Investor Sentiment and Pre-Issue Markets[*]

Francesca Cornelli
(London Business School and CEPR)

David Goldreich
(London Business School and CEPR)

Alexander Ljungqvist
(New York University and CEPR)

This Version: May, 2004

Abstract. What role do sentiment investors play in the pricing of newly listed stocks? We derive conditions under which we can distinguish between sentiment and rational pricing behavior and test for the rationality of small investors' demand for new stock issues using data from *pre*-issue (or 'grey') markets in Europe. Under sentiment, the model predicts asymmetric relations between the prices at which small investors trade new stock issues in the grey market and i) the subsequent issue price set by the investment bank, ii) prices in the early after-market, and iii) the degree of stock price reversal in the long run. Our empirical results suggest that sentiment demand is present and influences the pricing of newly listed firms.

1. Introduction

Behavioral biases such as 'conservatism' or 'over-confidence' have become popular explanations for a variety of asset pricing phenomena that are hard to reconcile with a rational decision-making framework.[1] However, the extent to which 'sentiment' (as opposed to rational) investor demand can account for these phenomena in equilibrium is controversial, not least given the difficulty in identifying empirically the demand curves and information processing of different investor groups. Our aim in this paper is to study the role (if any) of sentiment investors in the pricing of newly listed stocks, using a unique dataset that helps identify the valuations and informational updating among small investors.

The market for newly listed firms is a good setting in which to study the effect of sentiment investors on stock prices. Such firms by definition have no prior share price history and tend to be young, immature, and relatively informationally opaque. Not surprisingly, therefore, they are hard to value, and it seems reasonable to assume that investors will have a wide range of priors about their market values. Miller (1977) points out that this heterogeneity (coupled with short sale constraints) can lead to predictable patterns, such as poor returns in the long run. Similarly, Ritter and Welch (2002) argue that overenthusiasm among retail investors may explain the much documented price jumps once trading in newly listed stocks begins, as well as the subsequent low returns over the first few years of trading.

A major difficulty in testing such behavioral conjectures is disentangling the expectations of different types of investors when we observe only a single, market-clearing price. In this paper, we take advantage of a quirk in the trading environment of many European countries to isolate the *pre*-issue valuations of a group of predominantly small (largely retail) investors and link these to *post*-issue prices both in the short run and the long run. We are able to do this because many European countries have pre-issue (or 'grey') markets in which investors can speculate on the future stock prices of companies that are in the process of going public. Grey markets operate in parallel with institutional bookbuilding but are organized by unaffiliated brokers.[2] While grey market trading is widespread in Europe, regulations in the U.S. prohibit investors from trading IPOs before they are listed on an exchange. There is, however, a very active when-issued market for Treasury securities.

---

[1]See Shleifer (2000) for a survey of investor sentiment and its theoretical underpinnings. See Daniel, Hirshleifer, and Subrahmanyam (1998) and Barberis, Shleifer, and Vishny (1998) for models of investor sentiment.

[2]Section 4 describes grey markets in more detail.

1

Grey market prices are a matter of public record. They provide us a unique opportunity to observe the opinions of an important subset of mostly small (typically retail) investors. This allows us to examine the effect such investors have on i) the issue price the investment bank sets at the end of bookbuilding, ii) the prices at which newly listed companies trade in the aftermarket, and iii) the degree of stock price reversal in the long run.

Rather than simply describing these relations empirically, we impose some economic structure on the data. Our aim in doing so is to derive conditions under which we can distinguish between sentiment and rational trading behavior and thus test for the rationality of small investors' demand for new stock issues. In the model, small investors trade in the grey market, while the underwriter collects bids from large institutional investors during bookbuilding before pricing the issue.[3] Thus, there are two separate markets with two separate sets of investors. Bookbuilding investors are endowed with information about the fundamental value of the shares. Grey market investors also have information which may or may not be relevant to the fundamental value. In addition, grey market investors may overweight the relevance of their information (which is analogous to 'overconfidence' as in Daniel, Hirshleifer, and Subrahmanyam (1998) or 'conservatism' as in Barberis, Shleifer, and Vishny (1998)). Note that we allow for the possibility that grey market investors are sentiment investors, but we do not impose it.

To the extent that grey market investors are representative of small investors in general, their valuation is indicative of small investors' reservation price in the aftermarket. If that reservation price exceeds the fundamental value of the shares, bookbuilding investors can sell their shares in the immediate aftermarket at a price above the fundamental value. If instead small investors' reservation price is low, bookbuilding investors will hold on to the shares that they are allocated, valuing them at the fundamental value. As a result, bookbuilding investors value the shares at the maximum of the fundamental value and small investors' reservation price.

In this setting, the empirical relations that we expect between the grey market price and both the issue price and the aftermarket price depend on the degree to which the grey market price contains information about the fundamental value, and also on the possibility that grey market investors tend to overweight the relevance of their information.

---

[3]For a description of bookbuilding see Cornelli and Goldreich (2001) and Ljungqvist and Wilhelm (2002).

Consider the extreme case in which the grey market information is completely unrelated to the fundamental value, but grey market investors mistakenly use it in computing their reservation price. In this case, we expect a positive relation between the grey market price and the aftermarket price only when the reservation price is above the fundamental value. When the reservation price is below the fundamental value there should be no relation, because bookbuilding investors will not sell their shares to grey market investors.

The underwriter also observes the grey market price. When the grey market price is high, he will anticipate that the shares will be sold in the aftermarket and will set a higher offer price to extract the bookbuilding investors' trading surplus. When the grey market price is low, the underwriter will disregard the valuation of small investors when pricing the shares. Thus, to the extent that the underwriter anticipates a high reservation price, there will be a similar asymmetric relation between the grey market price and the issue price.

Next consider the case where the grey market price does contain fundamental information. Now there will be a positive relation between the grey market price and both the issue price and the aftermarket price even when the grey market price is low. However, to the extent that grey market investors overweight their information, there will still be an asymmetry. When the grey market price is high the aftermarket price will be very close to the grey market price, but when the grey market price is low the aftermarket price will only incorporate the component of the grey market price that is related to fundamentals.[4]

Note that the model predicts an asymmetric relation between these prices only if grey market investors are sentiment investors. Without sentiment their reservation price will always equal the expected fundamental value. In such case, the relation will depend on the presence of fundamental information in the grey market price regardless of whether the grey market price is high or low, i.e., there will be no asymmetry. Thus, whether or not grey market investors are sentiment investors and overweight their information is an empirical question which can be answered empirically in the context of our model.[5]

---

[4]Note that prices are sometimes biased upwards, but never downwards, even though the potentially irrational investors could be either excessively optimistic or pessimistic. In the presence of short sale constraints, excessively pessimistic investors are priced out of the market.

[5]Similar predictions would arise if grey market investors are less informed than bookbuilding investors, but do not fully update when they observe the aftermarket price. This form of irrational behavior is closely related to the one in the model except that the conservatism occurs when trading begins in the aftermarket, rather than before.

We test the predictions of the model using grey market price data for a large set of European IPOs completed between 1995 and 2002. We find that the grey market price is more correlated with the issue price and the aftermarket price when the grey market price is high, although there is a positive correlation even when the grey market price is low. This suggests that the grey market price contains information about the fundamental value, but also that grey market investors overweight their information in a way that is reflected in the aftermarket price and exploited by the underwriter when setting the issue price. We also find higher levels of aftermarket trading volume when the grey market price is high, consistent with bookbuilding investors selling their shares to grey market investors only when the grey market investors have higher valuations. In other words, our empirical findings support the view that sentiment investors can drive up prices in the short-run aftermarket.

The model also has empirical implications for long-run returns. Recall that when the grey market price is high, demand from sentiment investors will cause the shares to trade at a high price relative to fundamentals in the short run. In the long run, prices will revert to the fundamental value, as the true value is revealed through time, and we expect negative returns. On the other hand, when the grey market price is low, the aftermarket price is always based on fundamentals and we do not expect a reversal pattern. Our empirical results are consistent with this prediction.

We should point out one thing this paper does not do. Although the optimal mechanism in our model involves allocating underpriced shares to facilitate information extraction in bookbuilding (as in Benveniste and Spindt (1989)), it is not our goal to explain the magnitude of observed underpricing. Our aim is instead to show how high valuations among grey market investors can lead to a high aftermarket price, not why the issue price is often set well below this. A number of explanations given in the literature (for example, agency conflicts) could be added to our model to explain underpricing.

We stress that our results are also relevant for countries that do not have a grey market (such as the United States). As long as some investors are motivated by sentiment and the underwriter and the major institutional investors have some sense of what these investors are willing to pay, sentiment investors will drive short-run prices upwards when they are overly optimistic. The significance of the grey market is that it allows us to observe small investors' valuations easily and directly, enabling us to test for these effects.

Related literature

Recent literature investigating the role of sentiment in explaining asset pricing patterns includes Neal and Wheatley (1998) and Baker and Wurgler (2003). While this literature considers sentiment as a market-wide phenomenon, the grey market enables us to proxy for investor sentiment with respect to individual stocks.

Several papers on IPOs have documented empirical patterns that also motivate our study. Ritter (1991) presents evidence that abnormally high prices immediately after the IPO are followed by abnormally low returns in the long-run, and Ritter and Welch (2002) show that this pattern is particularly strong during 'hot market' periods.

Purnanandam and Swaminathan (2003) compare IPO offer prices to 'fair values' computed using various price multiples of non-IPO industry peers. They find that issues which are overpriced relative to fair value also have higher returns on the first day of trading but lower returns in the long run. These patterns are consistent with our results. If some investors are excessively enthusiastic about an issue, the underwriter sets the offer price above the fundamental value, expecting these investors to buy the shares in the aftermarket. This leads to a high short-run price and negative long-run drift as the price converges to the fundamental value.

Krigman, Shaw, and Womack (1999) find that a high level of 'flipping' predicts low returns in the long run. In the context of our paper, flipping can be interpreted as bookbuilding investors selling their shares to grey market investors, which is when we also find low long-run returns.

A large literature, both theoretical and empirical, has attributed these IPO patterns to the presence of sentiment investors. Lee, Shleifer, and Thaler (1991) show that the annual number of IPOs is negatively related to the closed-end fund discount which they argue is a measure of retail investor sentiment. Rajan and Servaes (2003) model two different types of irrational agents, feedback traders and sentiment investors. The latter are similar to our grey market investors. They proxy for investor sentiment using the market-to-book ratio and find that it correlates positively with first-day returns and negatively with long-run returns. In contrast, we use the grey market price to measure the valuation of sentiment investors.

Ljungqvist, Nanda, and Singh (2003) argue that an initial price run-up may be due to the existence of 'exuberant' investors and may lead to long-term underperformance. Their

model has similarities with ours, but focuses on explaining underpricing, which is needed to compensate regular investors for losses in case the 'hot' market ends prematurely.

Aggarwal, Krigman, and Womack (2002) relate the aftermarket price path to momentum traders, and focus on the role of research analysts and the media in creating momentum. They find that 'extra hot' issues tend to have low long-run returns.

Testing behavioral theories requires investigating the role of retail or small investors. While we use the grey market price as an indication of small investors' valuation, other studies have instead looked at who owns the shares in the aftermarket. Ofek and Richardson (2003) show that high initial returns occur when institutions sell IPO shares to retail investors on the first day. Similarly, Ben Dor (2003) looks at the level of institutional ownership shortly after the IPO and finds that high institutional ownership forecasts higher returns in 'hot' markets. These findings are consistent with our paper, since we predict that bookbuilding investors sell their shares to small investors when they are overvalued. Derrien (2004) develops a model where retail investors bid for shares in the bookbuilding process in the expectation that they can sell them at a higher price in the aftermarket. In a sample of French IPOs, he finds that retail investors' bookbuilding demand correlates positively with the issue price and initial returns, and negatively with long-run performance. These results are consistent with our findings. In addition, through the grey market we have the unique opportunity to observe the valuation of the investors who will be buying shares in the aftermarket, rather than just the investors who will be selling shares obtained in bookbuilding. In other words, we can measure the expectations of sentiment investors directly.

In an empirical study that is complementary to our findings, Dorn (2003) finds that the volume of grey market trading among the customers of a German retail brokerage is correlated with high initial returns and low long-run returns. This can be viewed as further evidence that participation by small investors in the grey market can be interpreted as sentiment. Besides Dorn (2003), two other papers study Germany's grey market. Löffler, Panther, and Theissen (2002) document that grey market prices are unbiased estimates of first-day prices. Aussenegg, Pichler, and Stomper (2003) also find that IPO issue prices are related to grey market prices. However, the estimated coefficient in their model is less than one, which they argue is consistent with bookbuilding being used for gathering information *before* the initial price range is set (consistent with the model of Jenkinson, Morrison, and Wilhelm (2003)). Finally, Pichler and Stomper (2004) model the interaction between bookbuilding and the

6

grey market when grey market investors have similar information to bookbuilding investors. They ask whether the existence of a grey market helps or hinders information aggregation in bookbuilding. In contrast, we introduce a class of investors who have (potentially) different information from bookbuilding investors, in order to explain certain IPO phenomena and to show how these (possibly biased) investors affect prices.

