

DISCUSSION PAPER SERIES

No. 8403

**PRIVATELY VERSUS PUBLICLY
OPTIMAL SKIN IN THE GAME:
OPTIMAL MECHANISM AND
SECURITY DESIGN**

Gilles Chemla and Christopher Hennessy

*FINANCIAL ECONOMICS and
INDUSTRIAL ORGANIZATION*



Centre for Economic Policy Research

www.cepr.org

Available online at:

www.cepr.org/pubs/dps/DP8403.asp

PRIVATELY VERSUS PUBLICLY OPTIMAL SKIN IN THE GAME: OPTIMAL MECHANISM AND SECURITY DESIGN

Gilles Chemla, Imperial College, London, DRM/CNRS and CEPR
Christopher Hennessy, London Business School, ECGI and CEPR

Discussion Paper No. 8403
May 2011

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **FINANCIAL ECONOMICS and INDUSTRIAL ORGANIZATION**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Gilles Chemla and Christopher Hennessy

ABSTRACT

Privately versus Publicly Optimal Skin in the Game: Optimal Mechanism and Security Design*

We examine screening incentives, welfare and the case for mandatory skin-in-the-game. Ex ante banks can screen, using interim private information to choose retentions and structuring. Ex post speculators trade with rational hedging investors. Absent regulation, there is a separating equilibrium with voluntary retentions. If funding value is high, banks may instead originate-to-distribute (OTD), selling the entire asset in opaque form, deterring informed speculation and destroying screening incentives. Under weaker conditions, banks instead sell the asset in transparent form, using tranching to increase hedging demand, informed speculation and price informativeness. With sufficient informed speculation, transparent OTD actually creates stronger screening incentives than voluntary retentions. In all unregulated market equilibria, interim adverse selection reduces screening incentives, so mandated retentions potentially increase welfare. To induce screening via pooling, banks should be required to retain a uniform junior tranche size which decreases in informational efficiency. However, uniform retention mandates may not be optimal. To improve risk-sharing, screening can instead be induced via separating contracts by compelling banks to choose from a menu of junior tranche retention sizes. In either case, efficiency of risk-sharing is maximized by splitting marketed claims into safe senior and risky mezzanine tranches. Finally, the separating (pooling) regulatory regime generally leads to higher welfare if efficient risk-sharing (bank investment scale) is the dominant consideration, and is always optimal in informationally inefficient markets.

JEL Classification: D82, G21, G32, G38 and I51

Keywords: adverse selection, originate to distribute, screening incentives, securitization, skin in the game, speculator and uninformed investors

Gilles Chemla
Imperial College Business School
Tanaka Building
South Kensington Campus
London SW7 2AZ

Christopher Hennessy
London Business School
Regent's Park
London NW1 4SA

Email: g.chemla@imperial.ac.uk

Email: chennessy@london.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=127839

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=158862

* We thank Sudipto Bhattacharya, Michel Habib, Sebastien Pouget, and Jean Charles Rochet for helpful feedback. We also thank seminar participants at the University of Zurich, University of Paris-Dauphine, and Queen Mary. An early draft was circulated under a different title.

Submitted 09 May 2011

In the aftermath of the subprime crisis, empirical researchers have searched through the rubble to locate root causes of high default rates on the loans underpinning residential mortgage backed securities and collateralized debt obligations. Securitization and the advent of the originate-to-distribute (OTD) business model have figured prominently in the list of causal factors. There is now a compelling body of empirical evidence establishing a positive relationship between securitization rates and default rates (see e.g. Keys et al. (2010) and Keys, Seru and Vig (2011)). This empirical evidence raises a natural question that has not been addressed: *What, if anything, should be done?*

Implicit in recent legislation in the United States is the view that government intervention in asset-backed securities (ABS) markets will increase social welfare. In particular, the Dodd-Frank Act, recently codified as Section 15G of the Securities Exchange Act, charges six federal agencies (the Federal Reserve, Treasury, FDIC, SEC, FHA, and HUD) with setting mandatory retention standards for ABS securitizers. Unfortunately, the formulation of optimal regulation of ABS is hindered by absence of a coherent theoretical framework allowing one to answer some fundamental questions. First, what are the market failures and, more importantly, can a regulator improve upon unregulated market outcomes when facing the same informational asymmetries confronting other agents? After all, while all agree mistakes were made in the past, it is unclear whether a government mandated retention scheme can improve upon what chastened free market investors can achieve via privately optimal contracting. Second, how do mandated retentions affect determinants of social welfare such as screening incentives, bank investment levels and risk-sharing? Third, given the trade-offs what are the central policy options and conditions under which each dominates? For example, what tranches should the originator hold? And should mandated retentions be uniform or should originators be offered a menu of retention options?

This paper develops a tractable mechanism and security design framework for analyzing the regulation of ABS markets. At the origination stage, there is a moral hazard problem in that the originator can either privately shirk or exert costly effort screening borrowers in order to increase the probability of a high payoff. At the interim stage, with the asset in-hand, the originator chooses

retentions and ABS security design. There is an adverse selection problem at this stage since the originator has private information regarding the probability of the asset delivering a high payoff (the “type” below). The originator is motivated to securitize a portion of the asset, despite this adverse selection problem, because he has a valuable scalable investment opportunity. Ex post, ABS are traded in competitive markets by an endogenously informed speculator and rational uninformed investors who have a hedging demand for exposure to the underlying asset. Thus, securitization structure influences risk-sharing across investors ex post.

In order to understand the merits, if any, of a mandatory retention scheme, we first characterize equilibria in unregulated markets. At the interim stage, privately informed issuers may implement the least-cost separating equilibrium (LCSE) in which the low type securitizes the entire asset while the high type signals by retaining the minimal junior tranche needed to deter mimicry. Ex post, the LCSE has the attractive feature that socially efficient risk-sharing is achieved across investors since they are symmetrically informed. However, in terms of ex ante incentives, the LCSE is problematic since the low type receives his first-best interim payoff. Although this deters mimicry at the interim stage, it mitigates screening incentives by rewarding shirking.

