The Pricing of IPO Services and Issues: Theory and Estimation

Ari Kang
The University of Texas-Austin Department of Finance

Richard Lowery
The University of Texas-Austin Department of Finance

We estimate a model for the process for setting IPO spreads and offer prices. We establish that the partially rigid spread schedule observed for IPOs, where over 90% of IPOs with proceeds between $20 and $80 million have a spread of 7%, can be rationalized as optimal collusion. Optimal collusion creates a rationale for high underpricing, and we use data on both spreads and underpricing to estimate structural parameters. Our estimates suggest that firms benefit from holding IPOs but that idiosyncratic manager preferences may drive much of the IPO market. Much of the money left on the table is estimated to accrue to underwriters. (JEL G240, L130, D430)

Managers of firms willingly pay high spreads and tolerate significant underpricing in order to take their firms public in underwritten initial public offerings (IPOs). Thus, the IPO process must be highly valuable to the decision makers within firms, either because the IPO increases the value of the firm or because of private benefits to management. Quantifying these benefits is a challenge, in part because the market for IPO services is characterized by somewhat puzzling behavior. Spreads, the percentage fee charged by underwriters for their services, are heavily concentrated at a single number, suggesting strategic behavior by underwriters while competing for IPOs; additionally, issues are systematically priced below the market value of the issue, indicating that the underwriter chooses to raise less money for the firm than it could, which also lowers the total fees collected by the underwriter. In this paper, we seek to match patterns in the IPO data with an empirically plausible model and to use

We thank the editor, Christopher Hennessy, and two anonymous referees for suggestions that led to substantial improvements in the paper. We also thank Paolo Fulghieri, William Wilhelm, and Zhaohui Chen for useful comments and suggestions. Supplementary data can be found on The Review of Corporate Finance Studies web site. Send correspondence to Richard Lowery, The University of Texas-Austin, McCombs School of Business, Department of Finance, 2100 Speedway, Austin, TX, 78712, USA; telephone: 512-232-6823. E-mail: richard.lowery@mccombs.utexas.edu.

© The Author 2014. Published by Oxford University Press on behalf of The Society for Financial Studies. All rights reserved. For Permissions, please e-mail: journals.permissions@oup.com.
doi:10.1093/rcfs/cfu005 Advance Access publication February 27, 2014
this model to uncover structural parameters governing the IPO process. With the model, we can present a possible decomposition of the benefits of an IPO, evaluate implications for the division of returns from underpricing between underwriters and investors who receive IPO shares, and investigate the likely effects of policies intended to decrease spreads or limit underpricing.

We model the selection of an IPO underwriter as an auction, where underwriters bid for the opportunity to take a firm public. We first show theoretically that the observed pattern of spreads, with heavy concentration at a single spread and a secondary concentration at a higher spread, can naturally arise under optimal collusion in an environment in which underwriters have only noisy, private information about the value of private firms. Empirically, spreads heavily concentrate on 7%, as previously documented by Chen and Ritter (2000). The second most common spread by a significant margin is 10%. Our theory, which relates to the price rigidity results of Athey et al. (2004) and Hanazono and Yang (2007),\(^1\) argues that underwriters will have valuable private information about the true value of a firm holding an IPO, but because of the need to avoid costly punishments arising from accidental apparent deviations from the collusive equilibrium underwriters largely ignore this information when bidding for the IPO. The coarse use of information implied by optimal collusion will then, in many cases, imply a relationship between IPO proceeds and spreads that resembles the step function observed in the data. In contrast, neither competition nor perfect collusion that would imitate the behavior of a monopolist underwriter are likely to explain the patterns observed in spread data. This result allows us to posit a data-generating process for the spread data, the first step toward estimating the parameters of the model.

We also derive equilibrium underpricing under a collusive equilibrium, which provides a crucial secondary source of identification. Underpricing plays a role in the model because underwriters neglect private information they have about the value of the firm when setting the spread. Because this information is valuable for determining how much a firm is willing to pay to go public, underwriters will seek to exploit this information by setting an issue price below the market value of the IPO as long as some share of the profits from the initial price run up are returned, directly or indirectly, to the underwriter. This secondary source of identification is needed because equilibrium bid behavior is degenerate. Structural models of auctions in general exploit variation in bidding

---

1 These papers deal with, respectively, a case in which each market participant has private information over a private value component of the payoff and a case with only two possible values for the imperfectly observed variable. Thus, both models are inadequate for investigating the full distribution of spreads in the IPO market.
behavior as a function of observable data to identify model parameters. This variation is not available to us, but our model does have implications for the relationship between IPO proceeds and underpricing, which can be used to identify parameters of interest.

By identifying the most plausible data-generating process for spreads and showing the implications of this process for underpricing data, we are able to in effect invert spread and underpricing data to recover structural parameters on both the value of the IPO process to firms and managers and the way in which profits from underpricing are divided between investors and underwriters. Our analysis suggests both that firms benefit from the IPO process, meaning that the firm should observe an increase in its value proportionate to the size of the IPO, and that at the same time managerial preference, represented as a fixed benefit to the decision maker that does not affect the proceeds of the IPO, also plays an important role in the decision to take a firm public.

Our results also indicate that much of the value from underpricing ultimately accrues to underwriters through the various channels identified in the academic literature, popular press, and regulatory enforcement actions, such as spinning (Liu and Ritter 2010), laddering Griffin et al. (2007), direct kickbacks, and other means. Although our estimates of the benefit to underwriters from underpricing are high, they are not out of line with anecdotal reports for smaller firms (see Liu and Ritter 2011), and, for the most part, they are consistent with the idea that underwriters and investors split the profits from underpricing. That is, our estimates do not show a benefit from underpricing that exceeds the direct profits available, a restriction that is not imposed ex ante. This result hinges in an important way on the fact that we account for optimal collusion on spreads. If underwriters were assumed to behave as monopolists, the high level of underpricing observed would suggest a greater preference for underpricing than for direct payments via spreads (i.e., a dollar of money left on the table would appear more valuable than an actual dollar paid to the underwriter), an unreasonable result that would then support the claim that underpricing arises from the preferences of issuing firms. We also find that the direct cost to the underwriter of holding an IPO should be small; otherwise, underwriters would show a greater preference for profiting directly via spreads versus indirectly via underpricing.

Beyond this, our estimates can be used to consider the potential drivers of the extremely high levels of underpricing during the dot-com boom in the 1990s. The increase in IPO underpricing during the dot-com boom appears to be driven by an increase in the average preference for IPOs rather than from changes in the productivity of IPOs. Then, having estimated the model, we can analyze the effect of different policy interventions in the IPO market. Policy intervention is potentially beneficial because, as we show, collusion best explains the pattern of spreads.
observed in data, and collusion will lead to inefficient underprovision of IPO services. Perhaps surprisingly, our empirical results indicate that directly intervening to reduce spreads would likely destroy value, whereas intervening to reduce the incentives available to underwriters to underprice issues would create value. The failure of a policy of capping spreads to create value comes as a result of our high estimate of the direct value underwriters place on underpricing. High spreads can dissuade firms that would value an IPO from going public, but high spreads also partially align the incentives of underwriters and firms, giving the underwriter an incentive to maximize proceeds from the issue.

Because we show how to combine the data on spreads with the data on underpricing to estimate the value of IPOs to firms, our work complements the structural estimation of French IPO data in Biais et al. (2002) (henceforth, BBR). That study assumes that the IPO mechanism is designed to maximize proceeds to the firm, whereas we solve for the optimal equilibrium of a fixed IPO mechanism from the perspective of the underwriters. In this way, we model the collusion among underwriters, which is implicit in BBR, but do not address the design of the institutional features of the IPO process. Our approaches are not contradictory; the optimal mechanism in BBR is not the mechanism employed in the United States. Because of this, their study considers data from French IPOs that use a particular market design that implements the optimal mechanism.

There are two other important distinctions between our work and this earlier structural estimation. First, our interest is in recovering parameters that describe the value of the IPO process itself, whereas the focus in BBR is on the sale of a quantity of shares with a fixed, but imperfectly observed, value in the presence of adverse selection and conflicts of interest between firms and underwriters. A second distinction is how we treat the conflict of interest between firms going public and underwriters. BBR assume an extreme degree of conflict, whereby the interest of the underwriter is perfectly aligned with professional investors rather than with the firm. In contrast, we remain agnostic on the severity of the conflict of interest and estimate parameters that determine the extent to which underwriters value the profits of the initial investors in the IPO. Interestingly, our estimates of the extent of this conflict support the assumption in BBR as our estimates indicate that a large fraction of the profits from underpricing are effectively returned to the underwriter.

In addition to providing a means to estimate the parameters of a model of IPOs, our analysis can inform the debate about the forces leading to underpricing in the IPO market. Ultimately, the question of whether underpricing is a mechanism for underwriters to extract rents from firms (as in Loughran and Ritter 2004; Cliff and Denis 2004) or a service provided by underwriters to firms (as in Rock 1986; Benveniste and Spindt 1989; Allen and Faulhaber 1989; Chemmanur 1993, among
others) must be decided at least in part on which model of underpricing can explain the empirical patterns observed. At first, the idea that underwriters use underpricing to extract rents from firms may appear implausible because this is a very inefficient means to charge firms for IPO services; underwriters could just charge higher spreads.

From the theoretical perspective, we show why underwriters might use the inefficient means to obtain compensation, and we are able to fit a model with plausible parameter values to the data. Out of the many explanations for underpricing, few have been taken to the data in a way that can be interpreted as providing a fit to the model, with the obvious exception being BBR in the context of French underpricing data.

We can also consider the implications of our model for those IPOs that do not conform to the dominant concentration on 10% and 7%. Virtually no IPOs have spreads above 10%, whereas for IPOs above $100 million in proceeds (which represent about 14% of IPOs in the sample period), spreads tend to be below 7% and are on average decreasing in proceeds. Finally, a subset of primarily smaller IPOs, which tend to be younger firms with low market capitalization, receives spreads between 7% and 10% (16%).

Although our estimation focuses only on the 10%/7% spreads, we also note that the model could be extended to rationalize the coexistence of the 10%/7% spread schedule for most firms with exceptions for the largest firms, where the profit from the IPO makes deviating too tempting to follow through on optimal collusion, and for some of the smaller firms when competition for the IPO is expected to be minimal. Firms receiving spreads between 7% and 10%, which appear to be deviations from our model, could be those in which underwriters anticipate little competition and thus bid according to a monopolist, rather than collusive equilibrium. We find some empirical support for this claim.

1. Stage Game

We model competition for providing IPO services as an infinitely repeated game among two long-lived underwriters and a new potential IPO firm arriving in each period. In this section, we describe the stage game. The firm has the opportunity to go public, selling an exogenously determined fraction of its equity, by hiring one of the two underwriters. The motivation of the firm, or more precisely the manager of the firm, for going public is twofold.

First, going public changes the value of the firm to the owner (or to the shareholders) of the firm, proportionate to the value of the equity sold at issuance, which we denote \( x \). Further, let \( w \) be the value of the share of the firm not to be sold at the issuance, which will not be affected by the
IPO process. Let $\beta$ represent the proportionate effect of the IPO on firm value. If $\beta > 1$, an IPO increases the value of the firm, perhaps through improved access to capital markets and the associated lower cost of capital for undertaking projects. Similarly, $\beta$ can be thought of as a reduced-form approach for modeling the incentives to go public, as outlined in Chemmanur and Fulghieri (1999). Alternatively, $\beta$ could be less than 1. This would represent a decrease in value resulting from less-effective monitoring by owners or greater regulatory costs. The manager has a claim or an intrinsic concern for some fraction $\phi$ of the value of the firm.

Second, the manager receives some private benefit or pays some private costs $\varepsilon$ when the firm goes public, where $\varepsilon$ is random. We will primarily focus on the case in which $\varepsilon > 0$, such that managers prefer to take the firm public, but this is not crucial for the results. The distribution of $\varepsilon$ is independent of the value of the firm and is privately observed by the firm. The private benefit $\varepsilon$ enters additively into the payoff function of the manager. Because the effect on the value of the firm is proportionate to the size of the IPO, this implies that the private benefit is relatively more important for smaller firms compared with larger firms. This structure is consistent with empirical findings on the scope for moral hazard in small versus large firms; Himmelberg et al. (1999) show that smaller firms tend to have larger managerial stakes, indicating that concerns about managerial diversion may be more severe in small firms. Slovin et al. (1992) also find evidence that smaller firms may be more subject to moral hazard concerns based on the stock price response to bank loan renewals.

The process of holding an IPO proceeds as follows. After the firm learns its value as a private entity, the proportion of the firm to be sold, and the preference of the manager for going public, each underwriter receives some information about the firm. Specifically, each underwriter $i$ privately observes a signal $\xi_i, i \in \{1,2\}$, which is informative about the value $x$, the value of the equity to be sold in the IPO. Underwriters then simultaneously submit bids in terms of a spread $\alpha \in [0,\infty)$, which represents the fraction of the realized value of the IPO proceeds that will accrue to the underwriter in fees. The firm chooses to either go public using one of the two underwriters or remain private. Because the underwriters are identical from the perspective of the firm, the firm will always choose the lowest spread if it decides to go public. Whichever underwriter is chosen incurs a fixed cost $\kappa$ to manage the IPO. For simplicity, the firm is assumed to randomize uniformly if both underwriters submit the same

---

2 Because the fraction of the firm to be sold publicly is assumed to be publicly observable, this is equivalent to assuming that the underwriter receives a signal about firm value.