The paper proceeds as follows. In Section 2 we present the model and derive the optimal mechanism. We discuss the empirical implications in Section 3. Section 4 describes the data. Section 5 presents the empirical results. In Section 6 we extend the model to allow bookbuilding investors to trade in the grey market. Section 7 concludes.

2. The model

An issuer wishes to sell S shares in an IPO. Each share has a fundamental value $v \in [0, \bar{v}]$. Since the underwriter does not know v before setting the issue price P_I, he conducts bookbuilding to collect information from institutional investors about v. Simultaneous with bookbuilding, a publicly observable grey market takes place in which a different group of investors trade the shares on a when-issued basis.[6]

The expected fundamental value of a share is a weighted average of the information arriving from bookbuilding s_B and the information from the grey market s_G:

$$E(v \mid s_B, s_G) = \alpha s_G + (1 - \alpha)s_B, \tag{1}$$

where $0 \leq \alpha < 1$. In the extreme case of $\alpha = 0$, grey market investors' information is irrelevant. We assume that bookbuilding investors' information is always relevant.[7]

The timing is as follows. First the underwriter sets an initial indicative price range, based on his prior beliefs. Then both bookbuilding and grey market trading begin. At the end of bookbuilding, the underwriter observes the bids in the book as well as the grey market price and sets the issue price. When the issue price is set, information about bookbuilding is revealed. Afterwards, aftermarket trading begins.

---

[6]Conversations with grey market brokers confirm that grey market investors are primarily smaller institutions and retail investors. However, in Section 6 we consider the case in which there is an overlap between grey market and bookbuilding investors.

[7]Cornelli and Goldreich (2003) show that bookbuilding aggregates information that is relevant for both the issue price and the long-run aftermarket price.

2.1. Bookbuilding. We model the bookbuilding process as in Biais and Faugeron-Crouzet (2002). Three investors take part in the bookbuilding process: two are informed and one is uninformed. The uninformed investor can buy at most $S(1-k)$ shares, with $0 < k < 1$. The information from bookbuilding, s_B, is the aggregate of the signals s_i observed by the two informed investors, $i = 1, 2$. s_i is i.i.d. and equals H with probability π and L with probability $(1-\pi)$. The distribution of the signals is such that if both investors observe a signal equal to H then $s_B = H$; if both observe L then $s_B = L$; but if one observes L and the other H then $s_B = M$, where $0 < L < M < H < \bar{v}$. The ex-ante expected value of the bookbuilding information is $E(s_B) = \pi^2 H + 2\pi(1-\pi)M + (1-\pi)^2 L$. For simplicity, we assume that $E(s_B) = M$.

From the point of view of an informed investor who observes s_i (as well as the grey market information s_G, which, as explained below, can be inferred from the grey market price) the expected fundamental value is

$$E(v \mid s_G, s_i) = \alpha s_G + (1 - \alpha)E(s_B \mid s_i) \tag{2}$$

Each informed investor submits a bid for shares, after observing his own signal, but not knowing the signal of the other informed investor. The underwriter designs a mechanism (described in Section 2.4) in which he sets the issue price and allocates the shares as a function of the bids. We assume that the underwriter acts in the issuer's interest, i.e., he maximizes IPO proceeds.

The bookbuilding set-up described so far is similar to the one in Benveniste and Spindt (1989). The main difference here is the existence of a grey market whose price is observed before the issue price is set.

2.2. Grey market. At the same time as bookbuilding there is a grey market in which investors trade the shares on a when-issued basis. For now, we assume that bookbuilding investors are not allowed to trade in the grey market.

Since in reality the grey market price is continuously and publicly observable, while bookbuilding is a confidential process controlled by the underwriter, we assume that grey market investors do not observe s_B. Instead, they only observe a signal about the value of the shares,

$s_G \in [0, \bar{v}]$.[8]

Grey market investors know that the fundamental value is a weighted average of their signal and the bookbuilding information, but we allow for the possibility that they overweight the importance of their signal. In other words, after observing s_G, their expectation of the fundamental value of the shares is

$$E_G(v \mid s_G) = \hat{\alpha} s_G + (1 - \hat{\alpha}) E(s_B) \tag{3}$$

where $\hat{\alpha} \geq \alpha$ and E_G refers to the expectation from the perspective of grey market investors. The difference $(\hat{\alpha} - \alpha)$ represents the extent to which grey market investors overweight their signal. If $\hat{\alpha} - \alpha > 0$ they are sentiment investors, while if $\hat{\alpha} - \alpha = 0$ they are fully rational. Note that only the expectation of s_B appears in equation (3), since grey market investors do not observe the bookbuilding information.

Trading in the grey market results in a price P_{GM}, reflecting investors' beliefs about the fundamental value of the shares. Thus, $P_{GM} = E_G(v \mid s_G)$. After observing P_{GM} the underwriter and the bookbuilding investors, knowing $\hat{\alpha}$, can perfectly infer s_G using the following relation:

$$s_G = \frac{P_{GM} - (1 - \hat{\alpha})M}{\hat{\alpha}} \tag{4}$$

After the underwriter aggregates the bookbuilding information into the issue price (and before the start of aftermarket trading), the bookbuilding information s_B is revealed.[9] Grey market investors then update their valuation, starting from their prior valuation, to

$$\hat{P}_{GM} \equiv \hat{P}_{GM}(s_G, s_B) = \hat{\alpha} s_G + (1 - \hat{\alpha}) s_B = P_{GM} + (1 - \hat{\alpha})(s_B - M). \tag{5}$$

$\hat{P}_{GM}$ differs from P_{GM} because it incorporates the observed s_B rather than its expectation. It differs from the fundamental value if $\hat{\alpha} \neq \alpha$. The difference between $\hat{\alpha}$ and α captures the extent to which grey market investors' expectations are biased (and thus they are sentiment

---

[8]The assumption that the signal of the grey market investors is continuous, while the signal of the bookbuilding investors is discrete, is made only in order to simplify the analysis.

[9]A more realistic assumption might be that grey market investors infer the information from the issue price. However, we assume s_B is revealed to avoid modelling situations in which the issue price is manipulated by the underwriter to hide information.

investors). This bias is supported by experimental evidence that individuals are slow to change their beliefs in the face of new evidence: they update their priors too little relative to Bayesian updating (see Barberis, Shleifer, and Vishny, 1998).[10]

2.3. Aftermarket. After the shares are allocated to the bookbuilding investors, trading in the aftermarket begins. At this point, both bookbuilding and grey market investors have observed both s_G and s_B. Grey market investors value the shares at $\hat{P}_{GM}$, while bookbuilding investors value the shares at the expected fundamental value, given in equation (1). Again, these two valuations differ if $\hat{\alpha} > \alpha$.

We assume that aftermarket participants include investors who have the same valuation as the grey market investors. They may be the grey market investors themselves or other (perhaps retail) investors who did not trade in the grey market. As a result, the grey market price is representative of the valuation of a larger set of investors. For simplicity, we continue to refer to this set of investors as grey market investors.[11]

Let P_{AM} denote the aftermarket price in the short-run. If the fundamental value of the shares exceeds the price grey market investors are willing to pay in the aftermarket, then the bookbuilding investors will not sell their shares to them. Thus, there will be no trading involving grey market investors and the aftermarket price will not depend on their valuation. The expected aftermarket price P_{AM} will then equal the expected fundamental value. If instead the price that grey market investors are willing to pay exceeds the fundamental value, the bookbuilding investors can sell their shares to the grey market investors at this higher price.

The price at which bookbuilding investors can sell their shares depends upon the depth of the market. If there are many investors willing to buy shares at the price $\hat{P}_{GM}$, bookbuilding investors will have all the market power and will set the price equal to $\hat{P}_{GM}$, extracting all the surplus from trading. However, if there are not enough investors willing to buy all S shares at $\hat{P}_{GM}$, for example if the demand for these shares is downward sloping, then bookbuilding investors will have to sell some of their shares at a lower price. Assuming a linear demand

---

[10]An alternative explanation could be the one studied in Harris and Raviv (1993), where individuals receiving common information differ in the way they interpret this information.

[11]Dorn (2003) finds a strong positive correlation between the volume of retail trade in the grey market and retail volume in the first-day of aftermarket trade. This supports our assumption that the opinion of grey market investors is indicative of the valuation of small investors in the aftermarket.

curve, bookbuilding investors expect to sell their shares in the aftermarket at $\hat{P}_{GM} - \lambda S$, where λS captures the discount necessary to sell all S in the aftermarket. Although more complex functional forms are possible, this simple linear form suffices to capture the idea that the market may not be very deep. If the market is deep enough to sell all the shares at $\hat{P}_{GM}$, then $\lambda = 0$.

In the long run, all uncertainty is resolved and the long-run price equals the fundamental value.

2.4. Optimal mechanism. We now characterize the optimal bookbuilding mechanism. This mechanism specifies the issue price and the number of shares to be allocated to the various bidders, as a function of their bids and the grey market price P_{GM}.[12]

To find the optimal mechanism, by the Revelation Principle we can restrict attention, without loss of generality, to a direct revelation mechanism where bookbuilding investors simultaneously announce their signals to the underwriter. The underwriter uses the announced signals $(\tilde{s}_1, \tilde{s}_2)$ (which aggregate to $\tilde{s}_B$) and the grey market information s_G to set the issue price and to allocate the shares. A direct revelation mechanism is described by the outcome functions (P_I, q, q_u), where $P_I(s_G, \tilde{s}_B)$ is the issue price; $q(s_G, \tilde{s}_i, \tilde{s}_j)$ is the allocation to an informed investor who announces signal $\tilde{s}_i$ when the other informed investor announces $\tilde{s}_j$; and $q_u(s_G, \tilde{s}_i, \tilde{s}_j)$ is the allocation to the uninformed investor when one informed investor announces $\tilde{s}_i$ and the other announces $\tilde{s}_j$. We look for an equilibrium of this mechanism in which buyers truthfully reveal their signals.

In order to derive the optimal mechanism, we must first determine the reservation price of the bookbuilding investors. After the issue price has been set, bookbuilding investors observe both s_B and s_G. They thus know the expected fundamental value of the shares (equation (1)), which is the value they will obtain if they hold the shares in the long run. Additionally, since they observe P_{GM}, they can also compute the price that grey market investors would be willing to pay in the aftermarket ($\hat{P}_{GM} - \lambda S$).

When $\hat{P}_{GM} - \lambda S$ exceeds the fundamental value, bookbuilding investors can sell their shares in the aftermarket to grey market investors, at a profit. Thus, the relevant value of the shares,

---

[12]By law, the underwriter cannot charge different prices to different investors, so we do not allow the mechanism to price discriminate among investors. Bennouri and Falconieri (2003) show that, if there is no limit to the quantities investors can be allocated, price discrimination is never optimal.

from their point of view, is the short-run aftermarket price. When $\hat{P}_{GM} - \lambda S$ is lower than the fundamental value, bookbuilding investors will not sell shares in the aftermarket. Therefore, after the end of bookbuilding, when a bookbuilding investor observes s_B, his valuation will be the maximum of the two possible valuations, i.e., $Max\{E(v \mid s_G, s_B), \hat{P}_{GM} - \lambda S\}$.

During bookbuilding, each informed investor will have observed his own signal s_i but not that of the other informed investor. Therefore, an investor with a signal $s_i = H$ will value a share at $\pi * Max\{E(v \mid s_G, H), \hat{P}_{GM} - \lambda S\} + (1 - \pi) * Max\{E(v \mid s_G, M), \hat{P}_{GM} - \lambda S\}$. An investor with a signal L will have a valuation of $\pi * Max\{E(v \mid s_G, M), \hat{P}_{GM} - \lambda S\} + (1 - \pi) * Max\{E(v \mid s_G, L), \hat{P}_{GM} - \lambda S\}$.

The underwriter, who also observes the grey market price, knows that if P_{GM} is high, he can increase the issue price above the fundamental value. That way, he can extract the surplus informed investors expect to gain from trading with the grey market investors in the aftermarket. Yet if P_{GM} is low, the underwriter does not need to lower the issue price, since he knows that the bookbuilding investors are still willing to buy and hold the shares at a price close to the fundamental value. In other words, the issue price (and the subsequent aftermarket price) will reflect the grey market price when it is high, but will reflect the fundamental value when the grey market price is low.

The following proposition presents the optimal mechanism.