If originators place extremely high value on funding, the set of interim equilibria also includes the low type’s preferred pooling equilibrium in which the entire asset is securitized in opaque form. In this case, there is no possibility for information acquisition by speculators and prices are uncorrelated with true asset value. This equilibrium is attractive from an ex post perspective since there is socially efficient risk-sharing given the absence of informational asymmetries between investors. However, this equilibrium destroys ex ante screening incentives since the two types get the same payoffs.

Under less restrictive conditions, the set of interim equilibria includes the high type’s preferred pooling equilibrium featuring OTD cum optimally tranced securities designed to stimulate informed trading. Here the high type finds it optimal to sell to investors safe senior debt and a mezzanine debt claim tailored to the demands of uninformed hedge traders. This increases uninformed trading volume and raises the gain to informed speculation, driving prices closer to fundamentals. Ex

post, risk-sharing is imperfect since uninformed investors scale back their purchases of mezzanine debt when facing informed trading. In this pooling equilibrium, market prices provide discipline and screening incentives may actually be stronger than in the separating equilibrium featuring voluntary retentions, but they are still below first-best.

This analysis reveals two valid arguments for government intervention in ABS markets. These arguments are best understood by placing them in the context of the canonical analyses of Maskin and Tirole (1992) and Tirole (2005). They show that if a privately informed principal is free to propose whatever menu of structures he desires, the set of equilibria in pure adverse selection settings consists of all structures Pareto dominating the LCSE *from the perspective of both originator types*. This characterization applies to the continuation game starting at the *interim-stage* of our model. The first problem with such equilibria is that the welfare of investors is not taken into account. Second, and more importantly, in all unregulated market equilibria, adverse selection reduces the spread between the interim payoffs to high and low types, which mitigates incentives for originators to screen.

A socially optimal mandatory retention scheme can increase screening incentives by increasing the spread between interim payoffs to high and low types, and does so in a way that accounts for costs imposed on investors as well as originators. Suppose first the social planner wants to induce screening by way of interim-stage separating equilibria. Here originators should be forced to choose from a *menu* of junior tranche retention sizes, with the high (low) type finding it optimal to hold the relatively large (small) tranche size. The optimality of using junior tranche retention size as the basis for sorting is a consequence of single-crossing, as distinct from a standard moral hazard argument calling for the agent to be residual claimant (see e.g. Innes (1990)). As in the LCSE, ex post efficient risk-sharing across investors is achieved by the separating regulatory scheme. However, unlike the LCSE, the low type is forced to retain a stake in order to increase the spread between high and low type interim payoffs and encourage screening ex ante. Significantly, unregulated markets could never implement this outcome since it is (interim) Pareto dominated from the perspective of

both originator types.

Alternatively, the social planner can induce screening via pooling equilibria. In this case, originators should be forced to hold a single junior tranche size. Intuitively, increasing the size of the marketed claim in low states reduces underinvestment and also relaxes the incentive constraint by increasing the gap between the interim payoffs of high and low types. Further, the size of the required retention should be larger when speculator effort and price informativeness are expected to be low. The disadvantage of the pooling regulatory scheme is that it entails costly speculator effort and distortions in risk-sharing across investors ex post. However, this scheme imposes lower underinvestment costs on originators when speculator effort and price informativeness are high.

The model delivers a rich set of policy prescriptions regarding how to motivate screening. First, originators should be forced to hold *junior* tranches. Second, when discretion is granted to originators, it should be over the *size* of the junior tranche, as distinct from proposals granting originators discretion over which tranches to hold. Third, in contrast to standard signaling models, optimal separating mechanisms impose underinvestment costs on even the lowest type in order to motivate screening. Fourth, the choice between separating versus pooling regimes generally trades off improved risk-sharing in the former against higher bank investment in the latter. Finally, regulation should vary according to the informational efficiency of the specific ABS market under review. If informational efficiency is high, it can be optimal to require originators to hold a small junior claim, relying on prices for discipline, consistent with the findings of Keys, Seru and Vig (2011) that market discipline appeared operative for full documentation loans but not low documentation loans. If informational efficiency is low, originators should be forced to choose from a granular menu of junior tranche retention sizes.

Our paper is most closely related to work by Gorton and Pennacchi (1995), Parlour and Plantin (2008), Plantin (2010), and Rajan, Seru and Vig (2010) who analyze the link between securitization and screening incentives. There are a number of important differences. Most importantly, these papers do not analyze the social welfare arguments for and against mandatory retentions. Second,

they do not analyze optimal security design from either a public or private perspective. Third, they abstract from the possibility of informed speculation and the role it can play in increasing screening incentives even if the entire asset is securitized. Finally, these papers abstract from the effect of ABS on risk-sharing by investors.

With its focus on the social welfare implications of ABS, our paper is also related to that of Dang, Gorton and Holmström (2010). However, their model abstracts from the ex ante moral hazard problem that is central to our analysis. In their setting, an opaque securities market is optimal since it improves risk-sharing and prevents costly information acquisition.¹ Our model accounts for the social cost of opacity in terms of inducing lender laxity. Further, in their economy optimal security design discourages informed speculation. In our model, informed speculation is a substitute for originator retentions as an ex ante motivator.

Boot and Thakor (1993) analyze security design in a pure hidden information setting with no screening. They show tranching can stimulate speculator effort, but focus on a different lever. In their model, tranching relaxes speculator wealth constraints as they trade against pure noise traders. Fulghieri and Larkin (2001) also analyze the role of security design in a setting with pure noise traders. DeMarzo and Duffie (1999) analyze optimal security design by a principal in advance of his acquiring private information, with debt being an optimal security to minimize price impact of future sales. There is no moral hazard in their model.

In the model of Gorton and Pennacchi (1990), uninformed investors carve out riskless debt in order to furnish themselves with a safe storage technology. Hennessy and Chemla (2011) show that a privately informed bank would not issue a safe claim in such a setting, relying on uninformed trade in risky debt to promote informed trading. Both papers abstract from screening incentives. Further, the theory of security design derived in the present paper differs fundamentally in that here uninformed investors trade in the mezzanine debt tranche in order to hedge endowments negatively correlated with asset payoffs.