3 This spread is allowed to exceed one for technical convenience, but this will have no effect in any relevant cases.
spread. This process is a procurement auction with, effectively, a security bid.\(^4\) We take the form of the auction as given and do not allow underwriters to optimize over all possible bidding mechanisms.\(^5\)

Following the competition for the opportunity to hold the IPO, the winning underwriter pays a fixed cost \(\kappa\) to run the IPO and learns the true value of the equity to be sold, \(x\), through the bookbuilding process. The underwriter then selects an offer price by choosing an amount of money to leave on the table for initial investors, \(u\). Some potentially non-linear share of this money left on the table is kicked back to the underwriter, which is given by the function \(\theta(\alpha, u, x)\). This results in a level of proceeds \(v\), which is a function of \(x\), \(u\), and \(\alpha\). Initially, we assume that the underwriter receives no benefits from underpricing the issue and simply selects the offer price that corresponds to the value of the proceeds, which is given by \(\beta x\).\(^6\) Later, we admit the possibility that the kickback from investors is positive, which gives the underwriter an incentive to underprice issues. In this variant of the model, the firm then has the opportunity to cancel the IPO. Note that the underwriter cannot commit to an offer price when bidding for the right to take a firm public and thus cannot compete on offer prices. In practice, the offer price is only set the day before the issue starts trading, months after the firm selects the underwriter.

Finally, the true proceeds and the lowest spread bid are revealed publicly.\(^7\) The revelation of the true proceeds represents the fact that, after the IPO, the firm is traded publicly which means there is a common knowledge signal of the true value of the equity. Even if the share price does not perfectly reveal this true value, it will provide a signal more precise than the private information of any one market participant and, more crucially, is publicly observable to all participants. The private signal of each underwriter is never revealed.

The timing of the stage game is summarized in Figure 1, and payoffs to each player in the stage game are given by

<table>
<thead>
<tr>
<th>Go public with (i)</th>
<th>Stay private</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>firm manager:</strong></td>
<td>(\phi((1 - \alpha')v(\alpha, u, x) + w) + \epsilon)</td>
</tr>
<tr>
<td><strong>underwriter (i):</strong></td>
<td>(\alpha'v(\alpha, u, x) + \theta(\alpha, u, x) - \kappa)</td>
</tr>
<tr>
<td><strong>underwriter (j):</strong></td>
<td>0</td>
</tr>
</tbody>
</table>

\(^4\) For reference, DeMarzo et al. (2005) provide a general treatment of auctions with security bids.

\(^5\) The selling mechanism that would permit the greatest collusive profits would condition spreads on realized price of the issue, rendering concerns about private information irrelevant. This mechanism is never observed in practice and does not appear to be a feasible choice for underwriters.

\(^6\) This expression for proceeds ignores dilution of firm value from spreads, which is unimportant in the analysis that does not consider underpricing.

\(^7\) As a technical convenience, we assume that the lowest bid and the value are revealed even when the firm chooses to remain private; this does not influence the main results.
Note that \( w \) will net out whenever the manager of the firm considers whether to take the firm public, so we can ignore this quantity. Allowing the firm to choose the fraction of equity to sell to raise new funds and introducing dispersed ownership, whereby existing private owners could choose to add their shares to the IPO, would potentially enrich the model. Unfortunately, the technical difficulties associated with including these additional choices in the model are formidable, and so we choose to maintain the simpler specification.

2. Spreads

In this section, we establish that the pattern of spreads observed in data is consistent with optimal tacit collusion but is not consistent with either competition or explicit collusion on monopoly outcomes.

Here, for simplicity, we eliminate the opportunity to underprice issues and restrict attention to the specific distributions of values, \( x \sim U[x, \bar{x}] \), and signals, \( \xi \) distributed as

\[
\xi_i = \begin{cases} 
  x & \text{with probability } p \\
  U[x, \bar{x}] & \text{with probability } 1 - p
\end{cases}
\]

We assume that manager preference is distributed \( \varepsilon \sim \exp(\lambda) \) and that underwriters are arbitrarily patient (the discount factor, \( \delta \), approaches 1). We set \( \phi = 1 \), which is largely without loss of generality; the only role of \( \phi \) is to scale the relative importance to the manager of private benefits and firm value, which is only relevant for interpreting the quantitative magnitudes implied by the later estimation of the model. Thus, we maintain this assumption until Section 4. At this stage we ignore dilution from spreads, with proceeds of the issue given as \( v = \beta x \).

2.1 Competitive and monopolist spreads

To show that tacit collusion provides a robust explanation for the observed pattern of spreads in contrast to the implications from competitive or monopoly IPO provision, we must first consider the implications
of the model for spreads under competitive and and monopolist provision of IPO services. The easiest way to see that both of these market structures will not predict rigid spreads is to derive the competitive and monopoly spread offer in the absence of uncertainty about the IPO value or manager preference:

**Proposition 1**
Competitive underwriters with perfect information would offer a spread of

\[ \alpha^c(x) = \frac{\kappa}{\beta x} \]  

(1)

to a firm selling a share of itself with value \( x \).

The proof is immediate from the zero profit condition for the underwriter. Unsurprisingly, spread offers decline in the value of the firm, as lower spread offers maintain the zero profit condition for larger firms because the cost of running an IPO is assumed to be fixed. More generally, only by assuming both that the costs of running an IPO are strictly proportional to the value of the firm and that underwriters observe the value of the firm without noise would competition generate rigid spreads. These assumptions are clearly unrealistic and in any case the rigidity will only exist in that single knife-edge case. Any small fixed costs component to the cost of running an IPO or any nonlinearity in costs would break rigidity, as would private information about the value of the firm.\(^8\) Further, any small change in parameters over time, as would likely occur over the course of several decades that include significant financial reforms, would imply that equilibrium levels of spreads would change over time. This has, dramatically, not been the case.

We now consider what a monopolist would choose for the spread, given perfect information about the firm.

**Proposition 2**
When underwriters act as a monopoly and have perfect information about \( x \) and \( \varepsilon \), if

\[ \varepsilon \geq \kappa - (\beta - 1)x, \]  

(2)

the spread function is given by

\[ \alpha^* = 1 - \frac{1}{\beta} + \frac{\varepsilon}{\beta x} \]  

(3)

\(^8\) An earlier version of this paper formally demonstrated that competition cannot generate spread rigidity generically.
and the firm chooses to hold the IPO. Otherwise,
\[ \alpha \geq 1 - \frac{1}{\beta} + \frac{\varepsilon}{\beta x} \]  
and the firm does not hold the IPO.

This proposition is established by finding the spread offer that leaves the firm just indifferent between holding and not holding the IPO; we omit this elementary calculation.

These limiting cases provide benchmarks for what monopolist or competitive underwriters would charge for IPO services. Absent dynamic effects associated with tacit collusion, underwriters with noisy, but informative, signals about value would use this information in setting spreads; for monopolists, signals about the value of the firm are informative about the reservation spread of the manager, even in the absence of any signal about manager preferences. For competitive underwriters, information about firm value is informative about the zero profit spread. This lack of rigidity holds very generally and does not depend on the specific distributions of values, signals, or preferences assumed. Only knife edge cases can produce rigidity under monopoly provision or competitive provision.

We now develop the argument for tacit collusion generating rigid spreads.

2.2 Optimal collusion
To investigate optimal collusion in the repeated game, we take a “mechanism design” approach to the problem, where long-lived players are assumed to solve an implicit mechanism design problem to choose the optimal equilibrium out of the large set of equilibria in the repeated game. The total payoff of the game to the long-lived player becomes the objective function the mechanism designer maximizes, while the constraint that all actions be part of an equilibrium serves the role of the incentive compatibility conditions. See Athey et al. (2004) for a more complete discussion of this approach. As is common in the literature on collusion, we focus on strongly symmetric perfect public equilibria (see Mailath and Samuelson 2006, for a detailed discussion of this equilibrium concept). Thus, an equilibrium of the game consists of a set of actions assigned to each player as a function of the public history of the game and his current period private information. The equilibrium also must specify a transition process to enter “punishment” phases, where the long-lived players proceed in an equilibrium with lower payoffs going forward. The risk of entering such an equilibrium gives the long-lived players the ability to construct equilibria with higher payoffs by deterring deviations that are profitable in the stage game. Crucially, because the
private information of the underwriter is never revealed, monitoring is
imperfect, and thus costly punishment phases may be triggered even when
all players adhere to the equilibrium strategy. This characteristic of the
model complicates the analysis because underwriters must take into
account the costs associated with entering these punishment phases
when selecting the optimal equilibrium. Confining attention to perfect
public equilibria allows for a recursive representation of the game.
Further, the bang-bang result of Abreu et al. (1990) greatly restricts the
set of punishment transitions that need to be considered.

To keep the analysis relatively simple and short, we consider two sets
of potential collusive equilibria: the optimal fully flexible spread schedule
and the optimal two-step self-enforcing spread. The optimal fully flexible
spread schedule assigns a different spread offer to every signal about firm
value observed by an underwriter. This flexibility allows underwriters to
maximize the joint payoffs to themselves from a static perspective, as they
can choose a spread schedule that closely matches what a monopolist
underwriter would select. Trying to collude on such a spread, however,
requires punishments for apparent deviations, which under strong sym-
metry implies that in some periods underwriters will revert to low profit,
competitive bidding for IPOs.

The alternative is to select two spreads: (1) the high spread, will be
offered by underwriters who receive a low signal about the value of the
equity stake to be sold, and (2) the low spread, will be offered by under-
writers who receive a high signal. Using such a spread schedule can com-
pletely avoid punishments that occur on the equilibrium path. To see
why, note that a monopolist will in general charge a very high spread
on IPOs in which the proceeds are expected to be very small; a low spread
will produce very low total payoffs when the proceeds are small, so the
underwriter prefers a high spread to profit from the occasional small IPO
in which the manager has a very strong preference for going public. A
carefully selected threshold above which the underwriter should offer the
low spread can guarantee that no underwriter with a signal below the
threshold will have an incentive to deviate to the low spread, pretending
to have received a higher signal. To see why, note that an underwriter
with a low signal about the size of the IPO will face a trade-off when
deciding whether to deviate to a low spread offer. On the one hand, the
deviating underwriter increases his chances of winning the IPO. On the
other hand, the lower spread offer will be less efficient in the sense of
extracting less of the total surplus from the firm. The trade-off arises
because, in the best equilibrium, the threshold for making the low offer
will be set low, in the region in which the first-best spread offer is high.
That is, an underwriter whose private information implies that he is just
under the threshold at which the equilibrium calls for him to offer 10% as
the spread could instead offer 7%. If his signal is correct and the other
underwriter also received this correct signal, deviating to 7% means that
the deviating underwriter will win the IPO. The deviating underwriter,
however, gives up the higher total fees associated with the 10% spread,
which at the threshold will be close to the first best level. Further, the
deviating underwriter must still split the winnings with those underwriters
who received a higher, and possibly incorrect, signal about the value of
the IPO proceeds. The optimal two-step self-enforcing equilibrium takes
account of this constraint and maximized total revenue over the set of
two-step spreads.

We present here the two programming problems that solve for the two
sets of equilibria discussed. Note that there will be a large number of
slack incentive constraints that we ignore: any deviation to a spread offer
that is never called for in equilibrium can be costlessly prevented through
a severe punishment threat because we assume the discount factor \( \delta \to 1 \).

To present the recursive representation of this class of equilibria, we
first define the necessary notation. Because the optimal (enforceable)
flexible spread schedule is strictly decreasing in the signal about firm
value,\(^9\) we can write the problem of finding the optimal flexible spread
by defining two functions, \( \alpha(\xi) \) and \( \rho(x_\xi) \), where \( \alpha \) maps signals into
spread offers and \( \rho \) maps the signal about value implied by the lowest
spread offer observed and the realized value of the firm into a probability
of reverting to the punishment phase. We also define \( R(\rho) \) as the uncondi-
tional probability along the equilibrium path of reversion to the pun-
ishment state, given a reversion probability function \( \rho \), and \( \Pi(\alpha) \) as the
unconditional value of the game for a given \( \alpha \).

Finally, we need the value of adhering to the proposed equilibrium for
any signal, \( \pi^A(\xi) \), and the value of deviating for any signal and any
possible deviation, \( \pi^D(\xi_\xi, \xi^\prime) \), where \( \xi^\prime \) is any misrepresentation of
the signal; that is, this is the payoff to an underwriter with a signal \( \xi \)
bids as if his signal were \( \xi^\prime \). The values \( v_h \) and \( v^\prime \) are the endogenous values
that are available in the cooperative phase and punishment phase,
respectively.