Proposition 1: *Assume that $\hat{\alpha} \geq \frac{1-\alpha}{2}$ and that H and L are equidistant from M.[13] In the optimal mechanism the quantities allocated to the bookbuilding investors are:*

$$
\begin{aligned}
&q(H, H) = S/2, \quad q_u(H, H) = 0, \\
&q(H, L) = S, \quad q(L, H) = 0, \quad q_u(H, L) = 0, \\
&q(L, L) = Sk/2, \quad q_u(L, L) = S(1 - k)
\end{aligned}
\tag{6}
$$

where we have suppressed the argument s_G since the quantities allocated do not depend on the grey market information.

---

[13] *These assumptions are sufficient but by no means necessary for Proposition 1 to hold.*

The issue price depends on the grey market and the bookbuilding information as follows:

$$P_I(s_G, H) = V_H(H) - \frac{1-\pi}{\pi} k \left[V_M(L) - V_L(L)\right]$$
$$P_I(s_G, M) = V_M(M) \tag{7}$$
$$P_I(s_G, L) = V_L(L)$$

where

$$V_{s_B}(\tilde{s}_B) \equiv Max\{E(v \mid s_G, s_B), \hat{P}_{GM}(s_G, \tilde{s}_B) - \lambda S\}. \tag{8}$$

Proof: See Appendix 1.

$V_{s_B}(\tilde{s}_B)$ represents the value of a share to an informed investor where s_B is the aggregate information from the two signals of the informed investors and $\tilde{s}_B$ is the aggregate information revealed through bookbuilding (and thus conveyed to grey market investors). When $s_B \neq \tilde{s}_B$, at least one investor misreported his signal. $V_{s_B}(\tilde{s}_B)$ is the maximum of the fundamental value and the sale value. Note that our notation suppresses the argument P_{GM} in the function V_{s_B} as the maximization is conducted for a given P_{GM}.

Although the quantities allocated do not depend on the grey market price, the issue price P_I depends on P_{GM} as well as on the announced signals. To highlight how the issue price, as expressed in (7), depends on the grey market price and the information in the book, we can divide the possible values of P_{GM} into intervals and present the issue price for each interval. The boundaries of the intervals and the actual expression for the issue price in each interval are derived in Appendix 1 and represented in Figure 1.

The vertical axis represents the possible values that P_{GM} can take. The expression for P_I is different in each interval. Moreover, within each interval, the issue price depends on the bookbuilding information, s_B. Figure 1 displays the asymmetry in the optimal mechanism. Let us start by looking at the two extreme intervals. When P_{GM} is very high (i.e., above v_1, where v_1 is defined in Appendix 1) the issue price is close to the grey market price (more precisely, the updated valuation of the grey market investors) and does not depend at all on the bookbuilding information. This follows because the grey market price is so high that bookbuilding investors are certain they can sell their shares in the aftermarket for more than

the fundamental value, regardless of what the fundamental value is. The underwriter takes advantage of this by setting P_I above the fundamental value. When P_{GM} is very low (below v_4), P_I is based solely on the fundamentals. This is because the grey market valuation is so low that grey market investors are unwilling to buy shares in the aftermarket at a price above the fundamental value. In this case, the grey market is irrelevant and the mechanism is the standard bookbuilding mechanism. In the middle are cases in which the issue price is set equal to the fundamental value (minus an informational rent to be left to bookbuilding investors for revealing their information truthfully) when the fundamental value is high, and equal to the updated grey market valuation otherwise. So, on average, P_I is close to P_{GM} when P_{GM} is high, but not when P_{GM} is low, creating an asymmetry in the relation between the issue price and the grey market price.

2.5. *Discussion.* Central to the arguments presented in this paper and formalized in the model is the presence of an asymmetric relation between the grey market price and the issue and aftermarket prices. The extent of the asymmetry depends on the weights α and $\hat{\alpha}$. When $(\hat{\alpha} - \alpha)$ is small, grey market investors overweight their signal only by a small amount. In this case, their valuation and the fundamental value will be similar and there will be less asymmetry. When $(\hat{\alpha} - \alpha)$ is large, grey market investors are much more overconfident in the relevance of their signal and their valuation can be much higher than the fundamental value. In this case, the issuer can take advantage of their overconfidence and set the issue price much higher than the fundamental value.

Recall that α indicates the relevance of the grey market signal for the fundamental value. In the extreme case of $\alpha = 0$, the grey market price is not relevant for the fundamental value, which is completely determined by the bookbuilding information. In this case, when P_{GM} is low and the issue price is set equal to the fundamental value, there will be no relation between P_{GM} and the issue price. However, when the grey market price is high, it will be closely related to the issue price even though it contains no fundamental information.

On the other hand, when $\alpha > 0$ the grey market price does include fundamental information. In this case, there will be a positive relation between P_{GM} and the issue price even when the grey market price is low. However, to the extent that $\hat{\alpha} > \alpha$, this relation will be weaker than when P_{GM} is high, so the asymmetry remains.

The issue price is based on the valuation of bookbuilding investors, which in turn depends

on their expectation about aftermarket prices. Thus the asymmetry in the issue price is driven by the asymmetry in the aftermarket price P_{AM}, and there will be a stronger relation between P_{GM} and P_{AM} when P_{GM} is high.

Although we allow for shares to be sold in the aftermarket at a discount λS to capture a potential lack of depth, the results also hold when $\lambda = 0$. A larger λ implies that shares will be sold to grey market investors in the aftermarket less often, but the basic asymmetry remains.

As explained above, the bookbuilding investors' expectations about the aftermarket are central to determining their willingness to pay and the choice of issue price. This aspect is similar in spirit to Busaba and Chang's (2002) model where investors have an incentive to misreport their information to the underwriter in order to fool uninformed investors and take advantage of them in the aftermarket. However, there are two main differences here. First, in our model the underwriter designs the optimal mechanism to account for bookbuilding investors' incentives to sell their shares. As a result, bookbuilding investors will not misreport their information. Second, since the grey market price signals the potential sale value in the aftermarket, the bookbuilding investors and the underwriter can take it into account when determining their actions.

This is why the existence of the grey market is beneficial to the issuer even when it does not contain any information about the fundamental value of the shares ($\alpha = 0$). The valuation of the grey market investors affects the bookbuilding investors' valuation (because it affects the short-run aftermarket price) and thus provides a lower bound on their willingness to pay. If grey market investors are willing to pay a high price, the surplus that can be appropriated increases. Moreover, since this part of the valuation is publicly observable, the issuer can extract a larger part of the trading surplus from the bookbuilding investors.

3. Empirical implications

The model allows us to make predictions about the relation between the grey market price P_{GM}, the issue price P_I, and the aftermarket prices in the short and long run, as well as other variables. Here we list the main empirical predictions. A more detailed analysis is conducted in Section 5, where we present the results.

Hypothesis 1: P_I is positively correlated with P_{GM}. Moreover, $\hat{\alpha} > \alpha$ implies that this correlation is larger when P_{GM} is high. $\alpha > 0$ implies that this correlation is positive even when P_{GM} is low.

Hypothesis 2: The short-run aftermarket price P_{AM} is positively correlated with P_{GM}. Moreover, $\hat{\alpha} > \alpha$ implies that this correlation is larger when P_{GM} is high. $\alpha > 0$ implies that this correlation is positive even when P_{GM} is low.

Hypothesis 3: When the reliability of the grey market signal s_G increases, the correlations of P_{GM} with P_I and P_{AM} increase.

A more reliable grey market signal s_G means that an investor should give additional weight to s_G. In other words, α should be higher. Thus, P_{GM} will be more closely related to the fundamental value. While the model does not necessarily imply it, presumably $\hat{\alpha}$ will also be higher in this case.

Hypothesis 4: When P_{GM} is high, P_I and P_{AM} are negatively correlated with the issue size (S) and positively correlated with the depth of the grey market ($-\lambda$).

Hypothesis 5: Aftermarket trading volume is higher when P_{GM} is high than when P_{GM} is low, since when P_{GM} is high bookbuilding investors sell their shares to grey market investors in the aftermarket.

Finally, the model has implications for long-run returns. Intuitively, when P_{GM} exceeds the fundamental value, the immediate aftermarket price (P_{AM}) is closely related to the grey market investors' willingness to pay ($\hat{P}_{GM}$), which differs from the fundamental value if grey market investors overweight their own signal, i.e., if $\hat{\alpha} > \alpha$. In this case, we expect reversal of the share price towards the fundamental value in the long run. In contrast, the difference between P_{AM} and P_{GM} captures grey market investors updating their valuation when they learn the bookbuilding information s_B. To the extent that they underweight the bookbuilding information, the share price movement from P_{GM} to P_{AM} is only a partial movement in the right direction and should continue in the same direction as the fundamental value is revealed over time. This discussion is summarized as:

Hypothesis 6: When P_{GM} is high, the long-run return (relative to P_{AM}) is negatively correlated with P_{GM} and positively correlated with the difference between P_{AM} and P_{GM} (to the extent that grey market investors do not fully update for s_B, i.e., if $\hat{\alpha} > \alpha$).

The correlations predicted in Hypothesis 6 can be derived more formally from the model as follows. The share price converges in the long run to the fundamental value $\alpha s_B + (1 - \alpha)s_G$. Recall that when P_{GM} is high, the short-run aftermarket price is $P_{AM} = \hat{P}_{GM} - \lambda S = \hat{\alpha}s_B + (1 - \hat{\alpha})s_G - \lambda S$ (i.e., the reservation price of the grey market investors). Thus the difference between the long-run price and the short-run price is $(\hat{\alpha} - \alpha)(s_B - s_G) + \lambda S$. This is the long-run return in dollars. To the extent that $\hat{\alpha} > \alpha$, this return is positively related to the bookbuilding signal s_B and negatively related to the grey market signal s_G. If $\hat{\alpha} = \alpha$, the short-run aftermarket price is already the expected fundamental value, so the long-run returns are zero.

The difference between P_{AM} and P_{GM} (for high P_{GM}) is $(1 - \hat{\alpha})(s_B - E(s_B)) - \lambda S$, which is also positively related to the bookbuilding signal. Thus, there is a positive relation between the long-run return and $(P_{AM} - P_{GM})$ if and only if $\hat{\alpha} > \alpha$.

Finally, P_{GM} equals $\hat{\alpha}s_G + (1 - \hat{\alpha})E(s_B)$ and so it is also related to the grey market signal s_G. Thus, the long-run return should be negatively related to P_{GM} as long as $\hat{\alpha} > \alpha$.

4. Sample and data

The dataset consists of 486 companies which went public in twelve European countries between November 1995 and December 2002 and for which we have grey market prices. The extent to which IPO shares are traded in grey markets varies from country to country. For instance, in Germany and Italy most IPOs trade in the grey market while in France or Sweden very few do. (As we will see, outside Germany and Italy, grey market trading is more likely in larger IPOs.) As a result, our dataset is a subset of the universe of 2,723 firms going public in the twelve countries over the sample period.

While we only consider firms that go public in Europe, our sample does include some non-European companies that obtained a first-time listing in a European country (typically Germany's *Neuer Markt*). Sample companies thus come from a total of 20 countries: Austria (13), Belgium (1), Canada (1), Denmark (1), Finland (3), France (13), Germany (321), Greece (2), Ireland (2), Israel (7), Italy (61), Lithuania (1), Luxembourg (1), Netherlands (11), Norway (2), Spain (5), Sweden (2), Switzerland (11), the United Kingdom (24), and the United States (4).

Grey markets are usually organized not by an exchange but by independent brokers who

make forward markets in IPO shares on a when-issued basis. Thus, the structure of grey markets differs across countries and even within countries depending on the broker. Brokers quote spreads and investors can take a long or short position depending on their expectations. Usually, grey market prices are public information: not only are they available from the broker, but they are often reported in the financial news media.

Grey market trading typically begins on the day the company publishes its initial indicative price range and concludes on the day before the stock begins trading on the stock market. Often, IPOs are priced a day or two before stock market trading begins, in which case grey market trading continues for a short while *after* the IPO has been priced.

Grey market prices were obtained from two large brokers, based in Germany and the United Kingdom, and supplemented with a news search. For every company in our sample, we have the last grey market price established *before* the IPO is priced, and for 262 companies we also have post-pricing grey market prices.[14]

Information on the IPOs is derived from an updated version of the dataset compiled by Ljungqvist and Wilhelm (2002), based on Dealogic's Equityware, Thomson Financial's SDC, information from national exchanges, and a comprehensive news search. Firm and offer characteristics are taken from the IPO prospectuses. Aftermarket trading prices and trading volumes are from Datastream. We convert monetary values – such as gross proceeds – into U.S. dollars using exchange rates on the first day of trading.