Hanson and Sunderam (2010) also analyze the link between security design and information

¹See also Pagano and Volpin (2010) for a model of tradeoffs associated with primary market opacity.

acquisition. In their model, debt claims have low informational-sensitivity during good times and speculators do not acquire information. The ABS market freezes in bad times since investors did not invest in slow-moving information systems during the boom. Screening incentives are not analyzed.

Shleifer and Vishny (2010) present a behavioral model in which securitization volume increases in response to price exceeding fundamentals. Although their model is silent on screening incentives, it is worth pointing out that overoptimism in securities markets need not destroy screening incentives. For example, proportional overvaluation of all securities actually increases screening incentives by increasing the absolute wedge between high and low type payoffs.

The role of price informativeness in alleviating moral hazard has been analyzed in other contexts. Holmström and Tirole (1993) present a model in which the equity float affects information acquisition, price informativeness, and managerial risk premia. Maug (1998), Aghion, Bolton and Tirole (2004) and Faure-Grimaud and Gromb (2004) show that price informativeness stemming from speculator monitoring promotes insider effort. Only Aghion, Bolton and Tirole analyze security design. Each of these papers assumes pure noise trading, precluding the use of tranching to increase uninformed demand and the gains to informed speculation. Social welfare analysis is impossible in such noise trader models, and there is no analysis of mandatory retentions.

The remainder of the paper is as follows. Section I describes the full game and timing. Section II analyzes the final continuation game in which market-makers set prices. Section III analyzes the penultimate continuation game in which the privately informed originator chooses retentions and security design. Section IV compares screening incentives across pooling and separating equilibria in unregulated markets. Section V pins down the privately optimal tranching structure from the perspective of high types. Section VI analyzes social welfare and socially optimal mandatory retention schemes. We conclude with detailed policy recommendations.

I. The Game

This section describes the structure of the Full Game, which itself features two proper continuation games. The equilibrium concept is perfect Bayesian equilibrium (PBE). A PBE demands: all agents have a belief at each information set; strategies must be sequentially rational given beliefs; and beliefs are determined using Bayes' rule and the equilibrium strategies for all information sets on the equilibrium path.

A. Screening Technology and Agent Preferences

Although stylized, the model's timing assumptions correspond to the real-world timing of mortgage securitization, as described in Keys, Seru and Vig (2011). In reality, a borrower first applies for a loan. The originator then screens by processing both hard and soft information. The originator then grants a loan or not. Originators will typically inventory the loan for a period of four to six months until they eventually securitize the loan using some preferred method. The key features of this market from our perspective are that the hidden effort (screening) stage occurs *ex ante*, before the interim contracting stage in which retentions and security designs are established, with hidden information being an important friction at that interim stage.

In the model, there are four periods, 0, 1, 2 and 3, with the Full Game starting at $t = 0$. In the last two periods, there is a single storable consumption good. Originator is time and risk-neutral, deriving utility from consumption equal to $C_2 + C_3$. At the start of period 0, Nature draws the non-pecuniary screening cost \tilde{c} from a set $\{0, c, \infty\}$ with each point in the set having respective probability $\pi_{\tilde{c}} \in (0, 1)$. The realized value of \tilde{c} is privately observed by Originator. Next, Originator decides whether to pay the non-pecuniary effort cost in order to screen prospective borrowers.

Screening is not observable but screening decisions are correctly inferred in equilibrium. Screening increases the quality of the underlying asset, e.g. a loan or pool of loans. If Originator screens, the asset delivers H with probability \bar{q} and L with probability $1 - \bar{q}$. If he does not screen, the asset instead delivers H with probability \underline{q} and L with probability $1 - \underline{q}$. The cash flow generated by the underlying asset accrues at time 3. It is assumed that $0 < \underline{q} < \bar{q} < 1$ and $L \in (0, H)$. If the realized

screening cost is zero (infinite), the originator will (not) screen, so the focus is on the incentive to screen if realized $\tilde{c} = c$. We assume that in the absence of any financial market imperfection, screening would be profitable:

$$A1 : c < (\bar{q} - \underline{q})(H - L).$$

At the beginning of period 1, Originator has produced the asset, say a set of loans, and can now go to the ABS market. At this interim-stage, he has private knowledge of the true asset type τ , with *type* $\tau \in \{\underline{q}, \bar{q}\}$ always denoting the probability of the payoff H . In a PBE, investors correctly infer the effort decision of an originator facing cost c and hence correctly infer the probability of the originator being type τ . Letting ρ denote the uninformed prior probability of the type being \bar{q} , $\rho = \pi_0 + \pi_c$ if screening is incentive compatible when $\tilde{c} = c$. Otherwise $\rho = \pi_0$.

The originator is motivated raise external funds due to the fact that he attaches value $\beta > 1$ to each unit of funding he raises in securities markets at time 2. To fix ideas, throughout we treat the originator as having access to a new scalable investment with an expected payoff of β units at time 3 per unit invested at time 2. Alternatively, one can think of β as measuring the shadow value of liquidity for a bank facing capital adequacy tests. As in Parlour and Plantin (2008) and Plantin (2010), $\beta > 1$ encourages securitization.

Agents receive endowments and consume at dates 2 and 3. They can store their date 2 endowment in order to carry resources to date 3, or they can buy securities sold by the originator. As in Allen and Gale (1988), limitations on the verifiability of endowments leads to endogenously incomplete markets. In particular, the endowments of the various agents are not verifiable by courts. Consequently, other agents cannot issue securities, borrow or short-sell. However, as in Maug (1998), results would not change if we allowed short-selling by the informed speculator.

The only verifiable quantity is the cash flow generated by the underlying asset originally owned by Originator. In contrast, courts cannot verify the payoff on the new investment opportunity. This distinguishes our securitization model from a model of optimal financial structure for a corporation, since the latter features claims written on both assets in place and growth options. In our setting, the

originator sells claims on the underlying asset in order to increase the scale of the new investment. In this respect, the model resembles the security design model of DeMarzo and Duffie (1999).