Details about the derivation of these and the following expressions,
discretized appropriately for the numerical implementation, are in
Internet Appendix A. The problem becomes

\[
\max_{v^h, v^\prime, \alpha, \rho} (1 - \delta) \Pi(\alpha) + \delta((1 - R(\rho))v^h + R(\rho)v^\prime)
\]  

\[\text{(5)}\]

\(^9\) This will not be true in general for all problems but is true for the cases we consider. The optimality of
decreasing spread offers is verified once the value of the game is computed.
subject to
\[ (1 - \delta)(\pi^D(\xi_i, \xi') - \pi^A(\xi_i)) \leq \delta \rho(\xi, \xi')(v^h - v') \quad \text{for all } \xi_i, \xi' \] (6)
\[ v^h, v' \leq v^* \] (7)
\[ v^h, v' \geq 0, \] (8)

and subject to the incentive compatibility constraints preventing deviations to spreads never observed in equilibrium, which will always be slack for \( \delta \) sufficiently large. Solving this problem amounts to finding a fixed point at which the optimized value calculated from Equation (5) equals \( v^* \). This solution gives an equilibrium spread offer function and an equilibrium reversion probability function.

Because punishments do not occur along the equilibrium path within the set of self-enforcing equilibria, the analysis for the optimal two-step self-enforcing spread is much simpler. Underwriters must choose the optimal combination of a high spread, a low spread, and a threshold for the signal such that an underwriter receiving the signal exactly at the threshold will be indifferent between offering the high and low spread even if no punishments are anticipated. We provide the precise statement of this program in Internet Appendix A.

2.3 Numerical analysis

These programs can be solved for any set of parameter values, and the resulting payoffs can be compared to determine whether flexible or rigid pricing dominates. Figures 2 and 3 present selected results from this numerical investigation. Panel (a) shows a horizontal line representing the value for the optimal two-step self-enforcing spread and a plot of the successive values calculated in the iterative procedure for determining the value of the fully flexible spread. In Figure 2, the value converges far above the value for the two-step spread, indicating that the flexible spread is preferred. In Figure 3, the value converges below the value for the two-step spread, indicating that the two-step spread dominates the flexible spread. Panels (c) and (d) present different views of the punishment probability function required to support the fully flexible spread. As expected, a less precise signal (in this case \( p = 0.37 \) versus \( p = 0.54 \)) can lead to cases in which the two-step spread dominates the flexible spread, while generating spreads similar to those observed.

From these above arguments, we establish that the optimal two-step self-enforcing spread can be optimal. Thus, we can use the model for a quantitative investigation of the IPO market under the assumption that the data are generated by the optimal collusive equilibrium. The first step in this process is presented in Figure 4. Here, we posit that spread offers of 10% for small IPOs and 7% for large IPOs are optimal, with a
threshold of $20 million; $20 million and 7% are are the basic characteristics of the data documented by Chen and Ritter (2000), whereas 10% is the most prominent mode observed in the data for smaller firms. We also cap total proceeds at $80 million, again based on the threshold presented by Chen and Ritter (2000). Above $80 million, the profits associated with a 7% offer appear to be too high to permit collusion at 7%, and so we see a decrease in the spread offer for these largest IPOs. This effect is directly related to the well-known equilibrium price war logic of Rotemberg and Saloner (1986) and could be modeled simultaneously with the spread rigidity for the smaller and medium-sized IPOs.

For simplicity, we do not present such an exercise here.

For different levels of the signal precision parameter $p$, panels (a)–(c) of Figure 4 show values for the lower bound on IPO size, $\bar{x}$, the cost of running the IPO, $\kappa$, and the inverse of the mean of the private payoff to the manager for an IPO, $\lambda$, that imply that the optimal two-step self-enforcing spread will call for 10% for IPOs below $20 million and 7% for IPOs above $20 million. Note that the value for the threshold is also chosen optimally. Panel (d) plots the payoffs of this two-step spread schedule against the payoff for the optimal fully rigid spread, where all

![Figure 2](image-url)

**Figure 2**
Payoff for flexible and rigid spreads: Flexible preferred.
For $\beta = 1.04$, $\bar{x} = 40$, $p = .54$, $\lambda = 14.9$, $\kappa = 1.36$, and $\lambda = 2.48$, panel (a) shows the value of the optimal self-enforcing two-step spread and the sequence of upper bounds for the value of a flexible spread. Panel (b) shows the optimal two-step and optimal flexible spreads. The two-step spread is not exactly at 10% and 7% because it is calculated for the discretized version of the problem. Panels (c) and (d) show two perspectives on the reversion probability function.
firms are offered the exact same spread. This alternative equilibrium is trivial to enforce because all deviations are “off path.” Unsurprisingly, as signals become less precise, the advantage of using the two-step spread rather than the fully rigid spread decreases.

2.4 Discussion: Three or more steps
Adding a third (or more) step will not necessarily increase the efficiency of the collusion. Formally, the additional step will introduce a new incentive compatibility constraint. Enforcing this new step may require distorting the optimal spreads further away from the first best spread offers. To see why, consider a three-step spread schedule. To enforce that the lowest signals make the highest offer, the second highest offer must be unappealing to underwriters with low signals. But, the lowest offer must also be unappealing to underwriters who receive a signal associated with the middle spread. To accomplish this the lowest spread offer may have to be distorted far downward; otherwise, it will be too appealing to underwriters with middling signals. This distortion may, and in our examples does, prove too costly to justify the slight increase in payoffs associated with a more nuanced set of spread offers for the smaller firms.

Figure 3
Payoff for flexible and rigid spreads: Rigidity preferred.
For $\beta = 1.04$, $\xi = 40$, $p = .37$, $\lambda = 16.1$, $\kappa = 1.42$, and $\lambda = 2.27$, panel (a) shows the value of the optimal self-enforcing two-step spread and the sequence of upper bounds for the value of a flexible spread. Panel (b) shows the optimal two-step and optimal flexible spreads. The two-step spread is not exactly at 10% and 7% because it is calculated for the discretized version of the problem. Panels (c) and (d) show two perspectives on the reversion probability function.
2.5 Discussion: Spread rigidity over time

In the data, spreads do not depend on the value of the issue, but they also show near-complete rigidity over time. This is perhaps surprising as the underlying parameters of the model likely fluctuate in response to the investment environment or general market conditions. Our model suggests an explanation for this as well. Even if underwriters believe that the parameters have changed, thus justifying a different collusive spread, they will not know if other underwriters have received the same information about the change in parameter values. Thus, we are again in a situation with imperfect monitoring; if underwriters want to collude on moving spreads in response to information about structural parameters, they will have to enforce the equilibrium by punishing apparent deviations, and these apparent deviations will occur even when all underwriters adhere to the equilibrium strategy. Unlike the value of the IPO, the true change in the parameter values will never be publicly observable, so enforcing a collusive agreement that responds to changes in parameter values over time requires private monitoring, which is more difficult to use to enforce efficient outcomes and thus likely will be even more costly than enforcing collusion that depends on the value of the IPO.

Figure 4
Two-step spread versus rigidity.
For $\beta = 1.01$ and $\bar{x} = 80$, panels (a)–(c) show the parameter values that support an optimal self-enforcing two-step spread that implies spreads of 10% and 7%, with the jump at a signal of the value of the equity stake being sold of $20$ million, for possible values of $p$. $\bar{x}$ is the lowest possible IPO size; $\kappa$ is the cost of running an IPO; and $\lambda$ is the inverse of the mean of manager preference. Panel (d) compares the value to the underwriters from this two-step spread to the value from a fully rigid spread, again as a function of $p$. 

The Pricing of IPO Services and Issues
203
This is not, however, a fully satisfying explanation. Although the parameters of IPO productivity certainly do not move around in a public fashion, there are public signals that likely correlate with these parameters. These signals themselves could be incorporated into the “mechanism design” problem because in principle actions can be conditioned on any public signal. We do not, in practice, observe any relationship between the spread level and, for example, the S&P 500 or other measures of market conditions.

In practice, this is not as surprising as it might appear in theory. The evolution of the concentration on the 7% and 10% spreads was gradual, indicating a slow convergence to the equilibrium without, necessarily, much communication. Indeed, such communication would have been illegal. Implicitly coordinating on how spreads should change over time in response to particular noisy signals about demand conditions would be far more difficult than coordinating on the most efficient fixed spread. Any attempt to coordinate on a dynamically changing spread would likely either unravel, as firms pretended to have beliefs about the effect of changes in market conditions that would call for a lower spread, or would result in a risk of costly price wars as punishments, leading to the same trade-off between short-term and long-term efficiency that holds in the model.

2.6 Discussion: Restriction to strongly symmetric equilibria
Because we impose strong symmetry, we rule out punishment of particular underwriters for apparent deviations from the assigned equilibrium spread offers. Thus, underwriters are assumed to only be able to use industry-wide punishments to deter opportunistic cheating on collusion. If underwriters were able to punish individual underwriters for apparent deviations, in certain cases higher payoffs would be available for flexible spread schedules and the rigidity result would be weakened. Note, however, that such punishments would take the form of temporarily excluding certain underwriters from the market following a spread offer that, ex post, looked too low. Such asymmetric equilibria would be, in effect, a bid rotation scheme and would be most likely detected as such by regulators. Thus, under the regulatory constraints facing underwriters, strong symmetric is a reasonable restriction.

Even without strong symmetry, however, the basic intuition for spread rigidity will go through as long as underwriters are not extremely patient. To see why, observe that under asymmetric punishments an underwriter must participate in his own punishment by refraining from bidding when public information identifies him as a potential deviator. For the underwriter to willingly refrain from bidding, he must anticipate being subject to an even harsher punishment in the future. With extreme patience, such equilibria can be constructed, but with less patient underwriters the
temptation to deviate to a competitive offer when excluded from bidding will be too strong. Thus, during phases in which underwriters must punish an apparent deviator, total profits must be distorted downward to keep that underwriter willing to participate in his own punishment. This cost will be borne on the equilibrium path, and thus a self-enforcing two-step spread can still dominate if we relax the assumption that the discount factor is arbitrarily close to one.

3. Underpricing

Under the rigid or partially rigid equilibrium described above, underpricing has a natural role. Underwriters collude to extract as much surplus from firms as possible. However, given the need to collude on a rigid or partially rigid spread and the imperfect information about firm value and preferences, those firms that choose to hold an IPO will still benefit. That is, firms are not pushed to their participation constraint, and indeed underwriters are not even extracting all surplus that they could given their information because they cannot use the information to set the spread. This section shows that underwriters may have an incentive to underprice issues in order to extract this additional surplus. The opportunity to underprice, in turn, reinforces the incentives to collude on a rigid or partially rigid spread; part of the loss from failing to charge the most efficient, flexible spread is recouped through the underpricing stage without requiring costly punishments.

To introduce underpricing into the model, we now assume that underwriters can choose to price the issue below (or above) the true market value. This offer price is selected after the firm has chosen an underwriter, as is the case of U.S. IPOs. In practice, offer prices are not set until immediately prior to the effective date of the IPO, and furthermore the offer price is not set or informally negotiated before the bookbuilding phase of the IPO, which must take place after the firm has committed to an underwriter and a spread. The underwriter chooses an amount of “money left on the table”, \( u \), which will be the total return to the initial investors on the first day of trading.

The benefit of underpricing to the underwriter represents the opportunity to receive direct or indirect kickbacks from investors who benefit from the initial increase in price. A number of sources support the idea that underwriters are able to extract such kickbacks. Specific examples include the fine levied on Credit Suisse First Boston for receiving kickbacks from investors benefiting from IPO underpricing through inflated commissions on future business (see Loughran and Ritter 2004, their footnote 1, for a summary of this case). Slightly less direct forms of kickbacks also appear to be common in the IPO market. Liu and
Ritter (2010) document “spinning” in the IPO market, a practice where underpriced shares are allocated to executives of other companies in exchange for implicit promises of future underwriting business. Griffin et al. (2007) find evidence of “laddering” in NASDAQ IPOs, where underwriters provide investors with underpriced shares in exchange for promises of investor participation in price support in the aftermarket. This results in a financial benefit for the underwriter because the underwriter does not have to expend his own resources to support cold IPOs and risk losing money if the stock loses value.

We parameterized these benefits as follows. The underwriter may choose to leave money on the table because we assume that some fraction \( \theta_1 \) of this return is kicked back to the underwriter. We also allow the underwriter to care directly about the percentage runup in the price on the first day. This allows the model to imply that underwriters may prefer to leave a certain amount of money on the table in larger IPOs versus smaller IPOs. This assumption proves essential for matching the qualitative characteristics of underpricing. We thus express the payoff to underpricing as

\[
\theta_1 u + \theta_2 \frac{u}{(1 - \alpha)v}
\]

where \( v \) is the proceeds. Proceeds will depend on both the value of the private firm and the level of underpricing, as we describe below.

The underwriter trades off the benefits of underpricing against the lower fees available on underpriced issues and against the risk that the firm will cancel the IPO when faced with too much underpricing. Because realized underpricing is variable, firms view it as possible that the actual level of underpricing will be zero. Thus, a firm will begin the IPO process when an IPO is profitable at a zero level of underpricing.

Note that we do not admit collusion on underpricing. When choosing the level of underpricing, the underwriter trades off the risk that the issuer will cancel the IPO against the higher payoff from higher underpricing. This is exactly the trade-off that the “mechanism designer” (who chooses the spread schedule) would consider when determining what policy the winning underwriter should use for pricing the issue. Thus, the collusive level of underpricing will correspond exactly to the level of underpricing that will be selected if the underwriters are left to their own devices. Thus, it is without loss of generality to ignore collusion on underpricing.