Table 1 shows descriptive statistics for the sample as a whole as well as broken down by the twelve countries on whose exchanges sample companies list. Most sample firms (75 percent) list in Germany, 54 companies list in more than one country (usually the home country plus Frankfurt or London), and 43 companies do not list in their home country at all (including non-European issuers from Israel, the U.S., and Eastern Europe).

Although the sample IPOs span the period from November 1995 to December 2002, the range of dates varies from market to market, depending on the IPOs for which we have grey market prices. In the UK, for instance, we have grey market prices for firms going public between June 1997 and July 2002. To allow the reader to assess how comprehensive our sample is, Table 1 reports the number of IPOs in each market during the entire period, as well

---

[14]Other than this, we do not have a time series of grey market prices for our sample companies. For an analysis of a limited sample of daily grey market prices, see Löffler, Panther, and Theissen (2002).

as during the sub-periods for which we have IPOs with grey market prices for each country.

Over our sample period, Germany and Italy have the most active grey markets, while London-based brokers frequently make grey markets in IPOs taking place in other countries. Except in Germany, grey market trading is more common in larger IPOs. Reflecting the fact that many of our sample IPOs were completed in the late 1990s, the initial return (P_{AM}/P_I-1) averages 36.3%. Bid-ask spreads in the grey market are quite wide, with quoted half spreads averaging 4.7%. Just over half the IPOs (54.1%) are priced at the high end of the initial indicative price range. On average, the last grey market price before the issue price is finalized exceeds the price range midpoint by 40.4%.

5. Empirical results

We now discuss the empirical results in light of our predictions. Since we pool data from several countries whose grey market and bookbuilding practices likely differ in subtle ways, we initially estimated all our models with country fixed effects but found these to be insignificant. Similarly, we obtain qualitatively similar results if we restrict the sample to firms going public in Germany, which has the most active grey market in our sample. We have also verified that our results are robust to outliers by winsorizing the price data at the 5% level. To conserve space, neither of these robustness tests is reported.

5.1. The offer price. Hypothesis 1 discusses the relation between the issue price and the grey market price. We normalize each by the midpoint of the initial indicative price range, P_{mid}, in order to reduce the impact of differences in scale and of heteroskedasticity. The grey market price that we use is the last reported transaction price before the issue price is set (or the midpoint of the bid-ask spread when transaction prices are unavailable). This corresponds to P_{GM} in the model.

It is well-documented that issue prices in Europe are rarely set outside the initial indicative price range; frequently they are set at the endpoints, especially at the top of the range (see Ljungqvist, Jenkinson, and Wilhelm (2003)). Consequently, the observed distribution of issue prices in our sample is censored at the range endpoints. To correct for this, we estimate censored regressions (Amemiya (1973)), with censoring both from above and from below. Censored regressions are similar to Tobit models, except that the point of censoring is

observation-specific. Note that 54.1% of our observations are right-censored, while 10.5% are left-censored.

Our model distinguishes between the cases where the grey market price is higher or lower than the fundamental value. Because fundamental value is unobservable to the econometrician, empirical studies usually take the midpoint of the initial indicative price range as a proxy for the underwriter's ex ante prior of the fundamental value. Thus if the grey market price is above the midpoint it is more likely to be above the fundamentals. To capture the predicted asymmetry, we interact the grey market price with an indicator function that equals one if the grey market price exceeds the range midpoint, and zero otherwise.[15]

Table 2 reports the results. Regression 1 examines the relation between the issue price P_I and the grey market price P_{GM}. Overall, the fit of the model is very good in view of the highly significant likelihood ratio test. As expected, we find a very significant relation between the issue price and the grey market price, and an even stronger relation when the grey market price is above the range midpoint, P_{mid}. This result is consistent with Hypothesis 1: the grey market price is positively correlated with the issue price. The fact that the correlation is positive even when the grey market price is low suggests that $\alpha > 0$: the grey market price contains information about the fundamental value. The stronger relation for high P_{GM} reflects an asymmetry in the relation between the grey market price and the issue price. This is consistent with the asymmetry in the model when $\hat{\alpha} > \alpha$, i.e., the underwriter bases the issue price on the grey market price when it is high, but when the grey market price is low, the underwriter uses it only to the extent that it contains (partial) information about the fundamental value. This suggests that not only are grey market investors sentiment investors, but also the underwriter and the bookbuilding investors are aware that the grey market price includes a bias. In fact, they include this bias in their valuation only when the bookbuilding investors can profit from it by selling shares to grey market investors in the aftermarket.

Regression 1 also includes the market index return (measured over the three-month period prior to the IPO) as a control variable. This variable has previously been associated with market sentiment (see, for instance, Derrien (2004)). We find a positive and statistically

---

[15]The large proportion of right-censored observations is the reason why we introduce the indicator function to capture the asymmetry rather than splitting the sample between high and low levels of the grey market price, as we do in later tables. If we were to estimate the censored regression model for the subsample where the grey market price exceeds the range midpoint, we would have little explanatory power since for most observations the issue price would equal the top of the range.

significant coefficient and yet we still find a strong positive and asymmetric relation between the grey market price and the issue price. This suggests that the grey market price contains additional information beyond the market-wide returns. This is not surprising, since market returns surely reflect more than just sentiment and so are at best a noisy proxy for investor sentiment, especially at the level of individual securities.

In Regression 2 of Table 2 we add the (logarithm of) gross issue proceeds and the bid-ask spread quoted by grey market brokers shortly before IPO pricing. A wider bid-ask spread may indicate a lack of depth in the grey market, due to either a scarcity of traders in the grey market or a diversity of opinion among investors.[16] Either way, bookbuilding investors may not be able to sell all their shares in the aftermarket at the (updated) grey market price $\hat{P}_{GM}$, causing the underwriter to price the IPO more conservatively. Similarly, when the issue size is large, the issue price should reflect the greater difficulty of selling the shares in the aftermarket. In the model, this is captured by the discount λS and gives rise to Hypothesis 4, which suggests the issue price should be negatively correlated with the issue size S and positively correlated with the depth of the market $(-\lambda)$.

Consistent with Hypothesis 4, in Regression 2 we find negative and statistically significant relations between the issue price P_I and the bid-ask spread, and between P_I and log proceeds S.

In Regression 3 we interact the bid-ask spread with log proceeds, since in the model the discount λS is the product of the two terms. We find a negative and statistically significant relation. The relations in Hypothesis 4 refer to the case when P_{GM} is high, since only then do bookbuilding investors sell their shares to grey market investors. To capture this asymmetry, Regression 4 includes the product of the bid-ask spread and log proceeds times an indicator function that equals one when the grey market price is above the range midpoint. Both the resulting coefficients are negative, but not statistically significant at conventional levels.

Finally, if a wide bid-ask spread also reflects greater divergence of opinion among grey market investors, it may indicate a less reliable grey market signal (i.e., a smaller α in the model). Hypothesis 3 predicts that the correlation between the issue price and the grey market price is weaker when the grey market signal is less reliable. In Regression 5, we attempt to

---

[16]An alternative measure of depth is trading volume in the grey market. However, grey market volume data are not available on a systematic basis.

capture this by interacting the bid-ask spread with the grey market price. We find that the coefficient of the interaction term is indeed negative and statistically significant, suggesting that the positive effect of the grey market price on IPO pricing is attenuated when the bid-ask spread is wider.

5.2.   The short-run aftermarket price.   The model suggests that when the grey market price exceeds the fundamental value, the first-day closing price P_{AM} reflects the price grey market investors are willing to pay for the shares (i.e., the grey market price adjusted for the information learned when the issue price is set). In Table 3, Regression 1, we see that the first-day closing price is indeed highly correlated with the grey market price, with a coefficient close to one. The adjusted R^2 of 75.4% indicates that we capture a sizable part of the variation in aftermarket prices using only information available before aftermarket trading begins.

Since Table 2 demonstrates a positive relation between the grey market price P_{GM} and the issue price P_I, it is possible that Regression 1 simply captures the well-documented positive relation between issue prices and after-market prices (Hanley (1993)): issue prices contain bookbuilding information and so affect aftermarket prices. To investigate this further, Regression 2 relates the aftermarket price to the issue price. As expected, we find that the aftermarket price is higher, the higher is P_I. The adjusted R^2, however, is much lower than in Regression 1. In Regression 3, we include both the grey market price and the issue price. The coefficient estimated for P_{GM} remains highly significant, but P_I loses most of its explanatory power. In sum, high grey market prices predict aftermarket prices independent of the level of the issue price.[17]

It is interesting to see how the coefficient of the prior three-month market index return varies across Regressions 1 through 3. Although the coefficient is both economically and statistically significant in Regression 2, it loses all its significance when the grey market price is included in Regressions 1 or 3. This suggests that while market-wide returns may capture general investor sentiment, as Derrien (2004) argues, they do not capture investor sentiment about specific IPOs very well – and certainly much less well than P_{GM}.

---

[17]Note that while the issue price depends on the grey market price, and the aftermarket price depends on the issue price and the grey market price, the system described by these two equations is triangular. Thus it can be consistently estimated recursively, that is, by equation-by-equation estimation. See Greene (2003), p. 383.

So far, our results might be interpreted as evidence that P_{GM} is a good predictor of the aftermarket price in general. However, Hypothesis 2 suggests an asymmetry: P_{GM} should be a better predictor of P_{AM} when P_{GM} is high, since in this case P_{AM} reflects grey market investors' (updated) valuation rather than fundamental value. When P_{GM} is low, on the other hand, P_{AM} equals the expected fundamental value, so the grey market price should only be related to P_{AM} to the extent that P_{GM} contains information about the fundamental value.

At the same time, we also expect an asymmetry in the effect of the issue price on the aftermarket price. Specifically, P_I should have a relatively stronger effect when P_{GM} is low. When P_{GM} is high, the bookbuilding information that is incorporated in the issue price affects the aftermarket price only through the updating of the grey market investors' valuation. When P_{GM} is low, the information in the book about the fundamental value is incorporated directly into the aftermarket price.

In Regressions 4 and 5 we capture the asymmetry by splitting the sample into two subsets based on whether P_{GM} is above or below the midpoint of the initial indicative price range, P_{mid}. We find a stronger relation between the aftermarket price and the grey market price when P_{GM} exceeds P_{mid} (coefficient of 0.95) than when it is below P_{mid} (0.56), consistent with Hypothesis 2. The fact that the coefficient is larger when P_{GM} is high implies that $\hat{\alpha} > \alpha$ (i.e., grey market investors are biased). The fact that the coefficient is positive and significant even when P_{GM} is low is consistent with $\alpha > 0$ (i.e., the grey market price contains fundamental information). As for the issue price, the coefficient is very similar in the two cases but it is more statistically significant (at 1%) when P_{GM} is low, as expected.

According to Hypothesis 4, when P_{GM} is high, a wide bid-ask spread or a large issue size reduces bookbuilding investors' ability to sell their shares in the aftermarket at the grey market price. The negative coefficients for these variables in Regression 4 support this prediction, though only the coefficient of issue size is statistically significant. When the grey market price is below the range midpoint, on the other hand, neither the bid-ask spread nor log proceeds have a significant effect on the aftermarket price, as expected (Regression 5).

5.3. Robustness: IPO withdrawals. Until now we have ignored the possibility that IPOs could be withdrawn after the start of grey market trading. If a combination of negative sentiment in the grey market and negative information in bookbuilding leads to IPOs being withdrawn, the remaining observations with a low P_{GM} would tend to have positive book-

building information. This could potentially bias the results in the direction of the asymmetry in the relations between P_{GM} and P_I, and between P_{GM} and P_{AM}. Since we do not observe P_I and P_{AM} for withdrawn IPOs, the distribution of observed prices has truncated support with the usual result that regression coefficients may be estimated with bias (Heckman (1979)).

It is an empirical question whether IPOs are withdrawn in response to negative sentiment in the grey market. Boehmer and Ljungqvist (2004) investigate a sample of 330 privately-held German firms that between 1984 and 1995 announced their intention to go public, a third of which remained private as of December 1999. In their sample, all but one company withdrew at a very preliminary stage, well before IPO marketing and bookbuilding (and thus grey market trading) had begun.

To investigate the possible extent of bias for our (later) sample period, we estimate the frequency with which IPOs are withdrawn after grey market trading has begun in Germany, the most active grey market in our sample. Between 1997 and 2002, there were 485 completed IPOs in Germany. Over the same period, a further 236 companies announced their intention to go public (according to Reuters and VWD, a German news wire service). Of these 236 withdrawn issues, only 20 (8.5%) were withdrawn *after* grey market trading had begun.[18] Thus, as in Boehmer and Ljungqvist's (2004) sample, the vast majority of IPOs are withdrawn at a very preliminary stage, and not in response to negative sentiment in the grey market. Therefore, our results should not be affected.