There is a continuum of uninformed investors (UI) having measure one. UI may have an insurance motive for purchasing securities delivering consumption at date 3. The UI are sufficiently wealthy in aggregate to buy the entire asset since each has a date 2 endowment $y_2^{ui} \geq H$.

UI face a common endowment shock, with their date 3 endowment y_3^{ui} being either ζ or $\zeta - \phi > 0$, where $\phi > 0$ is a key parameter capturing the size of negative endowment shocks. Just prior to securities market trading in period 2, UI privately observe a signal regarding their date 3 endowment. In particular, with probability one-half they find that they are “invulnerable” to a negative shock and know they will have endowment $y_3^{ui} = \zeta$. With probability one-half they realize they are “vulnerable” to a negative shock. As is standard in models with rational hedge trading, the endowment shock is negatively correlated with the value of the underlying asset, creating intrinsic demand. Conditional upon being vulnerable, if the underlying asset delivers H the date 3 endowment of the UI will be $\zeta - \phi$ and if the underlying delivers L the date 3 endowment will be ζ . Throughout, it is assumed that $H - \phi > L$ so that in terms of aggregate output, the high realized asset payoff is good news

In fact, this setup is a reduced-form representation of hedging clientele in housing markets. To see this, consider agents who invest in assets at date 2 in order to buy a (necessity) house and a consumption good at date 3. For these agents, mortgage defaults in the final period are good news since such an outcome would allow them to buy houses cheaply. Conversely, low defaults are associated with higher house prices. One can then think of ϕ as capturing the negative shock to non-housing consumption hitting prospective home buyers if defaults are low. Long positions in CDOs help such agents hedge. Of course, this is only one potential source of hedging demand.

UI are risk-neutral over date 2 consumption and risk-averse over date 3 consumption, with ζ being a critical threshold for final period consumption. They are indexed by the intensity of their risk-aversion as captured by a preference parameter θ . The utility function of an uninformed investor

of type θ is:

$$U(C_2, C_3; \theta) \equiv C_2 + \theta \min\{C_3 - \zeta, 0\}. \quad (1)$$

The preference parameters have support $\Theta \equiv [1, \infty)$ and density f with cumulative density F . This distribution has no atoms, with f being strictly positive. This tractable specification of risk-aversion is also used by Dow (1998). Other smooth utility functions could be assumed at the cost of more complex aggregate demand functions. The essential assumption is that UI are risk averse over final period consumption, so they are potentially willing to buy claims issued by Originator despite facing adverse selection in securities markets.

There is a risk-neutral speculator S with utility $C_2 + C_3$. At date 2 she is endowed $y_2^s \geq H$ units of the numeraire, so she can afford to buy the entire asset. Her final period endowment is normalized at zero. The speculator is unique in that she receives a noisy signal of asset type and can exert costly effort to increase signal precision. Letting $s \in \{\underline{s}, \bar{s}\}$ denote the signal and τ the true asset type, S chooses $\sigma \equiv \Pr(s = \tau)$ from the feasible set $[1/2, 1]$. Her non-pecuniary effort cost function e is strictly positive, strictly increasing, strictly convex, twice continuously differentiable, and satisfies

$$\begin{aligned} \lim_{\sigma \downarrow \frac{1}{2}} e(\sigma) &= 0 \\ \lim_{\sigma \downarrow \frac{1}{2}} e'(\sigma) &= 0 \\ \lim_{\sigma \uparrow 1} e'(\sigma) &= \infty. \end{aligned}$$

If S exerts effort, the signal becomes informative since the signal causes her to correctly revise upwards her estimate of the true type being \bar{q} , with

$$\sigma > \frac{1}{2} \Rightarrow \Pr[\tau = \bar{q} | s = \bar{s}] = \frac{\Pr[\tau = \bar{q} \cap s = \bar{s}]}{\Pr[s = \bar{s}]} = \frac{\rho\sigma}{\rho\sigma + (1-\rho)(1-\sigma)} > \rho. \quad (2)$$

The final set of agents in the economy is a continuum of deep-pocketed market-makers having measure one. They are risk-neutral and have utility equal to $C_2 + C_3$.

B. The Securitization Game

Since Originator is privately informed regarding the type of the underlying asset at the start of period 1, his choice of retention and security design is a signaling game. Maskin and Tirole (1992) and Tirole (2005) show the equilibrium set of such games can be narrowed and Pareto-improved (from the perspective of the privately informed originator) by altering the fundraising mechanism in a way that expands the set of feasible initial actions. We modify their approach to fit our game.

Absent government regulation, to be considered in Section VI, the sequence of events is as follows. The continuation game starting at date 1 is called the *Securitization Game*. It is a signaling game played between Originator and outside investors. This game begins with Originator informing the other agents of the *menu* of two securitization structures, say $\Sigma \in \{\Sigma_i, \Sigma_j\}$, that he will choose from subsequently, e.g. a shelf-registration. Each structure stipulates payoffs for claimants as a function of the verifiable asset payoff at date 3. All agents must have a belief regarding the asset type in response to any menu announcement, including those off the equilibrium path. Thus, the menu offered is itself a potentially informative signal. Next, the originator selects a securitization structure Σ from the menu he just proposed, with the choice being incentive compatible. After observing Originator's choice, all agents revise beliefs using Bayes' rule where possible.

It is worth stressing that both types can offer the same menu, but they do not necessarily select the same securitization structure. In a *separating equilibrium*, the initial menu is such that the securitization structure chosen by the originator reveals the true asset type. In a *pooling equilibrium*, both originator types propose the same trivial menu with $\Sigma_i = \Sigma_j$. In pooling equilibria, no information about the type can be inferred based upon the securitization structure chosen by the originator.

Anticipating, Section VI considers whether there is a role for regulators in restricting the set of allowable securitization structures.