When determining the value of the firm as a public entity it is necessary to take into account that the firm faces expenses associated with going public. Specifically, the firm must pay the underwriter a fraction of the issuing proceeds, and when the issue is underpriced the firm must effectively transfer value to the shareholders who buy the issue from the
underwriter at the IPO. We assume that both costs are subtracted from the value of the firm before the public technology is applied to the firm. To see the purpose of this setup, consider two firms that both have the same value \( x \) as a private entity. Assume that both go public at the same spread, but one is underpriced by more than the other. The total value of the firm as a public entity (i.e., the market capitalization of the firm after trading starts) should be lower for the underpriced firm than for the firm that was underpriced by less, because the firm that faced more underpricing raised less capital than the firm that was not underpriced. Similarly, a firm that pays a higher spread will have less capital to operate with as a public firm. We capture this relationship by positing that

\[
v + u = \beta(x - \alpha v - u).
\]

The left-hand side of the equation is by definition equal to the market capitalization of the equity share of the firm sold in the IPO after it has begun trading, whereas the right-hand side summarizes the technological effect of having an IPO.\(^\text{10}\) Total proceeds can then be expressed as

\[
v = \frac{1}{1 + \alpha \beta} (\beta(x - u) - u),
\]

whereas from Equation (9), the payoff to the underwriter for underpricing can be written as

\[
\theta_1 u + \theta_2 \frac{1 + \alpha \beta}{1 - \alpha} \frac{u}{\beta(x - u) - u}.
\]

From this, we can write the manager’s decision problem. Because the manager only owns a fraction \( \phi \) of the firm, the manager will agree to go public at a spread \( \alpha \) and a level of money left on the table \( u \) if

\[
\phi x + \phi \cdot 0 \leq \phi (1 - \alpha) v + \varepsilon,
\]

which reduces to

\[
\frac{\varepsilon}{\phi} \geq x - (1 - \alpha) v.
\]

Because neither \( \phi \) nor \( \varepsilon \) is observable, we simply posit that the quantity \( \frac{\varepsilon}{\phi} \) itself is distributed randomly according to an exponential distribution with mean \( \frac{1}{\phi} \). Because we cannot separately identify the absolute level of manager benefit and the relative weight the manager places on this quantity relative to the value of the firm, we will conserve on notation by setting \( \phi = 1 \) and treating \( \varepsilon \) as coming from an exponential with

---

\(^{10}\) For simplicity, in the previous section we ignored the dilution of the firm associated with spreads. That is, we assumed \( u = \beta x \). This is unimportant when we do not consider underpricing.
mean $1/\gamma$. When interpreting parameters, however, it will be necessary to keep the more structural definition in mind.

We solve for optimal underpricing under the assumption that the firms are in a partially rigid equilibrium. Thus, the firm faces the decision of how much to underprice the firm conditional on winning the IPO at a given spread. Having won the IPO, the underwriter learns the true value of the firm and that manager’s preference is above a certain threshold.

**Proposition 3**

Optimal underpricing is given by

$$u^* = v \left\{ -\left( 1 + \alpha \beta \right) \theta_2 + (\alpha - 1)v(\theta_1 + \alpha(\beta \theta_1 - \beta - 1 + (\alpha - 1)(1 + \beta)\lambda v)) \right\} / (1 + \beta)(\theta_2 + (\alpha - 1)\lambda \theta_2 v - (\alpha - 1)^2 \lambda \theta_1 v^2).$$

(12)

This proposition is established by directly solving for optimal underpricing; this derivation can be found in Internet Appendix B.

Two implications arise immediately from this expression. First, whenever underpricing is positive, underpricing is decreasing in IPO productivity, $\beta$. This is because underwriters lose more from underpricing in terms of fees when $\beta$ is large, so they actually prefer less underpricing. To see why, consider an underwriter who is deciding by how much to underprice an issue. He can choose a low price for the issue and profit by having a fraction of the initial trading profits returned to him. Alternatively, he can choose a price closer to the true value of the share of the firm being sold, which allows the firm to obtain greater total proceeds. The underwriter receives a share of these total proceeds through spreads, and when $\beta$ is higher the value of allowing the firm to raise more is higher at the margin; proceeds are more valuable to both the firm and the underwriter when the IPO process is productive; and large $\beta$ aligns the incentives of underwriters and firms.

Second, and more obviously, underpricing is increasing in the mean of managerial preference for IPOs, $1/\gamma$. Here, underwriters use underpricing to extract the larger rents available when IPOs are more appealing, because in this case the higher preference for an IPO is not driven by the increased productivity of the IPO but instead by the idiosyncratic preferences of the manager.

### 4. Empirical Analysis

In this section, we estimate the parameters of the model from moment conditions derived from the underpricing predictions and from the mechanism design problem. We posit that the underpricing equation holds up to some additive disturbance term that satisfies the orthogonality assumption.
in a nonlinear regression model; it is orthogonal to the first derivatives of expression for underpricing given in Equation (12). The underpricing equation links the observed data on underpricing to the structural parameters of the model, permitting us to estimate the model entirely from data on underpricing. This, however, would ignore the information contained in the mechanism design problem that generates the optimal spread schedule. Thus, we now derive additional restrictions on the parameters from the assumption that the collusive scheme used is optimal.

The mechanism design problem for a two-step self-enforcing spread schedule is given by:

$$\max_{\alpha^h, \alpha^l, \xi^*} \int_0^{\bar{x}} \left[ P(\xi \leq \xi^* | x)P\left(\varepsilon \geq x - \frac{(1 - \alpha^h)\beta x}{1 + \alpha^h \beta}\right)(\pi(\alpha^h, u^*, x) - \kappa) \\
+ P(\xi > \xi^* | x)P\left(\varepsilon \geq x - \frac{(1 - \alpha^l)\beta x}{1 + \alpha^l \beta}\right)(\pi(\alpha^l, u^*, x) - \kappa) \right] f(x)dx$$

subject to a self-enforcement constraint that the underwriter with signal $\xi^*$ will prefer an offer of $\alpha^h$ (the spread assigned to firms with signals below $\xi^*$) to an offer of $\alpha^l$ (the spread assigned to a firm with a signal above $\xi^*$). Here, $f$ is the distribution of IPO size. The self-enforcement constraint may or may not bind, depending on the distribution of firm value and the quality of the signal about firm value. Because we cannot identify parameters on the distribution of value or signal quality, we make the simplifying assumption that the self-enforcement constraint will not bind. Note that functional forms assumptions on the distribution of firm value and underwriter information are no longer imposed, and we no longer assume that the number of underwriters is two because this quantity does not enter into any of the moment conditions used in the estimation. The number of underwriters drops out because the mechanism design moments maximize industry-wide profits, and we assume that the discount factor is sufficiently high to dissuade deviations.

Here, $\pi(\alpha, u, x)$, the payoff to the underwriter conditional on winning the IPO, is defined in Internet Appendix Equation (30).

We derive two moment conditions from the mechanism design problem, based on an optimal selection of the high spread to charge small firms and the optimal selection of the low spread to charge larger firms. We do not derive a moment condition based on choosing the optimal value for the threshold; such a condition would require positing distributions both for the precision of the underwriters’ signals and for the distribution of private firms with the opportunity to go public.

**Proposition 4**

Assume that the self-enforcement constraint does not bind and the optimal equilibrium involves two steps, $\alpha^h$ and $\alpha^l$. Let $\Gamma$ be the set of
structural parameters. Then, the following moment condition must hold for $i \in \{h, l\}$:

$$E[h(\alpha^i, v, u; \Gamma) | \text{IPO with } u^* \text{ at } \alpha^i] = 0,$$

where

$$h(\alpha^i, u^*, v; \Gamma) = -\left(\alpha^i v + \theta_1 u^* + \frac{\theta_2 u^*}{(1 - \alpha^i)v}\right) \frac{\lambda(1 + \beta)}{1 + \alpha^i \beta} v + \frac{v}{1 + \alpha^i \beta} + \frac{\theta_2(1 + \beta)}{(1 - \alpha^i)^2(1 + \alpha^i \beta)} u^* + \frac{\lambda(1 + \beta) \kappa}{1 + \alpha^i \beta} \left(v + \frac{1 + \beta}{1 + \alpha^i \beta} u^*\right) + \exp\left\{\frac{\lambda(1 - \alpha^i)(1 + \beta) u^*}{1 + \alpha^i \beta}\right\},$$

$i \in \{h, l\}$.

This derivation can be found in Internet Appendix C.

The above proposition shows that we can arrange the first-order condition for the optimal spread at each step as the expectation of some function of data and parameters, conditional on the firm receiving a spread offer and actually going public. This conditional expectation is equal to zero, and thus we can set the empirical analog to zero as a moment condition. Because the mean value of the $h$ function is zero conditional on a firm being in the sample at a given spread (i.e., either the high spread or the low spread), we have that $E[h(\alpha^i, u, v) | \text{IPO with } u^* \text{ at } \alpha^i] = 0$. Note, however, that this is not the statement $E[h(\alpha^i, u, v) | \text{IPO with } u^* \text{ at } \alpha^i, v] = 0$, which will not, in general, be true. Thus, this moment condition is not of the form of an elementary zero function. Instead, the restriction is that over all the data observed, the average must be zero.

### 4.1 Time-varying parameters

Our empirical implementation of the model is hindered by the presence of an apparent structural break associated with the dot-com era. Underpricing increased markedly during the period from 1995 to 2000, as can be seen in the summary statistics in Table 1. It is unlikely that data from this period were generated by the same set of parameters that generate data during normal times. To deal with this break, we allow parameters to vary over time. Specifically, we posit that the IPO market exists in one of two possible states: “normal” and “hot.” In each period, each underwriter receives a precise, but private, signal of the state. Because signals are private, they cannot be used to adjust the collusive spreads as discussed in Section 2.5 but will be used in the underpricing stage because underpricing actions are not monitored. Thus, when establishing the optimal collusive spreads the underwriters will use the prior belief about the probability of each state occurring, whereas
when determining the level of underpricing, each underwriter will use his own signal about the true state. The change in underpricing, then, will allow separate identification of the parameters from the normal period and the hot period.

Conceptually, we view the switch between normal and hot markets as a two-state Markov transition process with the initial state drawn from the stationary distribution of the chain, but given the limited time series we do not attempt to estimate these transitions. All that is relevant for the estimation is the prior probability of being in each state. The realized probability is approximately 25% in our sample period. This level seems perhaps high as a prior expectation, so we calibrate the probability of being in the hot market, $p_{dc}$, as 0%, 5%, 10%, 15%, 20%, and 25%. This calibration permits us to evaluate the market under a range of plausible ex ante beliefs about the likelihood of entering a hot market.

Permitting this variation in parameters over time requires recognizing that the moment equation in Proposition 4 has the expectation taken over both firm value and the value of the parameters. We can, abusing notation slightly, define $\Gamma = \Gamma_{dc} \cup \Gamma_{ndc}$, where $\Gamma_{dc}$ contains the parameter