5.4. Updating. Our data allow us to investigate the extent to which grey market investors update their valuations upon learning the outcome of bookbuilding. Often, grey market trading continues for a short time after bookbuilding concludes and the issue price is set (but before aftermarket trading begins). For a subsample of 262 IPOs, we observe post-bookbuilding grey market prices, which correspond to $\hat{P}_{GM}$ in the model. To see if grey market investors incorporate the bookbuilding information revealed through the issue price P_I, we regress $\hat{P}_{GM}$ on P_I and P_{GM} (normalizing all three prices by the midpoint of the price range, P_{mid}). The estimated equation is:

$$\hat{P}_{GM}/P_{mid} = \underset{(-2.41)}{-0.14} + \underset{(2.96)}{0.23}\,P_I/P_{mid} + \underset{(33.82)}{0.92}\,P_{GM}/P_{mid}$$

---

[18]Among these 20 withdrawn IPOs, the last recorded grey market price before withdrawal was, on average, 8.3% below the midpoint of the initial indicative price range. Six of the 20 firms traded above the range midpoint at the time of withdrawal.

where heteroskedasticity-consistent t-statistics are shown in parentheses underneath the OLS coefficient estimates. The adjusted R^2 is 96.9%. The coefficient estimated for P_{GM} is significantly less than one ($p = 0.004$), while the coefficient for P_I is significantly greater than zero ($p = 0.003$). This suggests that grey market investors do adjust their expectations, and that bookbuilding information is incorporated in $\hat{P}_{GM}$.

The following alternative specification quantifies the extent to which grey market investors update upon learning P_I:

$$(\hat{P}_{GM} - P_{GM})/P_{mid} = \underset{(1.21)}{0.01} + \underset{(2.80)}{0.07}(P_I - P_{GM})/P_{mid}$$

The adjusted R^2 in this specification is 14.4%. The coefficient estimated for $(P_I - P_{GM})/P_{mid}$ suggests that for every dollar difference between P_I and P_{GM}, grey market investors increase their reservation price by seven cents. So although we find that grey market investors update when they observe the results of bookbuilding, they only update by a relatively small amount, consistent with conservatism.

5.5. Aftermarket trading volume. Table 4 examines the relation between the grey market price and aftermarket trading volume (as a fraction of the shares sold in the IPO). Hypothesis 5 suggests that the relation should be a step function. When P_{GM} is high, we expect high turnover because bookbuilding investors sell their shares to the grey market investors whose valuation exceeds the fundamental value. When P_{GM} is low, bookbuilding investors have no reason to sell their shares in our model and so trading volume will be lower.

We measure aftermarket trading volume both on the first day and over the first week following the IPO. To capture the step function, we use an indicator function that equals one when P_{GM} is above the initial price range midpoint P_{mid}, and zero otherwise. We find a positive and statistically significant relation between volume and the indicator function, both for first-day volume (Regression 1) and first-week volume (Regression 4). This suggests that when P_{GM} is high, bookbuilding investors are more likely to sell their shares in the aftermarket, consistent with Hypothesis 5.

A high grey market price could simply indicate that either the IPO or the market is 'hot,' which may lead to high volume for reasons outside our model. In Regressions 2 and 5 we include the market index return (measured over the three-month period before the IPO) to

capture a 'hot' market. In Regressions 3 and 6 we also include the (normalized) first-day closing market price P_{AM} to capture whether the IPO is 'hot.' Even after including these variables, the coefficient on the indicator function remains positive and significant. This implies that the positive relation between volume and the indicator function is not simply due to a high level of trading in 'hot' IPOs or in active markets.

5.6. Long-run returns. We now consider how the grey market price and the results of book-building are related to aftermarket returns in the long run. We test these relations using the following regression:

$$\frac{P_{LongRun} - P_{AM}}{P_{mid}} - \text{market index return} = \alpha + \beta_1 \frac{P_{GM} - P_{mid}}{P_{mid}} + \beta_2 \frac{P_{AM} - P_{GM}}{P_{mid}} + \epsilon$$
(9)

The dependent variable is the return measured from the end of the first aftermarket trading day until two, three, six or twelve months later (less the return on the domestic stock market index). In our model, $P_{LongRun} - P_{AM} = (\hat{\alpha} - \alpha)(s_B - s_G) + \lambda S$ when P_{GM} is above the fundamental value. When P_{GM} is below the fundamental value, $P_{LongRun} - P_{AM} = 0$. As before, we normalize all variables by P_{mid}, the midpoint of the initial indicative price range, in order to reduce the impact of differences in scale and of heteroskedasticity.[19]

The independent variables are the difference between the grey market price and the range midpoint $(P_{GM} - P_{mid})$ and the difference between the aftermarket price on the first trading day and the grey market price $(P_{AM} - P_{GM})$, both normalized by the range midpoint. Together, these two variables add up to the entire price movement from the prior expected value of the shares, P_{mid}, to the price at the end of the first day of aftermarket trading, P_{AM}.

By splitting the price movement in this way, we can relate long-run returns separately to the two different signals, s_G and s_B, of our model. $P_{GM} - P_{mid}$ reflects the information revealed through grey market trading. Specifically, $P_{GM} - P_{mid} = \hat{\alpha}(s_G - E(s_G))$ where P_{mid} is the ex-ante expected value, $E(v)$. The second variable, $P_{AM} - P_{GM}$, captures the price movement that occurs in response to bookbuilding being concluded and the issue price being set. When the grey market price is high, $P_{AM} - P_{GM} = (1 - \hat{\alpha}(s_B - E(s_B)) - \lambda S$, representing

---

[19]We normalize all variables by the same price, since this allows us to write the coefficients as simple functions of the model parameters.

the change in valuation due to the revelation of bookbuilding information (assuming that no other information arrives in this short interval).

According to Hypothesis 6, long-run returns relate differently to the grey market signal and the bookbuilding signal. When P_{GM} exceeds the fundamental value, the first-day aftermarket price will be close to the grey market investors' reservation price. To the extent that grey market investors overweight the grey market signal s_G (i.e., $\hat{\alpha} > \alpha$), their reservation price diverges from the fundamental value. Therefore, since the price should eventually revert to the fundamental value, Hypothesis 6 predicts a negative relation between long-run returns and the difference between the grey market price and the range midpoint. Specifically, the coefficient β_1 in (9) equals $-\frac{\hat{\alpha}-\alpha}{\hat{\alpha}} = \frac{\alpha}{\hat{\alpha}} - 1$ which is negative if $\hat{\alpha} > \alpha$.

Bookbuilding information, by contrast, is assumed to be about fundamental value. If so, the difference between the aftermarket price and the grey market price should not be reversed in the long run, that is, it should not be negatively correlated with long-run returns. Whether this correlation is zero or positive depends on how grey market investors update using the bookbuilding information, s_B. If they fully update after observing s_B (i.e., $\hat{\alpha} = \alpha$), there should be no correlation. If they instead exhibit conservatism (i.e., $\hat{\alpha} > \alpha$), then the movement from the grey market price to the first-day aftermarket price is only a partial movement towards the fundamental value (assuming P_{GM} is above the fundamental value). As the market updates further over time, the price continues to move towards the fundamental value. Thus, when P_{GM} is above the fundamental value and $\hat{\alpha} > \alpha$, we expect a positive correlation between long-run returns and the difference between the aftermarket price and the grey market price: $\beta_2 = \frac{\hat{\alpha}-\alpha}{1-\hat{\alpha}} > 0$ in (9).

On the other hand, when P_{GM} is below the fundamental value, P_{AM} already reflects the expected fundamental value so we expect neither reversal nor continuance in the long run.

In Table 5, we present the results of regression (9) for the full sample as well as a partition of the sample based on whether P_{GM} is above or below P_{mid}. The predicted relations should hold only when the grey market price is high, since only then does the first-day aftermarket price relate to the grey market investors' reservation value.

In the full sample, for all horizons, we find a statistically significant negative relation between $P_{GM} - P_{mid}$ and long-run returns. However, when we partition the sample, we find that the negative relation only holds when $P_{GM} > P_{mid}$, as predicted by Hypothesis 6. Since

$\beta_1 = \frac{\alpha}{\hat{\alpha}} - 1$ when P_{GM} is high, this suggests that $\hat{\alpha} > \alpha$: grey market investors overweight their signal. Thus, non-fundamental information is transmitted from the grey market to the aftermarket and is reversed in the long run. Moreover, depending on the horizon, the coefficients range from -0.25 to -0.77 (all significantly greater than -1), indicating that only part of the price difference between P_{GM} and P_{mid} is reversed. This can be interpreted as evidence that P_{GM} contains some information about the fundamental value (i.e., $\alpha > 0$) and so does not need to be reversed completely. When $P_{GM} < P_{mid}$, we do not find any reversal, consistent with Hypothesis 6.

Our second variable, $P_{AM} - P_{GM}$, is not negatively related to long-run returns. As argued earlier, this is consistent with the hypothesis that the information in the book pertains to the fundamental value and is not reversed in the long run. The coefficient β_2 is positive when the grey market price is high, consistent with investors updating only gradually over time, but it is mostly not statistically significant.

6. Allowing bookbuilding investors to trade in the grey market

In Section 2 we assumed that bookbuilding investors are not allowed to trade in the grey market. In reality, bookbuilding investors are able to participate in the grey market, but the underwriter actively discourages it (for example, by threatening to exclude them from future IPOs). Since the underwriter cannot directly observe whether a bookbuilding investor trades in the grey market, it is unclear how effective a prohibition would be. In this section we explicitly consider the possibility that bookbuilding investors trade in the grey market, and show that a bookbuilding investor with a position in the grey market may have additional incentives not to report his signal truthfully. To avoid this problem, the underwriter might prohibit the bookbuilding investors from trading in the grey market. If he cannot enforce such a prohibition, he has to modify the mechanism in a revenue-decreasing way.

To see whether trading in the grey market by bookbuilding investors reduces proceeds, we start by assuming the underwriter is using the mechanism in Proposition 1 and then look at whether it affects the incentive compatibility constraints in the maximization problem.

For simplicity, we consider the case in which bookbuilding investors can buy or sell only a limited number of shares, γ, in the grey market, so that they are price takers. That is, their trades do not affect the grey market price.

The details of the argument are given in Appendix 2. Here we sketch the main steps and give the intuition of the results. We solve the model backward: assuming that a bookbuilding investor has taken a position (long or short) in the grey market, we look at how his incentive to truthfully report his signal changes. Then, we look at the investor's decision to buy or short-sell shares in the grey market, given that he can anticipate the signal he will report during bookbuilding.

A bookbuilding investor who observed a signal H and has gone long in the grey market will always reveal his signal truthfully since this can only have the effect of raising the sale price in the aftermarket. However, if he has a short position in the grey market he may have an incentive to lie and report a signal L. If the grey market price is sufficiently high, he will have an incentive to report a signal L in order to manipulate the price downward and profit from his initial short position. But if the grey market price is low, the aftermarket price does not depend on the updated valuation of the grey market investors, so he has no incentive to misreport his signal. In contrast, a bookbuilding investor who observed a signal L, regardless of his position in the grey market, will never be induced to lie by a small position in the grey market.

Given these incentives to a bookbuilding investor who observed H, we look at whether he prefers to take a long position in the grey market (and truthfully reveal his signal) or whether he prefers to take a short position (and, when P_{GM} is high, misreport his signal). We find that the choice depends on the parameter λ.

If there is no lack of depth in the aftermarket ($\lambda = 0$), i.e., the bookbuilding investor knows that there are sufficient investors to whom he can sell shares at $\hat{P}_{GM}$ in the aftermarket, then an investor who has observed a signal H will find it more profitable to buy shares than to short sell them in the grey market. In such case, the investor will truthfully report his signal, so the underwriter need not be concerned if bookbuilding investors trade in the grey market.

The only situation in which trading in the grey market by bookbuilding investors could hurt the issuer arises when $\lambda > 0$, i.e., there is insufficient depth in the aftermarket. In this case, for sufficiently high P_{GM}, bookbuilding investors take advantage of the discount in the aftermarket (λS) by short selling in the grey market and covering this position at a lower price.

Therefore, although the existence of the grey market is beneficial for the issuer in that it

increases the expected IPO proceeds (as shown in Section 2), participation by bookbuilding investors in the grey market may not be. This result rationalizes investment bankers' efforts to discourage bookbuilding investors from participating in the grey market, particularly when P_{GM} is high. But the underwriter should only be concerned about bookbuilding investors participating in the grey market when they short sell and the aftermarket is expected to be thin.

7. Conclusion

We have taken advantage of the existence of a grey market for shares of companies about to go public to test whether behavioral biases among small investors can explain the well-known anomalies in post-IPO prices. When small investors are excessively optimistic, they are willing to pay a price above the fundamental value, resulting in a high aftermarket price. When they are pessimistic, and value the shares below the fundamental value, they will be priced out of the market, in which case we predict no bias in the aftermarket price. This argument implies an asymmetric relation between the grey market price and the aftermarket price.