C. The Market-Making Game

At date 2, play passes to a second continuation game, labeled the *Market-Making Game*.² In this

²This game is similar to that presented in Maug (1998), but we have endogenous security design and UI demand.

game, securities prices are set competitively by the market-makers facing the potentially-informed speculator. In a separating equilibrium, the true type has been revealed, so prices are set equal to fundamental value, and the speculator does not exert effort.

Consider next any pooling equilibrium. In this case, the Market-Making Game is a signaling game played between the informed speculator and market-makers. Here, the market-makers and uninformed investors enter the market-making game holding their prior belief that the asset type is \bar{q} with probability ρ . The market-making game starts with S choosing σ at personal cost $e(\sigma)$. Her equilibrium choice σ^{eq} is not publicly observable, but is correctly inferred by other agents in equilibrium. Then S privately observes her signal regarding the asset type.

Next, UI privately observe whether or not they are vulnerable to a negative endowment shock at date 3. Market orders are then submitted, with market-makers setting prices competitively. The market-making process is in the spirit of Kyle (1985) and Glosten and Milgrom (1985). The UI and S simultaneously submit non-negative market orders. Market-makers then set prices based upon observed aggregate demands in all markets, with no market segmentation. Market-makers clear all markets, buying all securities not purchased by UI or S .

Since Originator is the only agent capable of issuing claims delivering goods in period 3, market-makers cannot be called upon to take short positions. To this end, we impose a second technical assumption.

$$A2 : \phi \leq \frac{H - L - c/(\bar{q} - \underline{q})}{2} \Rightarrow c \leq (\bar{q} - \underline{q})[H - L - 2\phi].$$

The role of Assumption 2 is as follows. The aggregate demand of UI is weakly increasing in ϕ . To avoid the possibility of aggregate demand exceeding supply for any security, the endowment shock must be sufficiently small. Assumption 2 also implies that screening effort is socially profitable, even after accounting for potential losses incurred by UI when the realized asset payoff is H .

Figure 1 provides a review of the time-line.

II. Equilibrium of Market-Making Game

We solve for the set of PBE of the Full Game via backward induction. Consider first the Market-Making Game. This game is trivial in a separating equilibrium since in that case all securities are priced at fundamental value given the revealed type. Therefore, below attention is confined to pricing in pooling equilibria—where both issuer types pool at the same trivial menu and structure. We initially abstract from optimal security design and retentions, considering first the case in which the originator sells the entire asset in the form of a pass-through security. This simplifies the initial analysis and also allows us to isolate the effect of tranching in Section V. As a short-hand, *originate-to-distribute (or OTD)* denotes any securitization structure in which the issuer retains zero claim on terminal cash flows.

Since she cannot short, the optimal strategy for Speculator is to place a buy order if and only if she receives a positive signal regarding type $\tau \in \{\underline{q}, \bar{q}\}$. She attempts hiding her orders behind the UI orders. The optimal size of her order is equal to the aggregate buy order of UI when vulnerable to negative endowment shocks. This latter quantity is denoted X and is determined endogenously.

UI do not place orders if invulnerable to negative shocks since marginal utility of any increase in C_3 is then zero. An individual UI may place an order if vulnerable since there is a hedging motive. However, each UI weighs adverse selection costs against hedging motives in choosing demand.

Table 1 lists the possible aggregate demand (D) configurations confronting market-makers. Using Bayes' rule, market-makers form the following beliefs

$$\begin{aligned} \Pr[\tau = \bar{q} | D = 2X] &= \frac{\rho\sigma}{1 - \rho - \sigma + 2\rho\sigma} & (3) \\ \Pr[\tau = \bar{q} | D = X] &= \rho \\ \Pr[\tau = \bar{q} | D = 0] &= \frac{\rho(1 - \sigma)}{\rho + \sigma - 2\rho\sigma}. \end{aligned}$$

Market-maker beliefs and equilibrium prices (P) increase monotonically in aggregate demand with:

$$\begin{aligned} P(D) &= \underline{q}H + (1 - \underline{q})L + (\bar{q} - \underline{q})(H - L) \Pr[\tau = \bar{q} | D] \quad \forall D \in \{0, X, 2X\} & (4) \\ \Rightarrow & P(2X) > P(X) > P(0). \end{aligned}$$

To support the PBE conjectured in Table 1 it suffices to verify Speculator has no incentive to deviate regardless of her signal. To that end, off the equilibrium path market-makers form adverse beliefs from the perspective of the speculator, setting prices based upon:

$$\Pr[s = \bar{s}|D] = 1 \quad \forall \quad D \notin \{0, X, 2X\}.$$

Speculator has no incentive to deviate from the posited signal-contingent trading strategy when confronted with such beliefs.

A. Expected Revenue of Originator

Using equation (4) and Table 1, the expected revenue of the high type is:

$$\begin{aligned} E[R|\tau = \bar{q}] &= \underline{q}H + (1 - \underline{q})L \\ &+ (\bar{q} - \underline{q})(H - L) \left[\frac{\sigma \Pr[\tau = \bar{q}|D = 2X]}{2} + \frac{(1 - \sigma) \Pr[\tau = \bar{q}|D = 0]}{2} + \frac{\Pr[\tau = \bar{q}|D = X]}{2} \right]. \end{aligned} \quad (5)$$

Equation (5) can be rewritten as:

$$\begin{aligned} E[R|\tau = \bar{q}] &= Z(\sigma, \rho)[\bar{q}H + (1 - \bar{q})L] + [1 - Z(\sigma, \rho)][\underline{q}H + (1 - \underline{q})L] \\ Z(\sigma, \rho) &\equiv \frac{1}{2} \left[\frac{\rho\sigma^2}{1 - \rho - \sigma + 2\rho\sigma} + \frac{\rho(1 - \sigma)^2}{\rho + \sigma - 2\rho\sigma} + \rho \right]. \end{aligned} \quad (6)$$

The variable Z plays a critical role, measuring the high type's expectation of the market-makers' updated belief. The appendix shows:

$$\begin{aligned} Z_\sigma(\sigma, \rho) &> 0 \\ Z\left(\frac{1}{2}\right) &= \rho \\ Z(1) &= \frac{1 + \rho}{2}. \end{aligned}$$

Intuitively, high types expect favorable revisions in market-maker beliefs if the speculator gets more precise signals. Notice also that if the speculator does not exert effort, prices revert back to uninformed expected asset value.