<table>
<thead>
<tr>
<th>Year</th>
<th>Num.</th>
<th>Proceed(v) mean</th>
<th>Proceed(v) s.d.</th>
<th>MLT(u) mean</th>
<th>MLT(u) s.d.</th>
<th>Spread(\alpha) mean</th>
<th>Spread(\alpha) s.d.</th>
<th>$\alpha = 7%$</th>
<th>$\alpha = 10%$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985</td>
<td>22</td>
<td>24.10</td>
<td>22.24</td>
<td>1.98</td>
<td>2.57</td>
<td>7.81</td>
<td>1.20</td>
<td>18.18</td>
<td>30.32</td>
</tr>
<tr>
<td>1986</td>
<td>467</td>
<td>25.16</td>
<td>20.83</td>
<td>1.51</td>
<td>5.93</td>
<td>7.71</td>
<td>1.14</td>
<td>19.91</td>
<td>29.82</td>
</tr>
<tr>
<td>1987</td>
<td>349</td>
<td>24.32</td>
<td>20.57</td>
<td>1.17</td>
<td>4.36</td>
<td>7.84</td>
<td>1.23</td>
<td>24.07</td>
<td>34.38</td>
</tr>
<tr>
<td>1988</td>
<td>131</td>
<td>24.79</td>
<td>23.63</td>
<td>1.37</td>
<td>2.83</td>
<td>7.87</td>
<td>2.05</td>
<td>33.59</td>
<td>33.22</td>
</tr>
<tr>
<td>1989</td>
<td>114</td>
<td>29.27</td>
<td>23.17</td>
<td>2.52</td>
<td>3.95</td>
<td>7.69</td>
<td>1.17</td>
<td>51.75</td>
<td>39.32</td>
</tr>
<tr>
<td>1990</td>
<td>114</td>
<td>32.52</td>
<td>22.10</td>
<td>3.21</td>
<td>4.58</td>
<td>7.59</td>
<td>1.28</td>
<td>50.00</td>
<td>40.61</td>
</tr>
<tr>
<td>1991</td>
<td>251</td>
<td>35.75</td>
<td>21.74</td>
<td>4.66</td>
<td>7.60</td>
<td>7.44</td>
<td>1.01</td>
<td>66.93</td>
<td>41.78</td>
</tr>
<tr>
<td>1992</td>
<td>342</td>
<td>33.85</td>
<td>22.87</td>
<td>3.84</td>
<td>7.43</td>
<td>7.52</td>
<td>1.07</td>
<td>65.20</td>
<td>41.50</td>
</tr>
<tr>
<td>1993</td>
<td>441</td>
<td>36.39</td>
<td>23.35</td>
<td>5.03</td>
<td>9.99</td>
<td>7.43</td>
<td>1.01</td>
<td>68.25</td>
<td>42.94</td>
</tr>
<tr>
<td>1994</td>
<td>367</td>
<td>31.76</td>
<td>22.05</td>
<td>3.22</td>
<td>8.48</td>
<td>7.52</td>
<td>1.13</td>
<td>63.76</td>
<td>39.10</td>
</tr>
<tr>
<td>1995</td>
<td>389</td>
<td>38.33</td>
<td>22.77</td>
<td>8.85</td>
<td>12.75</td>
<td>7.52</td>
<td>1.15</td>
<td>73.52</td>
<td>45.82</td>
</tr>
<tr>
<td>1996</td>
<td>570</td>
<td>38.82</td>
<td>23.54</td>
<td>6.84</td>
<td>11.86</td>
<td>7.40</td>
<td>0.98</td>
<td>74.91</td>
<td>46.00</td>
</tr>
<tr>
<td>1997</td>
<td>378</td>
<td>38.46</td>
<td>22.87</td>
<td>6.18</td>
<td>10.87</td>
<td>7.40</td>
<td>0.97</td>
<td>77.78</td>
<td>45.31</td>
</tr>
<tr>
<td>1998</td>
<td>235</td>
<td>40.11</td>
<td>24.74</td>
<td>7.87</td>
<td>14.76</td>
<td>7.34</td>
<td>0.95</td>
<td>77.45</td>
<td>46.71</td>
</tr>
<tr>
<td>1999</td>
<td>273</td>
<td>54.34</td>
<td>22.54</td>
<td>21.72</td>
<td>27.81</td>
<td>7.11</td>
<td>0.55</td>
<td>89.74</td>
<td>58.32</td>
</tr>
<tr>
<td>2000</td>
<td>188</td>
<td>59.03</td>
<td>23.07</td>
<td>18.09</td>
<td>24.89</td>
<td>7.14</td>
<td>0.66</td>
<td>92.02</td>
<td>61.57</td>
</tr>
<tr>
<td>2001</td>
<td>32</td>
<td>55.90</td>
<td>23.85</td>
<td>8.94</td>
<td>12.10</td>
<td>7.23</td>
<td>0.93</td>
<td>78.13</td>
<td>62.47</td>
</tr>
<tr>
<td>2002</td>
<td>33</td>
<td>54.82</td>
<td>26.69</td>
<td>4.05</td>
<td>13.44</td>
<td>7.13</td>
<td>0.89</td>
<td>84.85</td>
<td>61.18</td>
</tr>
<tr>
<td>2003</td>
<td>32</td>
<td>60.68</td>
<td>29.98</td>
<td>6.95</td>
<td>11.33</td>
<td>7.17</td>
<td>0.73</td>
<td>87.50</td>
<td>66.80</td>
</tr>
<tr>
<td>2004</td>
<td>97</td>
<td>57.05</td>
<td>25.11</td>
<td>5.86</td>
<td>10.12</td>
<td>7.02</td>
<td>0.46</td>
<td>90.43</td>
<td>58.63</td>
</tr>
<tr>
<td>2005</td>
<td>80</td>
<td>50.40</td>
<td>28.12</td>
<td>4.69</td>
<td>8.41</td>
<td>7.03</td>
<td>0.43</td>
<td>78.75</td>
<td>56.91</td>
</tr>
<tr>
<td>2006</td>
<td>66</td>
<td>56.86</td>
<td>25.73</td>
<td>7.35</td>
<td>12.48</td>
<td>7.02</td>
<td>0.61</td>
<td>83.33</td>
<td>59.27</td>
</tr>
<tr>
<td>2007</td>
<td>76</td>
<td>59.26</td>
<td>26.79</td>
<td>7.06</td>
<td>16.18</td>
<td>7.00</td>
<td>0.60</td>
<td>84.21</td>
<td>65.69</td>
</tr>
</tbody>
</table>

All the monetary values are real in 2009 dollars and normalized by 1 million dollars. MLT: money left on the table.
values during the dot-com era hot market and $\Gamma_{ndc}$ the parameter values during the normal period. We can then express the expectation derived from the mechanism design problem as:

$$
(1 - p_{dc})E[h(\alpha', u^*, v; \Gamma_{ndc})] \text{ IPO with } u^* \text{ at } \alpha' \text{ in normal market } 
+ p_{dc}E[h(\alpha', u^*, v; \Gamma_{dc})] \text{ IPO with } u^* \text{ at } \alpha' \text{ in hot market } = 0.
$$

This moment condition guarantees that, when choosing the high and low spread, underwriters will take into account the subjective likelihood of being in each of the two possible states.

We posit that the IPO market may transition between normal and hot states, but we impose that the parameters governing the relationship between underwriters and investors are stable between the two states. That is, only the parameters on the productivity of IPOs ($\beta$) and the preference for IPOs ($\lambda$) are assumed to change between the hot market and the normal market. Thus, our complete set of parameters is given by $\{\theta_1, \theta_2, \beta_{dc}, \beta_{ndc}, \lambda_{dc}, \lambda_{ndc}, \kappa\}$.

4.2 Estimation

We estimate the model using GMM by combining the moment conditions associated with the orthogonal conditions from the underpricing equation of Equation (12) with the two moment conditions derived from the mechanism design problem. For underpricing, we assume that underpricing is given by this equation with some disturbance term:

$$
u = g(v, \alpha; \theta_1, \theta_2, \beta_k, \lambda_k, \kappa) + v \text{ for } k \in \{dc, ndc\},
$$

where $g$ is the expression from Equation (12), and, as mentioned above, $v$ satisfies the restrictions on errors for nonlinear regression. We assume that each underwriter’s private signal of the state of the market is sufficiently precise that we can use the parameters for the true state in the underpricing decision.

We have the following moment conditions to estimate the parameters:

$$
E\left[\nu \frac{\partial g}{\partial \Gamma}\right] = 0, \Gamma \in \{\theta_1, \theta_2, \beta_{dc}, \beta_{ndc}, \lambda_{dc}, \lambda_{ndc}, \kappa\},
$$

$$
E[h(\alpha', u^*, v)] \text{ IPO with } u^* \text{ at } \alpha' = 0, i \in \{h, l\}.
$$

Note that the derivative of $g$ during a normal market with respect to the parameters in the hot market is identically 0 and vice versa, so these moments drop out. Defining $n^k_m$ as the number of observations of

---

11 Under the assumption that the state follows a Markov process with transition matrix $\left(\begin{array}{cc} a & 1-a \\
1-b & b \end{array}\right)$ and that the initial state is drawn from the stationary distribution, then $p_{dc} = \frac{1-b}{1-a}$, where $dc$ is state 1. This structure thus admits both iid market states and the more reasonable partially persistent market states.
spread \( m \in \{l, h\} \) (the low spread and the high spread) in market, a sample analog of each of moment condition can be formed as follows:

\[
\frac{1}{n} \sum_j v_j \frac{\partial g_j}{\partial \Gamma}, \Gamma = \{\theta_1, \theta_2, \beta_k, \lambda_k, \kappa\}, k \in \{dc, ndc\},
\]

\[
\frac{1}{n_h} \left(1 - p_{dc}\right) \frac{1}{n_{dc}^h} \sum_j n_{dc}^h h(\alpha^h, u^*_j, v_j; \Gamma_{ndc}) + p_{dc} \frac{1}{n_{dc}^h} \sum_j n_{dc}^h h(\alpha^h, u^*_j, v_j; \Gamma_{dc}) \right),
\]

\[
\frac{1}{n_l} \left(1 - p_{dc}\right) \frac{1}{n_{dc}^l} \sum_j n_{dc}^l h(\alpha^l, u^*_j, v_j; \Gamma_{ndc}) + p_{dc} \frac{1}{n_{dc}^l} \sum_j n_{dc}^l h(\alpha^l, u^*_j, v_j; \Gamma_{dc}) \right),
\]

where \( j \) indicates an observation.\(^{12}\) Define \( G \) to be this vector of the sample analogs.

Thus, we numerically solve the following optimization problem to get parameter estimates:

\[
\min_{\theta_1, \theta_2, \beta_k, \lambda_k, \kappa, \lambda_{dc}, \lambda_{ndc}} G^T W G,
\]

where \( W \) represents a consistent weighting matrix.

We use efficient two-step GMM and impose that the moment conditions from the underpricing equation are uncorrelated with those from the mechanism design problem.

An immediate implication of the estimation equations derived is that \( \kappa \) must be vanishingly small as long as high levels of underpricing are observed. This can be seen from the last term in the \( h \) function, where \( \kappa \) is multiplied by the exponent of money left on the table. Because observed money left on the table is relatively large, this term explodes if \( \kappa \) is not almost zero. Intuitively, the high level of underpricing observed indicates that underwriters are willing to risk the failure of an IPO in the offer price stage. If \( \kappa \) were large, the underwriters would prefer to use higher spreads and less underpricing, thus facing fewer canceled IPOs.

It should also be noted that the costs of bookbuilding will likely depend on the equilibrium level of underpricing, because finding investors when the level of underpricing is high will be relatively easy. Although we

\(^{12}\) Note that the moment conditions from the underpricing equation and those from the mechanism design problem do not share the same number of observations. A subset of smaller IPOs do not receive either the 10% or 7% spread offer, which can be viewed as random deviations from the posited equilibrium, perhaps in response to some public signal; further discussion of these deviations can be found in Section 4.7. For these IPOs, the underpricing equation should still hold and so we do not exclude them from the underpricing moments. These deviations are more common in the earlier period, but again the optimal underpricing equation should still hold so we include this period in our estimation. Modeling the transition to the collusive equilibrium, although potentially interesting, is beyond the scope of the paper.
do not model the dependence, an estimate of \( \kappa \) is best seen as an estimate conditional on the level of underpricing. Thus, it is reasonable to think that \( \kappa \) would be quite small. Because \( \kappa \) must be estimated as vanishingly small, we set \( \kappa \) equal to zero to avoid computational problems that arise for large levels of underpricing. This simple identification strategy would not hold if the costs to the underwriter of running an IPO accrue after the stock is issued. At this point, however, the remaining role of the underwriter is primarily to make a secondary market in the stock. Although some forms of price support used following an IPO may end as a cost, overall the secondary market activity is more likely to be a profitable activity for the underwriter as he enjoys some temporary monopoly power in market making.

Still, the low level of underwriting costs needed to fit the model may appear surprising. We do note that many of the costs associated with IPOs are, in fact, borne directly by the issuing firm, rather than by the underwriter. This likely includes much of the costs associated with the road show. Further, the primary costs to underwriters throughout the IPO process is in the form of wages paid to investment bankers, which in our model may represent the distribution of monopoly rents arising from collusion, which should not necessarily be considered a cost. Further, a burgeoning literature on compensation in the financial sector (in particular, Philippon and Reshef 2012; Axelson and Bond 2012) supports the idea that salaries in finance, and investment banking in particular, may include rents. Axelson and Bond (2012), in particular, propose that the long hours put in by junior investment bankers on menial tasks may arise because moral hazard considerations push expected compensation far above the participation constraint of new employees; in their model, investment banks then have a pool of workers to complete easily monitored tasks, such as dealing with logistical issues for book building and road shows, at effectively no additional cost beyond what they already expect in incentive pay for other tasks.

We do note, however, that certain aspects of the IPO process that we do not model could bias our cost estimates. Our model assumes that firms cancel their IPOs when the level of proceeds raised falls below their participation constraint due to underpricing. We do not explicitly account for other sources of cancellation, such as market movements or accounting issues. We also assume that a withdrawn IPO still imposes the full costs on the underwriter; if a firm learns sufficiently early that the underwriter is likely to choose an unacceptably high level of underpricing, the firm may withdraw prior to the underwriter bearing most of the costs of the IPO. In this case, an underwriter would be more willing to risk cancellations even if the costs are not trivially small.

Note that we abstract away from cross-sectional parameter heterogeneity that very likely exists in the data; that is, firms in certain industries
may have different benefits from going public and a different degree of uncertainty about firm values. The fact that we do not observe different spread schedules for firms in different industries suggests that these cross-sectional-differences in parameter values are not common knowledge among underwriters. We leave exploration of such cross-sectional differences, which could manifest themselves in different underpricing patterns under the same spread schedule, to future work.

4.3 Identification

Identification of the remaining parameters comes from two sources: spreads and underpricing. Because spreads are set in a collusive equilibrium and thus do not vary continuously with proceeds or any other observable data, spreads alone are an insufficient source of identification. A number of parameter combinations could predict 10% and 7% for the high and low spread offers, as the calibration exercise in Section 2.3 shows. Without variation in spreads, distinguishing which of these values are most likely to have generated the observed data is impossible.