Using grey market price data for a large set of European IPOs, we find evidence of such an asymmetric relation. Moreover, when the grey market price is high, we find that long-run returns are negatively correlated with the grey market price, while this pattern does not arise when the grey market price is low. This suggests that when small investors drive the price upward in the short-run aftermarket, there is a reversal as the price converges towards the fundamental value in the long run.

The combination of the asymmetric effect of the grey market price and the long-run reversal provides evidence on the existence of sentiment investors. Moreover, the fact that when these sentiment investors are pessimistic about an issue their opinion has less of an effect on the aftermarket and issue prices, suggests that the underwriter and sophisticated investors can identify sentiment investors. Thus, they consider the opinion of sentiment investors biased, and they take it into account only when they can profit from it, by selling shares to them in the aftermarket.

Appendix 1

Proof of Proposition 1

Step 1. We focus on symmetric equilibria, i.e., informed investors are treated symmetrically. For a given s_G define $P_{\tilde{s}_B} \equiv P_I(s_G, \tilde{s}_B)$ as the issue price set when the aggregate information revealed through bookbuilding is $\tilde{s}_B$. The underwriter chooses the mechanism that maximizes expected proceeds:

$$\max_{P_{\tilde{s}_B}, q} S \left[\pi^2 P_H + 2\pi(1-\pi)P_M + (1-\pi)^2 P_L \right] \tag{10}$$

subject to the individual rationality constraints

$$V_{s_B}(s_B) \geq P_{\tilde{s}_B}, \tag{11}$$

the incentive compatibility constraint for informed investor i who observes a signal $s_i = H$,

$$\pi \frac{S - q_u(H,H)}{2} \left[V_H(H) - P_H \right] + (1-\pi) \left[S - q_u(H,L) - q(L,H) \right] \left[V_M(M) - P_M \right]$$
$$\geq \pi q(L,H) \left[V_H(M) - P_M \right] + (1-\pi) \frac{S - q_u(L,L)}{2} \left[V_M(L) - P_L \right] \tag{12}$$

and the incentive compatibility constraint for informed investor i who observes a signal $s_i = L$,

$$\pi q(L,H) \left[V_M(M) - P_M \right] + (1-\pi) \frac{S - q_u(L,L)}{2} \left[V_L(L) - P_L \right]$$
$$\geq \pi \frac{S - q_u(H,H)}{2} \left[V_M(H) - P_H \right] + (1-\pi) \left[S - q_u(H,L) - q(L,H) \right] \left[V_L(M) - P_M \right] \tag{13}$$

where the quantities are written so that the sum of the shares allocated to the two informed investors and the uninformed investor equals S.

We proceed as follows: we first ignore the second incentive compatibility constraint (13) and find the optimal solution. We then check that this constraint is in fact non-binding at the optimum.

Step 2. Since the number of shares sold always equals S and the underwriter charges all investors the same issue price, the quantity allocated to each investor does not directly affect proceeds. Thus, we choose the quantities to relax the incentive compatibility constraint (12) as much as possible since a slacker constraint will allow an optimum with a higher price.

This is achieved by setting $q_u(H, H) = q_u(H, L) = 0; q_u(L, L) = S(1 - k); q(H, H) = S/2; q(H, L) = S; q(L, H) = 0;$ and $q(L, L) = Sk/2$. Substituting and dividing both sides by S, the incentive compatibility constraint for type H (equation (12)) becomes

$$\frac{\pi}{2} [V_H(H) - P_H] + (1 - \pi) [V_M(M) - P_M] \geq (1 - \pi)\frac{k}{2} [V_M(L) - P_L] \tag{14}$$

When this incentive compatibility constraint is binding, we can write it as

$$P_H = V_H(H) + \frac{2(1 - \pi)}{\pi} [V_M(M) - P_M] - \frac{1 - \pi}{\pi} k [V_M(L) - P_L] \tag{15}$$

If we substitute constraint (15) into the objective function, we obtain

$$\max_{P_i} S \left[\pi^2 V_H(H) + 2\pi(1 - \pi)V_M(M) - \pi(1 - \pi)kV_M(L) + \pi(1 - \pi)kP_L + (1 - \pi)^2 P_L \right] \tag{16}$$

subject to the individual rationality constraint (11).

The maximization function is increasing in P_L, so it is optimal to set P_L as high as possible subject to the individual rationality constraint, i.e., $P_L = V_L(L)$. The choice of P_H and P_M is indeterminate, so we consider the example when $P_M = V_M(M)$. The incentive compatibility constraint thus implies $P_H = V_H(H) - \frac{1-\pi}{\pi} k [V_M(L) - V_L(L)]$.

Step 3. Since we derived the optimal mechanism ignoring constraint (13), we have to check that this constraint is not violated at the optimum. First, note that if we substitute in (13) the optimal quantities q and q_u and issue price P_I, the constraint can be rewritten as:

$$k [V_M(L) - V_L(L)] \leq \frac{\pi}{1 - \pi} [V_H(H) - V_M(H)] + 2 [V_M(M) - V_L(M)] \tag{17}$$

where all terms in brackets are non-negative. From the assumption of equidistance $k < \frac{\pi}{1-\pi} = 1$ it is enough to show that $[V_H(H) - V_M(H)] \geq [V_M(L) - V_L(L)]$ for the constraint to be satisfied.

To check whether this constraint is satisfied, we have to substitute the values of the functions $V_{s_B}(\cdot)$. Depending on the value of P_{GM} these functions take different values. In the next step we determine the different values of these functions for different ranges of P_{GM}. We can then proceed to check that constraint (17) is satisfied for all these areas.

Step 4. The value of $V_{s_B}(\cdot)$ depends on whether it is equal to the fundamental or the sale value. We first determine when $V_{s_B}(\cdot)$ is equal to the fundamental value.

For any s_B and $\tilde{s}_B$, $V_{s_B}(\tilde{s}_B) = \alpha s_G + (1-\alpha)s_B$ (the fundamental value) if and only if

$$P_{GM} \leq (1-\hat{\alpha})M + \frac{\hat{\alpha}}{\hat{\alpha}-\alpha}[(1-\alpha)s_B - (1-\hat{\alpha})\tilde{s}_B + \lambda S] \equiv B(s_B, \tilde{s}_B) \qquad (18)$$

Proof: Since $P_{GM} = E_G(v \mid s_G) = \hat{\alpha}s_G + (1-\hat{\alpha})E(s_B)$ we can derive that $s_G = \frac{1}{\hat{\alpha}}[P_{GM} - (1-\hat{\alpha})M]$. From the definition of $V_{s_B}(\tilde{s}_B)$ in (8), $V_{s_B}(\tilde{s}_B) = \alpha s_G + (1-\alpha)s_B$ if and only if

$$\alpha s_G + (1-\alpha)s_B \geq \hat{\alpha}s_G + (1-\hat{\alpha})\tilde{s}_B - \lambda S$$

Substituting for s_G and rearranging the terms gives (18).

Step 5. Let us distinguish three different areas:

1) $P_{GM} \leq B(M, H)$.

2) $P_{GM} > B(M, L)$.

3) $B(M, H) \leq P_{GM} < B(M, L)$.

where $B(M, H) < B(M, L)$ since, from (18), $B(s_B, \tilde{s}_B)$ decreases in $\tilde{s}_B$.

In the first area, from equation (18), $V_M(L) = V_M(H) = F_M$ and $V_H(H) = F_H$, where $F_{s_B} \equiv \alpha s_G + (1-\alpha)s_B$ (i.e., the fundamental value). Thus the constraint is satisfied if $(F_M - V_L(L)) \leq (F_H - F_M)$. Since $V_L(L) \geq F_L$ it is sufficient to show that $(F_M - F_L) \leq (F_H - F_M)$. Substituting for F_{s_B} we obtain

$$(1-\alpha)(M-L) \leq (1-\alpha)(H-M)$$

Since $M - L = H - M$, it is satisfied with an equality.

In the second area, $V_M(L) = V_L(L) = \hat{P}_{GM}(L) - \lambda S$. Thus the left-hand side equals 0. Since the right-hand side is always non-negative, the constraint is satisfied.

In the third area, $V_M(L) = F_M, V_H(H) = F_H, V_L(L) = \hat{P}_{GM}(L) - \lambda S$ and $V_M(H) = \hat{P}_{GM}(H) - \lambda S$. Thus the constraint is satisfied if

$$[F_M - \hat{P}_{GM}(L) + \lambda S] \leq [F_H - \hat{P}_{GM}(H) + \lambda S].$$

Substituting for F_M, $\hat{P}_{GM}(L)$, and $\hat{P}_{GM}(H)$, and using the assumption $\hat{\alpha} \geq \frac{1+\alpha}{2}$, the constraint is satisfied.

The optimal mechanism as a function of the grey market price

The expressions in Proposition 1 are written in terms of the functions $V_{s_B}(\tilde{s}_B)$, which depend on P_{GM}. Therefore the mechanism looks different over different ranges of P_{GM}. To see how the mechanism looks for different ranges of values of P_{GM} it is enough to substitute the actual values of the functions $V_{s_B}(\tilde{s}_B)$.

Since for a given P_{GM} we can always compute both the fundamental value and the sale value using equations (4) and (5), we can redefine the fundamental value and the sale value as functions of P_{GM}. The fundamental value is defined as $F_{s_B}(P_{GM}) \equiv E(v \mid s_G, s_B)$ and the sale value is defined as $\hat{P}_{GM}(P_{GM}, \tilde{s}_B) - \lambda S \equiv \hat{P}_{GM}(s_G, \tilde{s}_B) - \lambda S$.

We distinguish five different intervals of P_{GM}. The boundaries of these intervals are the values $\bar{v} > v_1 > v_2 > v_3 > v_4 > 0$, where

- v_1 is defined so that $\hat{P}_{GM}(v_1, H) - \lambda S = F_H(v_1)$. In other words, if the grey market price is $P_{GM} = v_1$, and the bookbuilding information is H, the fundamental value exactly equals the sale value of the shares;

- v_2 is defined so that $\hat{P}_{GM}(v_2, L) - \lambda S = F_M(v_2)$;

- v_3 is defined so that $\hat{P}_{GM}(v_3, M) - \lambda S = F_M(v_3)$; and

- v_4 is defined so that $\hat{P}_{GM}(v_4, L) - \lambda S = F_L(v_4)$.

The issue price in each interval is as follows:

1. If $P_{GM} \in [v_1, \bar{v}]$ then

$$\begin{cases} \text{if } \tilde{s}_B = H, & P_I = \hat{P}_{GM}(P_{GM}, H) - \lambda S, \\ \text{if } \tilde{s}_B = M, & P_I = \hat{P}_{GM}(P_{GM}, M) - \lambda S, \\ \text{if } \tilde{s}_B = L, & P_I = \hat{P}_{GM}(P_{GM}, L) - \lambda S. \end{cases}$$

2. If $P_{GM} \in [v_2, v_1]$ then

$$\begin{cases} \text{if } \tilde{s}_B = H, & P_I = F_H(P_{GM}), \\ \text{if } \tilde{s}_B = M, & P_I = \hat{P}_{GM}(P_{GM}, M) - \lambda S, \\ \text{if } \tilde{s}_B = L, & P_I = \hat{P}_{GM}(P_{GM}, L) - \lambda S. \end{cases}$$

3. If $P_{GM} \in [v_3, v_2]$ then

$$\begin{cases} \text{if } \tilde{s}_B = H, & P_I = F_H(P_{GM}) - \frac{1-\pi}{\pi} k[F_M(P_{GM}) - \hat{P}_{GM}(P_{GM}, L) + \lambda S], \\ \text{if } \tilde{s}_B = M, & P_I = \hat{P}_{GM}(P_{GM}, M) - \lambda S, \\ \text{if } \tilde{s}_B = L, & P_I = \hat{P}_{GM}(P_{GM}, L) - \lambda S. \end{cases}$$

4. If $P_{GM} \in [v_4, v_3]$ then

$$\begin{cases} \text{if } \tilde{s}_B = H, & P_I = F_H(P_{GM}) - \frac{1-\pi}{\pi} k[F_M(P_{GM}) - \hat{P}_{GM}(P_{GM}, L) + \lambda S], \\ \text{if } \tilde{s}_B = M, & P_I = F_M(P_{GM}), \\ \text{if } \tilde{s}_B = L, & P_I = \hat{P}_{GM}(P_{GM}, L) - \lambda S. \end{cases}$$

5. If $P_{GM} \in [0, v_4]$ then

$$\begin{cases} \text{if } \tilde{s}_B = H, & P_I = F_H(P_{GM}) - \frac{1-\pi}{\pi} k[F_M(P_{GM}) - F_L(P_{GM})], \\ \text{if } \tilde{s}_B = M, & P_I = F_M(P_{GM}), \\ \text{if } \tilde{s}_B = L, & P_I = F_L(P_{GM}). \end{cases}$$

Appendix 2: Allowing bookbuilding investors to trade in the grey market

In this Appendix we show when trading by bookbuilding investors in the grey market can affect the optimal mechanism and reduce expected IPO proceeds.