Using equation (4) and Table 1, the expected revenue of the low type is:

$$E[R|\tau = \underline{q}] = z(\sigma, \rho)[\bar{q}H + (1 - \bar{q})L] + [1 - z(\sigma, \rho)][\underline{q}H + (1 - \underline{q})L] \quad (7)$$

$$z(\sigma, \rho) \equiv \left(\frac{\rho}{1 - \rho} \right) [1 - Z(\sigma, \rho)].$$

Lemma 1, which follows from the fact that Z is increasing in σ , shows the high type benefits from increased speculator signal precision, since this drives prices closer to fundamentals.

Lemma 1 *The expected revenue of the originator of a high value asset is increasing in the precision of the signal received by the speculator.*

B. Incentive Compatible Speculator Signal Precision

Consider next the incentives of Speculator. From Table 1 it follows that her expected gross trading gain is

$$G(\sigma, \rho, X) = X \cdot \left[\begin{aligned} & \left(\frac{\rho\sigma}{2} \right) [\bar{q}H + (1 - \bar{q})L - P(2X)] + \left(\frac{\rho\sigma}{2} \right) [\bar{q}H + (1 - \bar{q})L - P(X)] \\ & + \left(\frac{(1-\rho)(1-\sigma)}{2} \right) [\underline{q}H + (1 - \underline{q})L - P(2X)] + \left(\frac{(1-\rho)(1-\sigma)}{2} \right) [\underline{q}H + (1 - \underline{q})L - P(X)] \end{aligned} \right].$$

It is readily verified that the first and third terms in the expression for G cancel, as one would expect given that equilibrium order flow fully reveals the speculator's private signal L when demand is $2X$. Thus, all speculator gains arise from the non-revealing aggregate order flow X . A bit of algebra reveals:

$$G(\sigma, \rho, X) = X \cdot \left[\frac{\rho(1 - \rho)(2\sigma - 1)(\bar{q} - \underline{q})(H - L)}{2} \right].$$

The incentive compatible signal precision (σ_{ic}) solves:

$$\sigma_{ic} \in \arg \max_{\sigma \in [1/2, 1]} X \cdot \left[\frac{\rho(1 - \rho)(2\sigma - 1)(\bar{q} - \underline{q})(H - L)}{2} \right] - e(\sigma). \quad (8)$$

The speculator properly treats the aggregate demand of vulnerable UI (X) as fixed when considering alternative levels of signal precision, since her choice of signal precision is not observable to UI, implying they cannot change demand in response to changes in σ . It follows that

$$e'(\sigma_{ic}) = \rho(1 - \rho)(\bar{q} - \underline{q})(H - L)X. \quad (9)$$

We summarize these results as follows:

Lemma 2 *The incentive compatible signal precision of the speculator is*

$$\sigma_{ic}(X) = [e']^{-1}[\rho(1 - \rho)(\bar{q} - \underline{q})(H - L)X].$$

C. Uninformed Demand for the Pass-Through Security

The next step in pinning down the equilibrium of the Market-Making Game is to determine aggregate UI demand (X) for the pass-through security in response to their being vulnerable to a negative endowment shock, recalling UI have zero demand if invulnerable. Letting $x^*(\cdot)$ denote the optimal θ -contingent demand, aggregate uninformed demand is

$$X \equiv \int_1^\infty x^*(\theta)f(\theta)d\theta. \quad (10)$$

Each UI has measure zero and acts as a price-taker. Each vulnerable UI expects the security to be overpriced since he knows a subset of vulnerable UI will submit positive orders, pushing prices higher as market-makers revise upward their assessment of the probability of the asset being of high value. Despite facing adverse selection, an individual UI is willing to submit a buy order if θ is sufficiently high.

In order to characterize UI demand, it is useful to compute the expected price of the security conditional upon the UI being vulnerable. This conditional expectation is:

$$\begin{aligned} E[P|\text{Vulnerable}] &= [\rho\sigma + (1 - \rho)(1 - \sigma)]P(2X) + [\rho(1 - \sigma) + \sigma(1 - \rho)]P(X) \\ &= L + (H - L)[\rho\bar{q} + (1 - \rho)\underline{q}] + \rho(1 - \rho)(2\sigma - 1)(\bar{q} - \underline{q})(H - L). \end{aligned} \quad (11)$$

Equation (11) shows vulnerable UI face adverse selection when submitting buy orders, since the asset is overpriced relative to its fundamental value. There is no adverse selection if the speculator fails to exert effort ($\sigma = 1/2$). Adverse selection increases in the speculator's signal precision.

Consider now the optimal portfolio for an individual vulnerable UI. He will not invest his endowment in the safe storage technology since this is an inefficient means of insuring given that he actually prefers an Arrow security paying off in the event that the realized payoff on the underlying asset is H . We now let x denote the number of units of the pass-through security purchased, and

confine attention to $x \leq \phi/H$ since x in excess of that cutoff has zero marginal utility in the final period. The objective of each vulnerable UI is to maximize his expected utility, implying

$$x^*(\theta) \in \arg \max_{0 \leq x \leq \phi/H} y_2^{ui} - xE[P|\text{Vulnerable}] - [\rho\bar{q} + (1 - \rho)\underline{q}]\theta[\phi - xH]. \quad (12)$$

It follows that:

$$\begin{aligned} x^*(\theta) &= 0 \quad \forall \quad \theta \in [1, \hat{\theta}) \\ x^*(\theta) &= \frac{\phi}{H} \quad \forall \quad \theta \in [\hat{\theta}, \infty) \\ \hat{\theta} &\equiv \frac{E[P|\text{Vulnerable}]}{[\rho\bar{q} + (1 - \rho)\underline{q}]H} \\ &= 1 + \frac{\rho(1 - \rho)(2\sigma - 1)(\bar{q} - \underline{q})}{\rho\bar{q} + (1 - \rho)\underline{q}} \\ &\quad + \frac{L}{H} \left[\frac{1 - \rho(1 - \rho)(2\sigma - 1)(\bar{q} - \underline{q})}{\rho\bar{q} + (1 - \rho)\underline{q}} - 1 \right]. \end{aligned} \quad (13)$$

The aggregate buy order of vulnerable UI decreases in the speculator's signal precision since $\hat{\theta}$ increases in σ .