Underpricing, on the other hand, behaves in a way that is much more amenable to standard empirical analysis. The level of underpricing varies depending on the proceeds of the issue and over time. Heuristically, then, we can view identification as a two-stage process. The fact that 10% and 7% are selected as the optimal two-step schedule restricts the possible set of parameters to those for which 10% and 7% are optimal. Then, within this constrained set, underpricing pins down the parameters most likely to generate the observed level of underpricing and the relationship between proceeds and underpricing. Of course, the estimation procedure is in fact simultaneous, and the trade-off between the fit for underpricing and spreads is given by the efficient GMM procedure.

In more detail, identification requires first finding an estimate of the overall benefits of an IPO to the firm and then decomposing that into firm benefits and idiosyncratic managerial benefits. Absent underpricing, the rigid spread level provides a good estimate of the total value produced by the IPO, as the willingness of a firm to accept that spread offer reveals the value created. This source of identification is similar to the standard way that bids identify parameters of the distribution of value in auction estimations. Decomposing the benefit of the IPO into the proportionate benefit to the firm and the idiosyncratic benefit to the manager can be partially accomplished by looking at the gap between the low spread offer and the high spread offer. For example, if the average level of managerial preference were low (i.e., \( \frac{1}{2} \) is large), the difference between the optimal spread to charge smaller firms and the optimal spread to charge larger firms would be small; the bid would simply reflect the proportionate increase in value arising from the IPO of a given size. Because the gap
between the spreads charged for small IPOs and large IPOs is positive, then manager preference must play some role; underwriters set spreads higher for small firms to capture the relatively large payoff from small IPOs in which the manager has a strong preference for going public.

The second source of identification of both the level and division of the benefit from IPO comes from the level of underpricing and the relationship between underpricing and the proceeds of an IPO. Dilution from selling shares below their value represents a significant share of the costs of holding an IPO and thus must be accounted for when determining the total benefit that should arise from an IPO. Further, the relationship between IPO proceeds and underpricing is useful for identifying the magnitude of the idiosyncratic benefit component. In particular, the slope of the relationship between proceeds and underpricing helps to pin down the value of $\lambda$.

The particular relationship between proceeds and underpricing that will be predicted for any given $\lambda$ depends on the payoff that underwriters receive for underpricing. Our analysis implies that the payoff to underpricing depends on the relative, as well as absolute, amount of money left on the table. Specifically, the value of a given level of money left on the table is lower when total proceeds are lower, indicating an aversion to particularly large price runups (i.e., $\theta_2 < 0$). This aversion is identified from the fact that the average level of money left on the table is rising in proceeds. A consequence of this result is that underpricing for smaller IPOs is restrained by the aversion to large price runups. Thus, the level of underpricing for small IPOs is not particularly sensitive to the parameter $\lambda$, which is informative about the probability that a firm will walk away from an IPO at a particular level of underpricing. For larger IPOs, however, the concern about the percentage runup in the price is less relevant, and underpricing is restrained instead by concern about cancellation if the price is set too low. Thus, for larger IPOs the level of underpricing depends much more on the risk of cancellation. For these larger firms a low level of $\lambda$, which corresponds to a high average level of managerial preference for an IPO, leads to higher underpricing. Thus, a steep, positive relationship between proceeds and underpricing implies a low $\lambda$, while a flatter relationship implies a high $\lambda$.

The effect of $\beta$, the productivity of the IPO for the firm, on the underpricing proceeds relationship is different. As described in section 3, higher $\beta$ makes underpricing less appealing to the underwriter. This effect is relatively uniform across moderate and large IPOs, such that changes in $\beta$ tend to lead to a more parallel shift in the relationship between proceeds and underpricing than do changes in $\lambda$. This difference in the effect of $\beta$ and $\lambda$, combined with the fact that total benefits to the IPO are pinned down through the total costs managers are willing to pay for the IPO, leads to a secondary source of information to separate $\beta$ and $\lambda$. 

216
Note that our model does not admit the various alternative explanations of underpricing where underpricing is, in effect, a service demanded by firms. For example, in the models of Rock (1986) and Benveniste and Spindt (1989), underpricing is a necessary part of the IPO process, designed to overcome the winner’s curse and the incentives for initial investors to hide their information, respectively. Allen and Faulhaber (1989), in turn, model underpricing as a signaling mechanism that provides a direct benefit to firms in the form of higher future proceeds in subsequent issues. Even more directly, Loughran and Ritter (2002) propose that the managers of firms obtain utility from underpricing due to preferences consistent with prospect theory. If the two latter explanations of underpricing are quantitatively relevant, our estimates of the benefit of an IPO, both from to the firm and to the manager, would tend to be biased upward. Because identification of the value of having an IPO comes, in part, from the willingness to tolerate underpricing, if some baseline level of underpricing is actually desirable this direct value for underpricing will be misattributed to the value from the IPO itself. Formally, ignoring these effects amounts to omitting a direct, concave benefit for underpricing in the payoff function of the firm. The Rock (1986) and Benveniste and Spindt (1989) stories would likely have a similar effect on our estimates, although in these cases the analysis would not be as clear because underpricing is treated as part of the selling mechanism and thus does not fit as directly into our framework.

We also note that the objective function is not guaranteed to be globally concave even in the relevant parameter region. We address this issue by using a large grid of starting values, which also confirms local identification of the model. Unsurprisingly, we do detect more than one local minimum in the estimation; different combinations of parameters are more or less likely to produce the observed behavior than nearby parameters, even if they are not the most likely parameter values globally.

In fact, we detect two close by local minima. We focus on the minimum of the objective function that is the global minimum for values of the prior on the dot-com period, $p_{dc}$, closest to that observed in data.\footnote{The minima are continuous in the calibrated parameter but for very low values of the calibrated parameter the identity of the global minimum switches.} Internet Appendix Table 1 presents the alternative estimates. All results are qualitatively similar using the other local minimum, and in particular the implications of the policy experiments, also presented in the Appendix, are the same.

It is also necessary to address how the data on underpricing identify the cause of the run-up in underpricing during the dot-com era. To capture the presumed stability of relationships between underwriters and professional investors, we impose that the parameters on the fraction of the first
day return do not change over time. This leaves two possible explanations for the runup in underpricing: a large increase in the average managerial preference for going public or a decrease in the productivity of IPOs, as described in the discussion of proposition 3. Our estimation can separate these two explanations through the relationship between underpricing, the change in underpricing, and the level of proceeds.

4.4 Data
From the SDC database, we collect data on proceeds, money left on the table and spread for those firms who have gone public in the U.S. market from 1985 to 2007. We exclude REIT, ADRs, and unit offerings.14 We also restrict attention to IPOs with proceeds less than $100 million, because this is the region over which the two-step spread schedule is observed. For the estimation, we normalize $100 million to one. We include stocks with zero underpricing and overpriced stocks; in the model, underwriters can sell an issue for more than it is worth, but they will bear a fraction of the losses borne by their clients. After restricting our sample, we end with 5,044 observations, among which 3,222 firms were charged 7% by underwriters and 547 firms were charged 10%.15 The descriptive statistics of our final sample are presented in Table 1 year by year. The number of IPOs was high during the 90s and low in the 2000s. The amount of money left on the table in 1999 and 2000 was very high, indicating the last period of the dot-com bubble. It is obvious that substantial fraction of the spreads was set at 7%. The trend is less prominent in the 80s, but after that, the percentage of IPO charged at 7% generally increases until 2000 and remains a large proportion until the end of our sample. On the other hand, the number of IPOs with a spread of 10% is much lower and has decreased markedly since the passage of the Sarbanes-Oxley act.16 This fact is consistent with the model in the sense that, if the benefit of going public declines but the level of this decline is not common knowledge, we expect to see small firms either choose to remain private or to only go public when offered the 7% spread.17 As expected, we see that the average proceeds for IPOs with a 7% spread is much larger than for those with 10%.

14 These restrictions follow Chen and Ritter (2000).
15 When the “dot-com bubble” period (1995–2000) is excluded, the number of observations is 3,011, among which 1,615 firms were charged 7% by underwriters and 376 firms were charged 10%.
16 Prior to this event, we cannot detect any change in the threshold between 7% and 10% offers; for example, there does not appear to be any trend when we look at IPOs by year in the smallest IPO to receive a 7% offer or the largest IPO to receive a 10% offer.
17 It appears from the data that the first effect dominates; the near disappearance of IPOs at 10% after 2001 appears to arise primarily because far fewer of these small IPOs take place at all.
4.5 Results
We estimate the structural parameters as a function of the calibrated parameter given the prior probability of the hot market. Our estimates of the structural parameters are qualitatively similar, in most cases, for different values of the ex ante probability of a hot market. In particular, the qualitative implications of the policy experiments are largely stable over the specifications.

Table 2 shows the estimates for different values of the calibrated prior on the hot market. The prior on the dot-com period is calibrated to 5%, 10%, 15%, 20%, and 25%, respectively. We estimate a productivity parameter $\beta$ between 0.97 and 1.45 for normal times, with higher estimates for lower priors on the hot market. The mean of manager preference ($\frac{1}{\gamma}$) varies between 0.25 and 0.29, indicating a strong but not overwhelming private motive for going public. Recall that this preference is denominated in lost investor dollars and should not be interpreted in raw dollar terms. Thus, these averages correspond to a willingness to forgo around $25–29$ million in value to pre-IPO shareholders in order to have an IPO.

The proceeds and the amount of money left on the table observed are plotted in Figure 5 along with the curve that is obtained using our estimated underpricing equation, Equation (12) for a prior probability of the hot market of 15% and with both the non-dot-com era and the dot-com era parameters and data. The estimated function is able to capture the general pattern of the data, where small issues leave less money on the table than do large issues, but mean money left on the table becomes nearly flat as a function of proceeds for the higher levels of proceeds. The estimated relationship between proceeds and money left on the table is noticeably steeper for the dot-com period.

All of the calibrations identify a similar explanation for the higher level of underpricing in the dot-com era. The increase is driven by a large increase in the expected value of the managerial preference term. Interestingly, this increase is coupled with a (perhaps implausibly) large increase in the value of the productivity parameter. This feature of the estimates is notable because an increase in IPO productivity in our model should decrease underpricing; instead, the decrease implied by this higher productivity is swamped by the very large increase in managerial preference for IPOs. These parameter changes are needed to capture the pattern in relationship between proceeds and underpricing during the dot-com era. The low $\lambda$ is needed to account for the steepness of the relationship in the middle range of our sample, but such a low $\lambda$ provides too steep a relationship at the higher end. Increasing $\beta$, however, flattens the predicted relationship at the high end without undoing the steepness in the middle. Intuitively, a high $\beta$ renders very high underpricing less appealing for larger IPOs; the effect of high $\beta$ is strongest for larger IPOs because
underpricing is already set at a level that makes cancellation unlikely, thus making high firms relatively more sensitive to the \( \beta \) effect than to the \( \lambda \) effect.

An interesting feature of our estimates is that the combined increase in managerial preference and IPO productivity should imply that many firms find it valuable to go public in the hot market. As is obvious from the summary statistics, the period we identify as the hot market based on the large first day runups in price also exhibits the highest level of IPO activity. This increased activity is not used in the

![Table 2](image)

### Parameter Estimates

<table>
<thead>
<tr>
<th>Parameters</th>
<th>(a) 5%</th>
<th>(b) 10%</th>
<th>(c) 15%</th>
<th>(d) 20%</th>
<th>(e) 25%</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_1 )</td>
<td>1.265</td>
<td>1.174</td>
<td>1.169</td>
<td>1.151</td>
<td>1.144</td>
</tr>
<tr>
<td>( \theta_2 )</td>
<td>-0.141</td>
<td>-0.136</td>
<td>-0.134</td>
<td>-0.132</td>
<td>-0.131</td>
</tr>
<tr>
<td>( \beta_{nk} )</td>
<td>1.451</td>
<td>1.29</td>
<td>1.198</td>
<td>1.08</td>
<td>0.972</td>
</tr>
<tr>
<td>( \lambda_{nk} )</td>
<td>3.388</td>
<td>3.411</td>
<td>3.59</td>
<td>3.8</td>
<td>4.042</td>
</tr>
<tr>
<td>( \beta_{dc} )</td>
<td>4.413</td>
<td>3.577</td>
<td>3.561</td>
<td>3.317</td>
<td>2.984</td>
</tr>
<tr>
<td>( \lambda_{dc} )</td>
<td>0.539</td>
<td>0.612</td>
<td>0.618</td>
<td>0.666</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Estimates from the primary objective function mode for different prior probabilities of the hot market. Standard errors are in parentheses.

![Figure 5](image)

**Figure 5**

Proceeds and money left on the table. Prior probability of hot market set to 15%. The non-dot-com era is on the left, and the dot-com era is on the right.
estimation procedure to identify parameters; identification obtained only from the level of spreads and the relationship between proceeds and underpricing, but the parameter estimates for the dot-com period imply that basically any firm that considers going public will do so. Thus, the highly active IPO market serves to some extent as validation of the model.