We proceed as follows. We assume that the underwriter is using the mechanism of Proposition 1 and check whether the incentive compatibility constraints are still satisfied.

We solve the model backward: assuming that a bookbuilding investor has taken a position in the grey market, we look at how his incentive to truthfully report his signal changes. Then, we look at the investor's decision to buy or short-sell shares in the grey market, given that he can anticipate the signal he will report during bookbuilding.

Consider an informed investor who has observed a signal $s_i = H$. We only need to look at the incentive compatibility constraint of this investor, since we showed in Appendix 1 that only this constraint is binding at the optimum.

If an investor who has observed a signal $s_i = H$ has no position at all in the grey market, his incentive compatibility constraint would be as in (13). However, if he has a position in the grey market, the incentive compatibility constraint would change.

If the investor has bought γ shares in the grey market, the constraint becomes

$$\pi \frac{S}{2} \left[V_H(H) - P_H\right] + (1 - \pi)S \left[V_M(M) - P_M\right] + \gamma \left[\pi V_H(H) + (1 - \pi)V_M(M) - P_{GM}\right]$$
$$\geq (1 - \pi)\frac{Sk}{2} \left[V_M(L) - P_L\right] + \gamma \left[\pi V_H(M) + (1 - \pi)V_M(L) - P_{GM}\right]$$

(19)

The last term on the left-hand side represents the expected profit from buying γ shares in the grey market if he truthfully declares H. The last term on the right-hand side represents the expected profit from the grey market position if he falsely declares L.

Since the incentive compatibility constraint (13) was satisfied with equality in the mechanism of Proposition 1, constraint (19) is satisfied if and only if

$$\gamma \left[\pi V_H(H) + (1 - \pi)V_M(M) - P_{GM}\right] \geq \gamma \left[\pi V_H(M) + (1 - \pi)V_M(L) - P_{GM}\right].$$

Since $V_H(H) \geq V_H(M)$ and $V_M(M) \geq V_M(L)$ the inequality is always satisfied. Thus, an investor who observed a signal H and bought γ shares in the grey market will still tell the truth and declare H.

If instead the bookbuilding investor has a short position of γ shares in the grey market, his incentive compatibility constraint is satisfied if and only if

$$\gamma \left[P_{GM} - \pi V_H(H) - (1-\pi)V_M(M) \right] \geq \gamma \left[P_{GM} - \pi V_H(M) - (1-\pi)V_M(L) \right],$$

where the left-hand side is the expected profit from the short position if he truthfully announces his signal H, and the right-hand side is the profit if he misreports his signal as L. In both cases, the profits are given by the difference between the grey market price and the expected value to the investor of the shares he must deliver to cover the position. Since $V_H(H) \geq V_H(M)$ and $V_M(M) \geq V_M(L)$ this inequality is always weakly violated. If $P_{GM} \leq F_M + \lambda S$ (corresponding to $P_{GM} \leq v_3$, where v_3 is defined in Section 2.4), it is satisfied with equality, so the investor does not lie. However, if $P_{GM} > F_M + \lambda S$ the inequality is violated and he misreports his signal.

To summarize, when a bookbuilding investor with signal H is long in the grey market, he will have greater incentive to truthfully reveal his signal in order to raise the aftermarket price. If he is short in the grey market and the grey market price is sufficiently high, he will have an incentive to report a signal L in order to manipulate the price downward.

We now look at whether this bookbuilding investor prefers to take a long position in the grey market (and truthfully reveal his signal) or whether he prefers to take a short position (and misreport his signal when $P_{GM} > F_M + \lambda S$). Since the original constraint (13) is satisfied with equality, we only have to compare the additional profits from the respective grey market positions. The investor will prefer to buy shares in the grey market (and report truthfully) if and only if

$$\gamma \left[\pi V_H(H) + (1-\pi)V_M(M) - P_{GM} \right] \geq \gamma \left[P_{GM} - \pi V_H(M) - (1-\pi)V_M(L) \right] \qquad (20)$$

In order to see when this inequality is satisfied, start by considering the special case of $\lambda = 0$. By substituting for the functions $V_{s_B}(\tilde{s}_B)$ one can verify that the inequality is always satisfied. Thus, an investor who has observed a signal H will buy shares in the grey market and declare the truth.

However, when $\lambda > 0$ the inequality may be violated. In particular, when P_{GM} is very high (i.e. $P_{GM} \geq F_H + \lambda S$), any $\lambda > 0$ results in a violation of the inequality. For intermediate

values of P_{GM} ($F_M + \lambda S < P_{GM} < F_H + \lambda S$), the inequality will be violated for sufficiently large λ. For lower levels of P_{GM} ($P_{GM} < F_M + \lambda S$) the inequality is irrelevant since, as shown above, even with a short position the investor has no incentive to misreport his signal.

Thus, when P_{GM} is high and $\lambda > 0$, allowing bookbuilding investors to trade in the grey market results in lower IPO proceeds.

References

Aggarwal, R., L. Krigman, and K. Womack, 2002, "Strategic IPO underpricing, information momentum, and lockup expiration selling," *Journal of Financial Economics* 66, 105-137.

Amemiya, T., 1973, "Regression analysis when the dependent variable is truncated Normal," *Econometrica* 4, 997-1016.

Aussenegg, W., P. Pichler, and A. Stomper, 2003, "IPO pricing with bookbuilding and a when-issued market," *mimeo*, University of Vienna.

Baker, M., and J. Wurgler, 2003, "Investor sentiment and the cross-section of stock returns," *mimeo*, Harvard University.

Barberis, N., A. Shleifer, and R. Vishny, 1998, "A model of investor sentiment," *Journal of Financial Economics* 49, 307-343.

Ben Dor, A., 2003, "The performance of initial public offerings and the cross section of institutional ownership," *mimeo*, Northwestern University.

Bennouri, M., and S. Falconieri, 2003, "Price versus quantity discrimination in optimal IPOs," *mimeo*, Tilburg University.

Benveniste, L.M., and P.A. Spindt, 1989, "How investment bankers determine the offer price and allocation of new issues," *Journal of Financial Economics* 24, 343-61.

Benveniste, L.M., and W.J. Wilhelm, 1990, "A comparative analysis of IPO proceeds under alternative regulatory environments," *Journal of Financial Economics* 28, 173-207.

Biais, B., and A.M. Faugeron-Crouzet, 2002, "IPO auctions: English, Dutch, ...French, and Internet," *Journal of Financial Intermediation* 11, 9-36.

Boehmer, E., and A. Ljungqvist, 2004, "On the decision to go public: Evidence from privately-held firms", *mimeo*, New York University.

Busaba, W., and C. Chang, 2002, "Bookbuilding vs. fixed-price revisited: The effect of aftermarket trading," *mimeo*, University of Western Ontario.

Cornelli, F., and D. Goldreich, 2001, "Bookbuilding and strategic allocation," *Journal of Finance* 56, 2337-70.

Cornelli, F., and D. Goldreich, 2003, "Bookbuilding: How informative is the order book?" *Journal of Finance* 58, 1415-44.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam, 1998, "Investor psychology and security market under- and over-reactions," *Journal of Finance* 53, 1839-1886

Derrien, F., 2004, "IPO pricing in 'hot' market conditions: Who leaves money on the table?" *Journal of Finance*, forthcoming.

Dorn, D., 2003, "Does sentiment drive the retail demand for IPOs?" *mimeo*, Columbia University.

Greene, W.H., 2003, *Econometric Analysis*, Prentice Hall.

Hanley, Kathleen W., 1993, "The underpricing of initial public offerings and the partial adjustment phenomenon," *Journal of Financial Economics* 34, 213-250.

Harris, M., and A. Raviv, 1993, "Differences of opinion make a horse race," *Review of Financial Studies* 6, 473-506.

Heckman, J., 1979, "Sample selection bias as a specification error," Econometrica 47, 153-161.

Jenkinson, T.J., A. Morrison, and W.J. Wilhelm, 2003, "Why are European IPOs so rarely priced outside the indicative price range?" *mimeo*, Oxford University.

Krigman, L., W.H. Shaw, and K.L. Womack, 1999, "The persistence of IPO mispricing and the predictive power of flipping," *Journal of Finance* 54, 1015-1044.

Lee, C., A. Shleifer, and R. Thaler, 1991, "Investor sentiment and the closed-end fund puzzle," *Journal of Finance* 46, 75-109.

Ljungqvist, A., T.J. Jenkinson, and W.J. Wilhelm, 2003, "Global integration in primary equity markets: The role of U.S. banks and U.S. investors," *Review of Financial Studies* 16, 63-99.

Ljungqvist, A., and W.J. Wilhelm, 2002, "IPO allocations: Discriminatory or discretionary?" *Journal of Financial Economics* 65, 167-201.

Ljungqvist, A., V. Nanda, and R. Singh, 2003, "Hot markets, investor sentiment, and IPO pricing," *mimeo*, New York University.

Löffler, G., P. Panther, and E. Theissen, 2002, "Who knows what when? The information content of pre-IPO market prices," *mimeo*, University of Bonn.

Miller, E.M., 1977, "Risk, uncertainty, and divergence of opinion," *Journal of Finance* 32, 1151-1168.

Neal, R., and S. Wheatley, 1998, "Do measures of investor sentiment predict stock returns?" *Journal of Financial and Quantitative Analysis* 34, 523-547.

Ofek, E., and M. Richardson, 2003, "Dot-com mania: The rise and fall of internet stock prices," *Journal of Finance* 58, 1113-1138.

Pichler, P., and A. Stomper, 2004, "Primary market design: Direct mechanism and markets," *mimeo*, University of Vienna.

Purnanandam, A., and B. Swaminathan, 2003, "Are IPOs really underpriced?" *Review of Financial Studies*, forthcoming.

Rajan, R., and H. Servaes, 2003, "The effect of market conditions on initial public offerings," in J. McCahery and L. Renneboog, eds., *Venture Capital Contracting and the Valuation of High-tech Firms*, Oxford University Press.

Ritter, J., 1991, "The long-run performance of initial public offerings," *Journal of Finance* 46, 3-27.

Ritter, J., and I. Welch, 2002, "A review of IPO activity, pricing, and allocation," *Journal of Finance* 57, 1795-1828.

Shleifer, A., 2000, *Inefficient Markets: An Introduction to Behavioral Finance*, Oxford University Press.
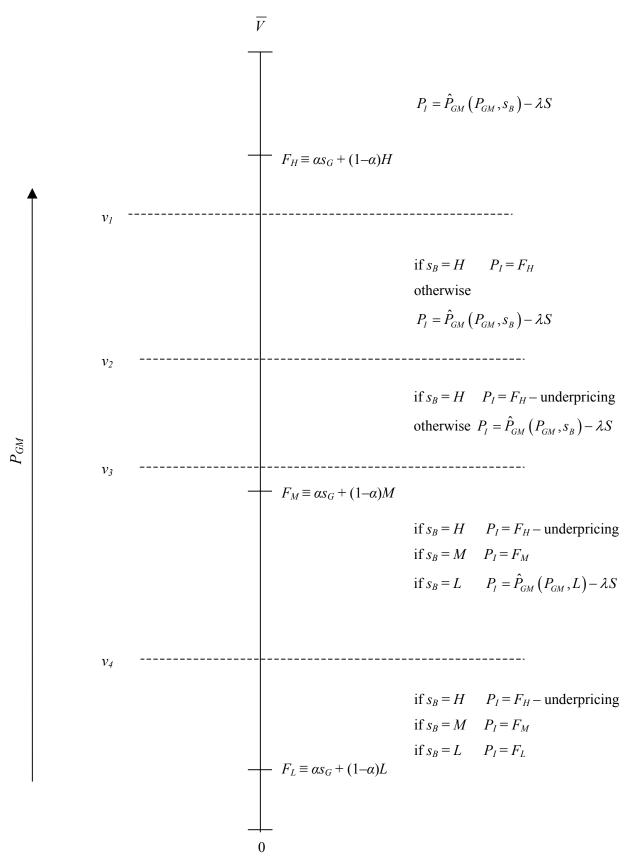
$$\overline{V}$$

$$P_I = \hat{P}_{GM}\left(P_{GM}, s_B\right) - \lambda S$$

$$F_H \equiv \alpha s_G + (1-\alpha)H$$

v_1 ---

if $s_B = H$ $P_I = F_H$

otherwise

$$P_I = \hat{P}_{GM}\left(P_{GM}, s_B\right) - \lambda S$$

v_2 ---

if $s_B = H$ $P_I = F_H -$ underpricing

otherwise $P_I = \hat{P}_{GM}\left(P_{GM}, s_B\right) - \lambda S$

v_3 ---

$$F_M \equiv \alpha s_G + (1-\alpha)M$$

if $s_B = H$ $P_I = F_H -$ underpricing

if $s_B = M$ $P_I = F_M$

if $s_B = L$ $P_I = \hat{P}_{GM}\left(P_{GM}, L\right) - \lambda S$

v_4 ---

if $s_B = H$ $P_I = F_H -$ underpricing

if $s_B = M$ $P_I = F_M$

if $s_B = L$ $P_I = F_L$

$$F_L \equiv \alpha s_G + (1-\alpha)L$$

0

P_{GM}

Figure 1. The issue price in the optimal mechanism. The figure illustrates the issue price (for $s_B = H, M, L$), as derived in Appendix 1, over the possible ranges of the grey market price. The boundaries v_2, v_2, v_3, v_4 are defined in Appendix 1. F_{sB} is the fundamental value when the bookbuilding signal is s_B (given P_{GM}). The figure is drawn for small positive values of λS. (When $\lambda = 0$, $F_M = v_3$.)

