Finance* 61 (3):1073–1117.

———. 2008. Advisors and asset prices: A model of the origins of bubbles. *Journal of Financial Economics* 89 (2):268–287.

Janeway, W. 2012. *Doing capitalism in the innovation economy.* Cambridge, UK: Cambridge University Press, 1st ed.

Kidd, J.B. 1970. The utilization of subjective probabilities in production planning. *Acta Psychologica* 34:338–347.

Lamont, O.A. and R.H. Thaler. 2003. Can the market add and subtract? mispricing in stock markets carve-outs. *Journal of Political Economy* 111 (2):227–268.

Lewis, M. 2004. The Irresponsible Investor. *New York Times Magazine* June 6.

Mackay, C. 1932. *Extraordinary popular delusions and the madness of crowds.* New York: Farrar, Straus and Giroux.

Mei, J., J.A. Scheinkman, and W. Xiong. 2009. Speculative trading and stock prices: Evidence from Chinese AB share premia. *Annals of Economics and Finance* 10 (2):225–255.

Murphy, A.E. 1986. *Richard Cantillon: entrepreneur and economist.* Clarendon Press Oxford.

Neal, L. 1990. *The rise of financial capitalism: international capital markets in the age of reason.* Cambridge, UK: Cambridge University Press.

Nocera, J. 2010. A Wall street invention let the crisis mutate. *New York Times* April 17:B1.

Ofek, E. and M. Richardson. 2003. Dotcom mania: The rise and fall of internet stock prices. *The Journal of Finance* 58 (3):1113–1138.

Santos, M.S. and M. Woodford. 1997. Rational asset pricing bubbles. *Econometrica* 19–57.

Scheinkman, J.A. and W. Xiong. 2003. Overconfidence and Speculative Bubbles. *Journal of Political Economy* 111 (6).

Scott, W. R. 1910-12. *The constitution and finance of English, Scottish and Irish joint-stock companies by 1720*, vol. 3. Cambridge, UK: Cambridge University Press.

Shiller, R.J. 2006. *Irrational exuberance*. Crown Business.

Tetlock, P.E. 2005. *Expert political judgment: How good is it? How can we know?* Princeton University Press.

Thomas, J. and R. Van Order. 2010. Housing policy, subprime markets and Fannie Mae and Freddie Mac: What we know, what we think we know and what we dont know. In *St. Louis FRB Conference*.

Wolfe, Tom. 1975. *The painted word*. New York: Farrar, Straus and Giroux.

Xiong, W. and J. Yu. 2011. The Chinese warrants bubble. *American Economic Review* 101:2723–2753.

Yan, Hongjun. 2008. Natural Selection in Financial Markets: Does It Work? *Management Science* 54 (11):1935–1950.