We now evaluate the implications of the model estimates for relevant economic quantities; we address the payoff to underwriters for underpricing, the share of the total money left on the table that is returned to the underwriter, and the probability of the IPO continuing as a function of proceeds. For the parameters on the benefit of underpricing to the underwriter, the curvature parameter $\theta_2$ indicates that gross underpricing is less costly for large firms, consistent with the potential for regulatory pressure or bad publicity associated with very large runups in stock prices and contrary to the claim that underpricing itself is desirable. Figure 6 shows the payoff from underpricing (conditional on the IPO being completed) as a function of the level of money left on the table for a project of size $x = $50 million. The dot marks the optimal level of underpricing.

The underwriter captures almost all of the surplus from underpricing for moderate-sized IPOs for all of the calibrations of the prior probability. When the prior probability is calibrated at a higher level, the share of underpricing profits estimated to accrue to the underwriter is lower. Figure 7 shows the predicted fraction of the value of underpricing returned to the underwriter as a function of the project value for the calibrated prior probability of the hot market. The smallest IPOs, with $x$ around $20 million, show a lower fraction of the money left on the table returned to the underwriter. This level for smaller firms is in line with estimates from practitioners reported in Liu and Ritter (2011).

We can derive additional empirical implications of the model as well, in light of the estimated parameters. In Figure 8 we show the implied relationship between proceeds and IPO cancellation. Our results indicate that the probability of cancellation in normal times increasing in proceeds initially but with a slight drop between moderate and high values. For the dot-com era, the probability of continuing the IPO is monotonically decreasing in proceeds. Empirically, Busaba et al. (2001) (see Table 4) find an increasing relationship between proceeds and the probability that the IPO is canceled; in more recently reported results, Hao (2011) finds the opposite to be true. Overall, the probability of cancellation that we estimate appears to be higher than the unconditional probability of cancellation observed in data as presented in Busaba et al. (2001) and Bernstein (2012), who find that about 15–18% of IPOs are withdrawn.
Figure 6
Payoff from underpricing for a $50 million IPO with a 7% spread. The left figure represents normal time estimates, and the right figure estimates for the hot market parameters. Estimates are for a 5%, 15%, and 25% prior on the hot market. The x-axis is the level of money left on the table in $100 million units, and the y-axis is the payoff to the underwriter, also in $100 million units.

Figure 7
Expected fraction of money left on the table returned to the underwriter. Assumes a 7% spread. The left figure represents normal time estimates, and the right figure estimates for the hot market parameters.

Figure 8
Probability of continuing with the IPO following the expected level of underpricing. The left figure represents normal time estimates, and the right figure estimates for the hot market parameters.
For all but the smallest firms our estimation suggests a higher cancellation probability of around 30%.¹⁸

4.6 Policy analysis
Because we recover estimates of structural parameters, we can use our model to investigate policy interventions in the IPO market. Two types of intervention naturally present themselves: reducing spreads and reducing underpricing. Both of these polices relate to proposed policies in the IPO market. For example, the Department of Justice investigated, but ultimately did not act against, the high, rigid spreads in the IPO market (WSJ 2001). Additionally, the SEC has taken actions against some of the more-transparent forms of kickbacks in the IPO markets (SEC 2004). Motivated in part by these potential interventions, we investigate the effect of a spread cap and a decrease in the incentives to underprice issues. Our results indicate that directly intervening on spreads is unlikely to increase either the frequency or proceeds raised in IPOs, whereas discouraging underpricing may be a more effective policy response.

For our policy experiments, we consider the effect of the interventions on a firm with a privately held project who has the opportunity to go public. We consider the effect of the policies on the probability that the firm will take the opportunity to have an IPO and on the proceeds that will be raised in the IPO. Because the costs to the underwriter of holding an IPO must be quite small relative to the benefits of the IPO for our model to fit the data (i.e., the cost parameter κ must be close to zero), increasing IPO frequency is a reasonable policy objective. Even without this result, IPO services will, in general, be underprovided in the collusive oligopoly modeled here. A policy maker also may naturally seek to increase the proceeds of an IPO because the productivity of IPO proceeds within the firm is high; thus, underpricing should not be necessarily viewed as a pure transfer but may also have a social cost. We do not consider the effect on the total level of private benefits that will accrue to the manager of the firm. These private benefits could be a source of value, but they may also represent a moral hazard problem within the firm. For the purpose of evaluating policy, then, we prefer to consider only the level of proceeds raised. We also do not net out the benefit to investors of gaining underpriced shares because we view maximizing the proceeds available to firms, which presumably have productive uses for the capital, as the goal of intervention in the market. We focus attention primarily on moderate-sized firms during normal times. Because we do not focus much

¹⁸ While this number exceeds the average cancellation probability over the entire sample period, in some years the cancellation probability is as high as 59%, indicating that our estimates generate qualitatively realistic cancellation probabilities.
on the smallest firms, we assume a spread of 7% in the absence of direct intervention on spreads.

The simplest policy experiment is a price control on spreads. Because costs of running an IPO are estimated to be vanishingly small, price controls would not appear to have a deleterious effect on IPO volume or proceeds. The supply of IPO services should not be reduced by decreasing spreads, and reducing monopoly rents acquired through collusion should lead to more IPOs and greater proceeds. This intuition fails for the parameter values implied by our estimates.

The reason for the failure of this policy is that underwriters will respond to being forced to offer a lower spread by increasing underpricing. At lower spreads, firms will, on average, be further from their participation constraint. Underwriters will respond to this by increasing underpricing. This response should partially offset the beneficial effect of lower spreads. Unfortunately, in the parameter region implied by our estimates, another effect becomes relevant and leads to even greater underpricing. This second effect arises from the role of the spread as an equity share of the IPO project. Reducing the spread makes compensation through underpricing more valuable relative to compensation through fees. The incentives to underprice thus become even stronger, and the total effect of the increased underpricing reduces the proceeds from an IPO and the probability that the firm will go public.

This result comes from the empirical analysis, not simply from the model. For some parameter values other than those we recover from our estimation, forcing spreads lower than 7% has a beneficial effect. Specifically, for higher values of the linear underpricing payoff term, $\theta_1$, forcing spreads lower does lead to more IPOs and higher proceeds. In this case, the concavity of the payoff to underpricing restrains underpricing sufficiently to weaken the second effect, allowing lower spreads to generate more IPOs.

We present selected results from this spread cap experiment in Table 3. We impose a spread cap of 5% and 3% for each calibrated value of the prior probability of the dot-com era and for $x = \{20 \text{ million}, 40 \text{ million}, 60 \text{ million}, 80 \text{ million}\}$ in normal times. For the probability of going public, almost all cases have the result that a spread cap decreases the frequency of IPOs, but the effect is very small. For the largest firms, at low prior probabilities, the effect is reversed and goes in the more intuitive direction, but here the effect is still very small. The worst result of a spread cap is a drop of approximately 0.6 percentage points in the probability of holding an IPO for smaller firms ($20 \text{ million}$ and $40 \text{ million}$) under the higher hot market priors ($\geq 20\%$). The best outcome is a 2.6 percentage point increase for a 3% cap for IPOs with a size of $60 \text{ million}$ under the 5% hot market prior. For the parameters estimated for the dot-com
period, the probability is estimated to be one in almost all cases with and without policy intervention.

Table 4 shows the effect on proceeds for the same experiment. Here, the spread cap uniformly decreases proceeds regardless of the size of the IPO. The effects work against the policy objective and are small in most cases. Overall, a spread cap decreases proceeds as the increased incentives to underprice issues overwhelms the value of leaving more of the IPO proceeds to the issuers. Because this policy is dominated by no intervention, we do not complete the analysis of aggregate changes.

Our second policy experiment is to reduce the share of the underpricing profits that investors return to the underwriter. This intervention should increase the frequency of IPOs because the benefit of underpricing will fall, while the cost (from an increased probability of IPO cancellation) will remain the same. Thus, underwriters should decrease the level of underpricing, which will make firms more likely to proceed with an IPO. In our analysis, we consider a policy that reduces \( \theta_1 \), the linear term on the payoff to underpricing for the underwriter. There would be a number of ways to achieve this reduction. Increased scrutiny and penalties for apparent quid pro quo could produce a lower \( \theta_1 \), although this would depend on the effectiveness of the enforcement strategy. Another alternative that would be more certain to have the intended effect and that would not depend on successful investigations into particular business practices would be to require that some fraction of every issue be allocated by lottery to investors willing to pay the IPO price; this approach would largely guarantee that at least these investors would not be in established relationships with underwriters.\(^{19}\) In contrast to the spread cap, decreasing the value of underpricing to underwriters

\[^{19}\text{Such a policy would be costless under our assumptions but could reduce the capacity of underwriters to gather information if the underpricing-as-compensation models of Benveniste and Spindt (1989) and Benveniste and Wilhelm (1990) actually hold. Because a quantitative model of this effect is not, to our knowledge, available it is not possible to calculate the potential costs under this alternative rationale for underpricing.}\]
improves outcomes for potential IPO firms. This beneficial effect comes partially from the assumption that underwriters cannot react by moving to a new, higher spread equilibrium, which is consistent with the argument for time-series spread rigidity in Section 2.5.20 But, it also comes from the fact that decreasing the benefit of underpricing does not generate negative incentives in the way cutting spreads does.

Table 5 presents the results of the experiment cutting $\theta_1$ to 80% and 60% of $\theta_1$. During normal times, the change to $\theta_1$ appreciably increases the probability of holding an IPO and increases the proceeds available. Motivated by these results, we attempt to calculate an estimate of the value that could be created by instituting such a policy. Specifically, we ask how many more IPOs would have occurred over the sample period if $\theta_1$ was lower and how much more proceeds would have been raised. We divide the sample into the two normal periods: prior to the dot-com boom and after the dot-com boom. Using the estimates from the model, we calculate an empirical distribution of IPOs based on proceed deciles. We focus on the eight largest deciles, which correspond to the IPOs that are very likely to pay a 7% spread. First, we calculate the probability that a firm will go public given the opportunity based on our estimates, and we use this to recover the number of firms that could have gone public. Then we apply the counterfactual probability of taking the opportunity to go public to this number of potential IPOs to obtain the counterfactual number of IPOs for each decile. Tables 6 and 7 present the results. The effect of a 20% reduction in $\theta_1$ is strong, particularly when the prior probability of the dot-com era is calibrated at a high level. For example, in the pre-dot-com era each decile, except the two highest, shows an increase of at least 20 IPOs under the policy experiment when the prior probability is calibrated to 25% (the lower deciles have an increase of over 30 IPOs). The effect on proceeds is also large, as presented in Tables 8 and 9.

20 We also assume that underwriters continue to offer the 7% spread, rather than increasing the spread to 10% for a significant fraction of the firms. Thus, we may overstate the beneficial effect of the underpricing intervention on the smallest cohort of IPOs we consider.
We conclude that intervention on reducing underpricing is a more effective approach to increasing the frequency and proceeds from IPOs than is decreasing spreads. Although high spreads transfer wealth from potentially productive firms to collusive underwriters, the spread partially

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Underpricing Policy: Probability of Completing IPO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
</tr>
<tr>
<td>IPO Size</td>
<td></td>
</tr>
<tr>
<td>$20 mil</td>
<td>0.981</td>
</tr>
<tr>
<td>$40 mil</td>
<td>0.872</td>
</tr>
<tr>
<td>$60 mil</td>
<td>0.97</td>
</tr>
<tr>
<td>$80 mil</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Effect of a reduction of the payoff to underpricing on the probability of completing an IPO of size $20 million, $40 million, $60 million and $80 million, with 80% and 60% of the estimated $t_i$. Results are for the normal period.

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Aggregate Effect of Underpricing Policy: Pre-Dot Com Era</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
</tr>
<tr>
<td>Size</td>
<td>$t_i$</td>
</tr>
<tr>
<td>$0.22$</td>
<td>297</td>
</tr>
<tr>
<td>$0.28$</td>
<td>274</td>
</tr>
<tr>
<td>$0.35$</td>
<td>214</td>
</tr>
<tr>
<td>$0.42$</td>
<td>151</td>
</tr>
<tr>
<td>$0.49$</td>
<td>132</td>
</tr>
<tr>
<td>$0.56$</td>
<td>114</td>
</tr>
<tr>
<td>$0.63$</td>
<td>80</td>
</tr>
<tr>
<td>$0.69$</td>
<td>59</td>
</tr>
</tbody>
</table>

Aggregate effect of decreasing the linear underpricing payoff term by 20% on the number of IPOs in the pre-dot com era. Size is in $100 million units.

<table>
<thead>
<tr>
<th>Table 7</th>
<th>Aggregate Effect of Underpricing Policy: Post-Dot Com Era</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
</tr>
<tr>
<td>Size</td>
<td>$t_i$</td>
</tr>
<tr>
<td>$0.36$</td>
<td>59</td>
</tr>
<tr>
<td>$0.46$</td>
<td>45</td>
</tr>
<tr>
<td>$0.57$</td>
<td>54</td>
</tr>
<tr>
<td>$0.67$</td>
<td>47</td>
</tr>
<tr>
<td>$0.78$</td>
<td>25</td>
</tr>
<tr>
<td>$0.88$</td>
<td>22</td>
</tr>
<tr>
<td>$0.99$</td>
<td>16</td>
</tr>
<tr>
<td>$1.09$</td>
<td>13</td>
</tr>
</tbody>
</table>

Aggregate effect of decreasing the linear underpricing payoff term by 20% on the number of IPOs in the post-dot com era. Size is in $100 million units.