Table 1: Descriptive Statistics of IPOs with Available Grey Market Prices

We have grey market prices for 486 (mostly European) IPOs completed between November 1995 and December 2002. Sample companies are incorporated in the following 20 countries: Austria (13), Belgium (1), Canada (1), Denmark (1), Finland (3), France (13), Germany (321), Greece (2), Ireland (2), Israel (7), Italy (61), Lithuania (1), Luxembourg (1), Netherlands (11), Norway (2), Spain (5), Sweden (2), Switzerland (11), the United Kingdom (24), and United States (4). Note that there is no grey market in the U.S.; the four American companies are in the sample because they go public in Europe. Most companies go public in their home country, but some do not. Where a company goes public on more than one exchange, we take the listing country to be its home country or (if it doesn't list on a home-country exchange) the country in which most of the shares are placed. The table shows descriptive statistics for the sample as a whole as well as broken down by the twelve countries on whose exchanges sample companies list. We also show, for each listing country, the first and last date for which we have an IPO with grey market prices. This sample window varies from country to country. The sample for which we have grey market prices is a subsample of the 2,723 IPOs completed in the twelve listing countries shown between November 1995 and December 2002. Gross proceeds are shares sold (including the overallotment option if exercised) times the issue price, converted into U.S. dollars using exchange rates on the first trading day. Initial returns are computed using the closing price on the first trading day. Quoted half spread refers to the quoted bid-ask spread in the grey market, just before the IPO issue price is set. It is computed as half the difference between the bid and the ask divided by the midpoint of the spread.

	Sample window		No. of IPOs…			Gross proceeds ($m)		Initial return (%)			Fraction priced at high end of range	Grey market price rel. to midpoint of price range, mean (%)
	first date	last date	w/ grey market prices	during sample window	in Nov '95-Dec 2002	mean	median	mean	st.dev.	Quoted half spread (%). mean		
Total	Nov 1995	Dec 2002	486	1,755	2,723	343.7	53.0	36.3	65.6	4.7	54.1	40.4
By country of listing												
Austria	Nov 1997	Nov 2000	2	18	25	654.5	654.5	-2.5	6.6	6.9	0.0	5.1
Finland	Nov 1998	Dec 1999	3	22	58	686.7	531.0	74.6	91.9	1.6	33.3	53.0
France	Oct 1997	Dec 2001	14	409	544	1715.3	650.0	6.5	12.9	4.2	42.9	24.6
Germany	Nov 1995	Jul 2002	363	489	504	169.9	42.2	41.5	67.8	5.1	63.4	46.6
Greece	Oct 2000	Dec 2001	2	23	180	423.4	423.4	-4.6	0.6		0.0	-7.3
Italy	Nov 1995	Dec 2002	61	132	133	599.0	106.1	20.2	63.4	4.3	27.9	13.1
Netherlands	Mar 2000	Mar 2000	1	1	56	2829.0	2829.0	0.5			100.0	117.9
Norway	Mar 2000	May 2000	2	5	107	139.7	139.7	29.5	44.5	0.6	0.0	85.7
Spain	Jun 1999	May 2001	5	12	38	1374.0	915.7	10.5	13.2	5.1	20.0	11.8
Sweden	Jun 2000	Jun 2001	2	22	196	4405.4	4405.4	9.9	8.2	0.9	50.0	19.2
Switzerland	Dec 1996	Dec 2001	8	59	67	1097.4	153.8	50.1	99.8	1.7	12.5	36.0
United Kingdom	Jun 1997	Jul 2002	23	563	815	566.8	265.3	21.5	35.0	1.5	21.7	32.9

Table 2: Determinants of the Issue Price

The dependent variable in these regressions is the IPO issue price P_I normalized by the midpoint of the initial price range P_{mid}. The explanatory variables are the last grey market price before the issue price was set P_{GM} (also normalized by the midpoint of the initial price range), the last bid-ask spread in the grey market (divided by its midpoint), and the logarithm of the IPO proceeds. We also include the domestic market index return over the three-month period before the IPO as a control variable. To capture the predicted asymmetry, we define an indicator function set to one when P_{GM} is above P_{mid}. Grey market prices are available for 486 IPOs. Nine of these are fixed-price offerings, so we lack information on their initial price ranges. This reduces the number of observations in model (1) to 477. Models (2) through (5) include the bid-ask spread, which is available for 442 IPOs. We use censored regressions because European IPOs are rarely priced outside the initial price range. t-statistics are reported in parentheses. Three, two, and one asterisks indicate significance at the 1%, 5%, and 10% level, respectively. Intercepts are not shown.

	Dependent variable: Normalized Issue Price (P_I / P_{mid})				
	(1)	(2)	(3)	(4)	(5)
P_{GM} / P_{mid}	0.29***	0.31***	0.31***	0.31***	0.37***
	(6.01)	(5.60)	(5.70)	(5.49)	(6.34)
P_{GM} / P_{mid} x Indicator($P_{GM} > P_{mid}$)	0.15***	0.14***	0.14***	0.16***	0.13***
	(7.29)	(6.43)	(6.44)	(5.50)	(6.24)
Market index return	0.20**	0.25***	0.26***	0.26***	0.25***
	(2.39)	(2.72)	(2.87)	(2.88)	(2.72)
Grey market bid-ask spread		-0.36***			
		(-3.24)			
Grey market bid-ask spread x P_{GM} / P_{mid}					-0.37***
					(-3.57)
Log gross proceeds		-0.01*			-0.01*
		(-1.83)			(-1.91)
Grey market bid-ask spread x Log gross proceeds			-0.08***	-0.05	
			(-3.13)	(-1.23)	
Grey market bid-ask spread x Log gross proceeds x Indicator($P_{GM} > P_{mid}$)				-0.06	
				(-1.15)	
LR test: all coeff. = 0 (χ^2)	488.9***	454.3***	453.2***	454.5***	456.6***
No. of observations	477	442	442	442	442
No. of left-censored observations	51	50	50	50	50
No. of right-censored observations	263	246	246	246	246

Table 3: Determinants of the First-Day Aftermarket Price

The dependent variable in these regressions is the stock price at the end of the first day of aftermarket trading (normalized by the midpoint of the range) adjusted for the market index return from the pricing date to the end of the first day of aftermarket trading: $P_{AM} / P_{mid} - (1 + \text{market index return})$. The explanatory variables are the normalized last grey market price before the issue price was set P_{GM} / P_{mid}, the normalized issue price P_I / P_{mid}, the last bid-ask spread in the grey market (divided by its midpoint), and the logarithm of the IPO proceeds. We also include the domestic market index return over the three-month period before the IPO as a control variable. White heteroskedasticity consistent t-statistics are given in parentheses. Three, two, and one asterisks indicate significance at the 1%, 5%, and 10% level, respectively. Intercepts are not shown.

	Full sample			$P_{GM} > P_{mid}$	$P_{GM} \leq P_{mid}$
	(1)	(2)	(3)	(4)	(5)
P_{GM} / P_{mid}	0.98***		0.95***	0.95***	0.56***
	(14.87)		(13.14)	(12.59)	(3.62)
P_I / P_{mid}		2.60***	0.44**	0.51*	0.53***
		(11.50)	(2.46)	(1.66)	(3.12)
Market index return	0.05	2.12***	-0.01	0.03	-0.15*
	(0.23)	(5.10)	(-0.07)	(0.11)	(-1.78)
Grey market bid-ask spread	-0.62	0.41	-0.46	-0.73	-0.06
	(-1.52)	(0.80)	(-1.09)	(-1.11)	(-0.49)
Log gross proceeds	-0.04***	-0.05***	-0.04***	-0.05***	-0.01
	(-3.06)	(-3.00)	(-3.12)	(-3.02)	(-1.50)
Adjusted R^2	75.4 %	27.5 %	75.7 %	70.7 %	68.4 %
F-test: all coeff. = 0	77.6***	52.7***	164.5***	64.4***	55.2***
No. of observations	442	442	442	330	112

Table 4: Aftermarket Volume

The dependent variable in these regressions is the natural logarithm of first-day volume (as a percentage of the shares sold in the IPO), measured over the first day and first week of aftermarket trading. The main explanatory variable is an indicator function set to one when the last grey market price before the issue price was set (P_{GM}) exceeded the midpoint of the initial price range (P_{mid}). The controls in models (2)-(3) and (5)-(6) are the domestic market index return over the three-month period before the IPO and the normalized first-day after-market price (P_{AM} / P_{mid}). White heteroskedasticity consistent t-statistics are given in parentheses. Three and two asterisks indicate significance at the 1% and 5% level, respectively. Intercepts are not shown.

	Log first-day volume			Log first-week volume		
	(1)	(2)	(3)	(4)	(5)	(6)
Indicator(P_{GM} > P_{mid})	1.08***	0.89***	0.70***	1.00***	0.84***	0.65***
	(7.64)	(6.15)	(4.59)	(8.72)	(7.25)	(5.25)
Market returns		2.84***	2.12*		2.47***	1.74**
		(4.48)	(3.20)		(4.69)	(3.19)
P_{AM} / P_{mid}			0.33***			0.33***
			(4.70)			(5.34)
Adjusted R^2	12.0 %	15.7%	18.0%	14.2 %	18.0%	21.3%
F-test: all coeff. = 0	58.3***	40.7***	44.3***	76.1***	48.7***	50.9***
No. of observations	443	443	443	443	443	443

Table 5. Long-Run Returns

The dependent variables are market-adjusted long-run returns measured from the first day of aftermarket trading, defined as $(R_{LR} - R_{mkt})(P_{AM}/P_{mid})$ where R_{LR} is the long-run return over the first two, three, six or 12 months of aftermarket trade, R_{mkt} is the contemporaneous return on the domestic market index, and the multiplier (P_{AM}/P_{mid}) is used to ensure that the dependent variables are consistent with the normalization of the independent variables. White heteroskedasticity consistent t-statistics are given in parentheses. Three and two asterisks indicate significance at the 1% and 5% level, respectively. Intercepts are not shown.

	Full sample				$P_{GM} > P_{mid}$				$P_{GM} \leq P_{mid}$			
Horizon:	42 days	63 days	126 days	252 days	42 days	63 days	126 days	252 days	42 days	63 days	126 days	252 days
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$(P_{GM} - P_{mid})/P_{mid}$	-0.20***	-0.19**	-0.39***	-0.65***	-0.26***	-0.25**	-0.48***	-0.77***	0.42	1.00	1.86**	0.98**
	(-3.43)	(-2.12)	(-3.51)	(7.38)	(-3.68)	(-2.36)	(-3.67)	(-7.05)	(1.48)	(1.97)	(2.14)	(1.98)
$(P_{AM} - P_{GM})/P_{mid}$	0.12	0.66**	0.67	-0.01	0.09	0.63**	0.64	-0.03	1.23	2.20	1.33	-0.80
	(1.01)	(2.32)	(1.37)	(-0.02)	(0.76)	(2.12)	(1.26)	(-0.10)	(1.29)	(1.60)	(0.91)	(-1.00)
Adjusted R^2	2.5%	6.0%	5.2%	4.8	3.2%	6.6%	6.3%	5.5%	5.9 %	8.1 %	2.4 %	0.7 %
F-test: all coeff. = 0	8.5***	10.8***	18.8***	43.7***	9.1***	11.6***	20.6***	39.0***	1.2	2.0	2.3	3.0
No. of observations	477	477	477	477	358	358	358	358	119	119	119	119