Lemma 3 summarizes the properties of aggregate uninformed demand.

Lemma 3 *If types pool by selling the underlying asset in the form of a pass-through security, aggregate uninformed demand when vulnerable to a negative endowment shock is*

$$X(\sigma) = \frac{\phi}{H} \cdot \left[1 - F(\hat{\theta}(\sigma)) \right].$$

Figure 2 sketches the proof for existence of a unique equilibrium in the Market-Making Game. To understand the figure, recall the incentive compatible signal precision from Lemma 2 and the uninformed demand from Lemma 3. We know $X(1/2) > 0$ and that X is strictly decreasing in σ , yet strictly positive for all $\sigma \in [1/2, 1]$. Plotting the incentive compatible signal precision, we know σ_{ic} is strictly increasing in X with $\sigma_{ic}^{-1}(1/2) = 0$, with the limit as $\tilde{\sigma}$ converges to one of $\sigma_{ic}^{-1}(\tilde{\sigma}) = \infty$. Thus, the two curves intersect once only, implying uniqueness.

Proposition 1 *In the Market-Making Game following pooling, there exists a unique equilibrium pair (σ^{eq}, X^{eq}) satisfying*

$$\begin{aligned}\sigma_{ic}(X^{eq}) &= \sigma^{eq} \in \left(\frac{1}{2}, 1\right) \\ X(\sigma^{eq}) &= X^{eq} \in \left(0, \frac{\phi}{H}\right).\end{aligned}$$

III. Equilibrium of the Securitization Game

Continuing the backward induction, suppose play has reached the Securitization Game. At the start of this game, all players have correctly inferred the probability ρ of the asset having type \bar{q} . However, Originator is privately informed about the true type. This sets up a signaling game in which Originator can either separate (by keeping sufficient skin-in-the-game) or pool by selling the entire underlying asset (originate-to-distribute) in the form of a pass-through security.

We begin by evaluating the least-cost separating equilibrium (LCSE). The LCSE minimizes the low type's incentive to mimic by giving him his first-best allocation in which he sells the entire asset to outside investors. An originator with positive information signals by retaining sufficient cash flow rights. In terms of signal credibility and originator payoffs, all that matters is originator retentions. Thus, we assume for now that all marketed cash flows are packaged into some bundle B , paying outside investors a total amount B_i for each possible final asset payoff $i \in \{L, H\}$. When analyzing the risk-sharing properties of alternative equilibria, we consider the socially optimal bifurcation of this bundle.

The LCSE makes the high type as well off as possible subject to the constraint that the low type does not mimic, as well as standard limited-liability constraints. The program is:

$$\max_{(B_L, B_H)} \quad \bar{q}(H - B_H) + (1 - \bar{q})(L - B_L) + \beta[\bar{q}B_H + (1 - \bar{q})B_L] \quad (14)$$

subject to

$$\begin{aligned}\beta[\underline{q}B_H + (1 - \underline{q})B_L] &\geq \underline{q}(H - B_H) + (1 - \underline{q})(L - B_L) + \beta[\bar{q}B_H + (1 - \bar{q})B_L] \\ B_L &\leq L; \quad B_H \leq H; \quad B_L \geq 0; \quad B_H \geq 0.\end{aligned}$$

To pin down the LCSE we solve a relaxed program that ignores the last three constraints, and then verify the neglected constraints are slack. With this in mind, first note the nonmimicry constraint must bind in the solution to the relaxed program since otherwise the objective function could be increased by raising B_H by an infinitesimal amount. From the nonmimicry constraint, B_H can be expressed as a function of B_L :

$$B_H(B_L) = B_L + \frac{(\beta - 1)[\underline{q}H + (1 - \underline{q})L - B_L]}{\beta\bar{q} - \underline{q}}.$$

Substituting $B_H(B_L)$ into the objective function and ignoring constants, the relaxed program can now be expressed as:

$$\max_{B_L \leq L} \quad \bar{q}B_H(B_L) + (1 - \bar{q})B_L.$$

The objective function is strictly increasing in B_L , implying $B_L = L$ is optimal. Substituting this value into $B_H(B_L)$ and verifying the neglected constraints are indeed slack, it follows that the LCSE entails:

$$(B_L^{lcse}, B_H^{lcse}) = \left(L, L + \frac{(\beta - 1)\underline{q}(H - L)}{\beta\bar{q} - \underline{q}} \right).$$

Since we are interested in social welfare, as well as the originator's payoff, it is sensible to consider bifurcating the marketed claims to (B_L^{lcse}, B_H^{lcse}) to maximize social welfare. With this in mind, note that in the LCSE the welfare of the originator, speculator and market-makers are unaffected by any bifurcation of marketed claims. However, UI are made strictly better off by carving out their desired Arrow security paying off iff the asset payoff is H . The following Proposition summarizes this LCSE with Pareto-improving bifurcations of marketed cash flows.

Proposition 2 *In the least-cost separating equilibrium of the Securitization Game, the entire asset is sold if type \underline{q} , being bifurcated into a senior claim with face L and a junior claim with face $H - L$. If the type is \bar{q} , investors are sold a senior claim with face L and a mezzanine claim with face $(\beta - 1)\underline{q}(H - L)/(\beta\bar{q} - \underline{q})$. The originator retains the residual equity tranche.*

The next lemma, due to Maskin and Tirole (1992), shows the LCSE is always a candidate

equilibrium of the Securitization Game. Further, the LCSE payoffs represent a lower bound on each type's payoff in any PBE.