We conclude that intervention on reducing underpricing is a more effective approach to increasing the frequency and proceeds from IPOs than is decreasing spreads. Although high spreads transfer wealth from potentially productive firms to collusive underwriters, the spread partially
aligns the incentives of the underwriter and the firm when the underwriter sets the price of the issue. More generally, our results can be interpreted as favoring a requirement that at least some fraction of IPO shares be allocated in a nondiscriminatory fashion, reducing the incentives for underpricing and thus raising the level of IPO activity and the proceeds raised. Notably, this policy proposal suggests moving to a system more in line with the optimal IPO mechanism derived in Biais et al. (2002), despite the significant modeling differences between the two papers.

4.7 Empirical distribution of spreads
Our model suggests that proceeds should primarily determine the level of spreads. Tables 10 and 11 present some summary statistics confirming that this is true. The first panel summarizes the frequency of different

---

Table 8
Aggregate Effect of Underpricing Policy: Proceeds Pre-Dot Com Era

<table>
<thead>
<tr>
<th>Size</th>
<th>5%</th>
<th>10%</th>
<th>15%</th>
<th>20%</th>
<th>25%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\theta_1$</td>
<td>$0.8\theta_1$</td>
<td>$\theta_1$</td>
<td>$0.8\theta_1$</td>
<td>$\theta_1$</td>
</tr>
<tr>
<td>0.22</td>
<td>62.1</td>
<td>75.1</td>
<td>0.24</td>
<td>63.9</td>
<td>81.3</td>
</tr>
<tr>
<td>0.28</td>
<td>70.3</td>
<td>89.4</td>
<td>0.31</td>
<td>72.1</td>
<td>92.1</td>
</tr>
<tr>
<td>0.35</td>
<td>67.2</td>
<td>86.2</td>
<td>0.39</td>
<td>70.8</td>
<td>90.6</td>
</tr>
<tr>
<td>0.42</td>
<td>57.8</td>
<td>74.</td>
<td>0.46</td>
<td>58.4</td>
<td>74.7</td>
</tr>
<tr>
<td>0.49</td>
<td>61.</td>
<td>75.8</td>
<td>0.54</td>
<td>61.6</td>
<td>78.7</td>
</tr>
<tr>
<td>0.56</td>
<td>62.6</td>
<td>73.7</td>
<td>0.61</td>
<td>62.3</td>
<td>79.6</td>
</tr>
<tr>
<td>0.63</td>
<td>51.3</td>
<td>56.</td>
<td>0.69</td>
<td>52.5</td>
<td>67.1</td>
</tr>
<tr>
<td>0.69</td>
<td>43.6</td>
<td>47.2</td>
<td>0.76</td>
<td>46.8</td>
<td>60.1</td>
</tr>
</tbody>
</table>

Aggregate effect on proceeds from decreasing the linear underpricing payoff term by 20% during the pre-dot com era.

Table 9
Aggregate Effect of Underpricing Policy: Proceeds Post-Dot Com Era

<table>
<thead>
<tr>
<th>Size</th>
<th>5%</th>
<th>10%</th>
<th>15%</th>
<th>20%</th>
<th>25%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\theta_1$</td>
<td>$0.8\theta_1$</td>
<td>$\theta_1$</td>
<td>$0.8\theta_1$</td>
<td>$\theta_1$</td>
</tr>
<tr>
<td>0.36</td>
<td>18.7</td>
<td>24.</td>
<td>0.39</td>
<td>18.2</td>
<td>23.2</td>
</tr>
<tr>
<td>0.46</td>
<td>19.3</td>
<td>24.4</td>
<td>0.5</td>
<td>20.4</td>
<td>26.1</td>
</tr>
<tr>
<td>0.57</td>
<td>30.3</td>
<td>35.4</td>
<td>0.62</td>
<td>29.9</td>
<td>38.2</td>
</tr>
<tr>
<td>0.67</td>
<td>33.3</td>
<td>36.2</td>
<td>0.73</td>
<td>32.5</td>
<td>41.6</td>
</tr>
<tr>
<td>0.78</td>
<td>21.5</td>
<td>23.2</td>
<td>0.85</td>
<td>24.8</td>
<td>30.5</td>
</tr>
<tr>
<td>0.88</td>
<td>22.4</td>
<td>24.</td>
<td>0.96</td>
<td>21.1</td>
<td>23.7</td>
</tr>
<tr>
<td>0.99</td>
<td>18.8</td>
<td>20.1</td>
<td>1.08</td>
<td>18.6</td>
<td>19.8</td>
</tr>
<tr>
<td>1.09</td>
<td>17.4</td>
<td>18.5</td>
<td>1.19</td>
<td>18.4</td>
<td>19.6</td>
</tr>
</tbody>
</table>

Aggregate effect on proceeds from decreasing the linear underpricing payoff term by 20% during the post-dot com era.
types of spreads over the entire period we study, whereas the second panel shows data for the more recent period, after 1995, where concentration on 7% was higher. Note that the fact that 7% is more common than 10% even for the range of proceeds below $20 million is not inconsistent with our model; the threshold for a 10% spread offer is on the signal of the underwriter, and an underwriter receiving an incorrect, high signal calling for an offer of 7% will win the IPO. For those firms that conform to the 10%/7% schedule, proceeds play by far the largest role in determining the spread. No firm with proceeds above $22.4 million paid a 10% spread.

An alternative explanation for concentration at two spreads would be that certain types of firms are assigned one spread and others receive another spread. Although the above data indicate that this factor is unlikely to be driving the pattern of spreads, it is still possible if certain types of firms almost always have small IPOs. Table 12 presents evidence against this alternative. Here, we separate firms by Fama-French industry classification (at the 12 industry level) and report the frequency of each type of spread (below 7%, 7%, between 7% and 10%, and 10%). All industries are represented in each spread category, and with the exception of industry 8 (utilities, which has few observations), the average proceeds

<table>
<thead>
<tr>
<th>Spread</th>
<th>Mean</th>
<th>Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>&lt;$20M</td>
</tr>
<tr>
<td>Below 7%</td>
<td>$193M</td>
<td>175 (6.87%)</td>
</tr>
<tr>
<td>7%</td>
<td>$45M</td>
<td>988 (38.8%)</td>
</tr>
<tr>
<td>Between 7% and 10%</td>
<td>$13M</td>
<td>852 (33.5%)</td>
</tr>
<tr>
<td>10%</td>
<td>$5M</td>
<td>532 (20.9%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spread</th>
<th>Mean</th>
<th>Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>&lt;$20M</td>
</tr>
<tr>
<td>Below 7%</td>
<td>$296M</td>
<td>44 (5.29%)</td>
</tr>
<tr>
<td>7%</td>
<td>$53M</td>
<td>467 (56.1%)</td>
</tr>
<tr>
<td>Between 7% and 10%</td>
<td>$15M</td>
<td>167 (20.1%)</td>
</tr>
<tr>
<td>10%</td>
<td>$6M</td>
<td>152 (18.3%)</td>
</tr>
</tbody>
</table>
within each spread category are similar. All industries show the pattern that a higher spread category is associated with lower proceeds. In the Internet Appendix, we present a multinomial logit estimation confirming that the amount of proceeds is the primary determinant of whether the spread is 10% or 7%.

Another thing to note from the above summary statistics is that some subset of IPOs do not conform to our posited 10%/7% equilibrium. Whereas about 90% of IPOs in the later part of the sample with proceeds between $20 and $100 million receive a 7% spread, above and below these cutoffs, and particularly above $100 million, there is variation in spreads. Above $100 million, firms tend to pay spreads below 7%. In this range higher proceeds imply lower spreads,\(^{22}\) which is consistent with the

---

\(^{22}\) Internet Appendix Table 9 presents a linear regression of gross spreads on log proceeds for IPOs with proceeds above $100 million and confirms the negative relationship between proceeds and spreads.
“price war during booms” effect in Rotemberg and Saloner (1986). Specifically, when underwriters are not perfectly patient, the very largest firms will provide such high profits to underwriters at 7% that deviations will become too profitable. For simplicity, we do not model this effect and thus exclude firms with proceeds in excess of $100 million from our main empirical analysis; including such firms would potentially permit identification of the effective discount factor used by underwriters but at the cost of significantly complicating the model. Excluding these large IPOs and the dynamics driving the spread increase are potential sources of error in our estimates because the optimal spread for medium-size firms would in principal take into account the effect of this choice on incentive compatibility for larger IPOs. We expect this source of error to be relatively small.

The second set of IPOs that do not conform to the 10%/7% equilibrium are smaller IPOs that pay a spread between 7% and 10%. Although these spreads do not fit in our model, there are three natural explanations that would require only minor modifications to the model. First, these deviations could represent off-equilibrium path deviations from the spread schedule, which should trigger punishment periods. We do not observe any periods in which spreads are consistently below 7%; if anything, the concentration on 7% is increasing over time, suggesting that this is not a likely explanation. Second, there may be some public signal for a subset of firms that identifies them as firms with little uncertainty about value, which might then make collusion on a flexible spread for this subset of firms more efficient. Finally, and we think most plausibly, there could be some public or private signal about certain firms that suggests that competition for these IPOs will be limited. If competition is limited and the equilibrium spread offer should be 7%, an underwriter bidding for the IPO would consider the possibility that no other underwriter would compete for the IPO; if this probability is high enough, the underwriter would prefer to charge a spread higher than 7%. An optimal equilibrium of this metamodel would not punish such a deviation. Empirically, we find that smaller firms (by market capitalization) and younger firms are more likely to receive spread offers between 7% and 10% versus a 7% offer, which could plausibly support the claim that IPOs with fewer potential underwriters might face elevated spreads closer to what a monopolist would charge.

We also note that there is a relatively small number of firms with very low proceeds who receive spreads below 7%. These firms tend to be older and larger; the average market capitalization for a firm with proceeds below $22.4 million and a spread below 7% is $187 million, whereas the average market capitalization for firms receiving 7% or above in this proceeds range is $48 million. The existence of these firms suggests that certain types of firms may have greater bargaining power with
underwriters, and that the collusive equilibrium admits a small degree of flexibility in order to profit from the IPOs of these exceptional firms.

4.8 Further empirical implications and directions for future work

Perhaps the most natural extension of our work would be to estimate a model of spreads and underpricing for the largest IPOs, meaning those with proceeds above $100 million. In the context of our model, these IPOs would provide such high profits at a 7% spread that the assumption that observable deviations to lower spreads would be successfully deterred through punishments in the repeated game would be violated, as in Rotemberg and Saloner (1986). This violation requires that profits, and therefore spreads, be below the optimally collusive level in this range. Extending the model in this way could provide a means to identify the effective discount factor used by underwriters in the collusive equilibrium, which may be an object of independent interest. Further, comparing the estimates of parameters from a model of spread and underpricing behavior above $100 million to those from the data we use could serve to validate or call into question our estimates because the change in equilibrium behavior should not necessarily imply a change in deep structural parameters in this context.

Another interesting direction for further work would be to estimate models of the SEO market and the corporate bond market, both for initial and for seasoned public debt issues. Both of these markets involve a substantially similar set of players with the possibility of collusion. Spreads, fees, and underpricing each behave differently in each of these markets in ways that can be potentially linked to the parameters governing the qualitative nature of optimal collusion. For example, spreads on SEOs, where the prior signal about the value of the issue is already public, do not exhibit the rigidity observed for IPO spreads, which is consistent with our model. Determining whether collusion or competition characterizes behavior across all of these markets and comparing the estimates of related parameters would contribute greatly to understanding the industrial organization of investment banking.

Our model also suggests a framework for further empirical analysis of the value of going public to firms and of the potential moral hazard in the IPO decision market. Specifically, our model suggests that IPO underpricing should be predictable; information about firms that is informative about their preferences for IPOs but would not be precise enough for coordinating on different spread offers should influence the level of underpricing observed. Bradley and Jordan (2002) and others find general support for the idea that underpricing is ex ante predictable, which is broadly in line with our model. Under our model, there should be a link between predictable differences in underpricing and the cross-sectional
differences in the productivity of IPOs, which could be explored in a structural fashion. For example, firms with limited access to bank capital should face higher levels of underpricing because the value of entering the public market should be higher, whereas firms with tighter governance structures should see less underpricing because the manager is likely to be more focused on maximizing the proceeds of the issue.

5. Conclusion

We have developed a model of IPO spreads and issue pricing that can explain the puzzling concentration of spreads and the high average level of underpricing of IPOs. We establish theoretically that simple models of perfect competition and perfect monopoly cannot explain the pricing of IPO services, but a model of optimal collusion with imperfect information can. We estimate the model and obtain additional empirical results. We find that the costs of running IPOs must be small, that managerial preference for IPOs appear to have driven the large increase in underpricing during the dot-com bubble, that underwriters extract a significant fraction of the profits from underpricing and that a policy of inhibiting underwriters from obtaining profits from underpricing is the most effective way to improve outcomes in the IPO market.

References