Lemma 4 *The equilibrium set of the Securitization Game always includes the least-cost separating equilibrium. It also includes pooling equilibria provided the pooling contract weakly Pareto dominates the least-cost separating equilibrium from the perspective of both originator types.*

Given the importance of the LCSE payoffs in pinning down the equilibrium set, it is worth specifying them. Substituting the LCSE policies into originator objective functions yields:

$$\begin{aligned}\tau &= \underline{q} \Rightarrow U_{lcse}^O = \beta[\underline{q}H + (1 - \underline{q})L] \\ \tau &= \bar{q} \Rightarrow U_{lcse}^O = \beta[\bar{q}H + (1 - \bar{q})L] - (\beta - 1) \left[\frac{\bar{q}\beta(\bar{q} - \underline{q})(H - L)}{(\beta\bar{q} - \underline{q})} \right].\end{aligned}\tag{15}$$

As shown above, in the LCSE the low type receives his first best payoff, a standard result. The high type receives his first best payoff minus the foregone NPV due to retentions. To see this, note that the foregone NPV is proportional to the value of retained equity tranche, which is captured by the last bracketed term in equation (15). The notion that signaling motives lead to retentions and underinvestment was first discussed by Leland and Pyle (1977) and Myers and Majluf (1984).

IV. Originate-to-Distribute versus Least Cost Separating Equilibrium

This section compares screening incentives in the LCSE with *candidate* pooling equilibria in which the entire asset is sold as a pass-through security. Such pooling equilibria represent one form of the originate-to-distribute (OTD) business model. To facilitate comparison with Dang, Gorton and Holmström (2010), and understand the effects of opacity, two variants of the OTD business model are considered: transparent OTD and opaque OTD. Our baseline modeling framework, in which the set of feasible σ is $[1/2, 1]$, captures a *transparent* securities market in which the speculator can acquire informative signals. If originators instead choose *opacity*, the speculator cannot possibly acquire informative signals, implying $\sigma = 1/2$ and $Z = z = \rho$. For example, one can think of low

documentation and coarse disclosure as forms of opacity, since they make it hard for investors to make informed inferences about fundamental value.

A. Is Originate-to-Distribute an Equilibrium Business Model?

Lemma 4 established pooling at OTD is a potential equilibrium of the Securitization Game provided both originator types are weakly better off than in the LCSE. The low type is strictly better off with pooling at OTD since he sells the entire asset, just as in the LCSE, but benefits from overpricing. From equations (6) and (15) it follows the high type is better off pooling at OTD if his loss from underpricing (left side of equation below) is less than his LCSE loss stemming from underinvestment (right side):

$$\beta(1 - Z)(\bar{q} - \underline{q})(H - L) \leq (\beta - 1) \left[\frac{\bar{q}\beta(\bar{q} - \underline{q})(H - L)}{(\beta\bar{q} - \underline{q})} \right]. \quad (16)$$

Rearranging terms in the equation above, we arrive at the following characterization.

Lemma 5 *There exists a pooling equilibrium of the Securitization Game in which the originator uses the originate-to-distribute business model and sells opaque securities iff*

$$\rho \geq \frac{\bar{q} - \underline{q}}{\beta\bar{q} - \underline{q}} \Leftrightarrow \beta \geq 1 + \left[1 - \frac{\underline{q}}{\bar{q}} \right] \left[\frac{1 - \rho}{\rho} \right]. \quad (17)$$

There exists a pooling equilibrium in which the originator uses the originate-to-distribute business model and sells transparent securities iff

$$Z[\sigma^{eq}(\rho), \rho] \geq \frac{\bar{q} - \underline{q}}{\beta\bar{q} - \underline{q}} \Leftrightarrow \beta \geq 1 + \left[1 - \frac{\underline{q}}{\bar{q}} \right] \left[\frac{1 - Z[\sigma^{eq}(\rho), \rho]}{Z[\sigma^{eq}(\rho), \rho]} \right].$$

The intuition for Lemma 5 is as follows. The condition for pooling at opaque OTD is similar to those derived by Parlour and Plantin (2008) and Plantin (2010). High types are willing to pool at OTD, cum opacity, provided the percentage of high types is sufficiently large and provided banks place sufficient value on immediate funding. From Proposition 1 it follows that in a transparent securities market the speculator exerts effort, $\sigma^{eq} > 1/2$, and market-makers set prices closer to fundamentals, with $Z > \rho \geq \pi_0$. This increases high type payoffs in pooling equilibria, so *pooling*

at OTD cum transparent securities is possible under weaker conditions than pooling at OTD cum opaque securities. Formally, we see this effect as implied by the fact that the second condition for pooling is less stringent than the first.

B. Screening Incentives and Equilibria

We now evaluate the conventional wisdom that the OTD business model necessarily destroys screening incentives. We focus on the incentives of the originator facing cost c , since only he faces a nontrivial screening decision. Three simplifying assumptions are maintained in this subsection (only). First, for each business model considered, screening is assumed to occur if there exists at least one concomitant equilibrium in which it is incentive compatible. Second, under the transparent OTD model it is assumed the equilibrium value of $Z - z$ is larger for $\rho = \pi_0 + \pi_c$ than for $\rho = \pi_0$. Finally, it is assumed that the first condition in Lemma 5 is satisfied so that OTD can always be supported as an equilibrium of the Securitization Game.

Consider first incentives to screen if the LCSE will be played at the interim-stage. Screening is incentive compatible if the cost is less than the difference between Securitization Game payoffs for high and low types, or:

$$\begin{aligned}
 c &\leq \left[\beta[\bar{q}H + (1 - \bar{q})L] - (\beta - 1) \left(\frac{\bar{q}\beta(\bar{q} - \underline{q})(H - L)}{(\beta\bar{q} - \underline{q})} \right) \right] - \beta[\underline{q}H + (1 - \underline{q})L] & (18) \\
 &\Downarrow \\
 c &\leq \hat{c}_{lcse} \equiv [\beta(\bar{q} - \underline{q})(H - L)] \left[\frac{\bar{q} - \underline{q}}{\beta\bar{q} - \underline{q}} \right].
 \end{aligned}$$

Here \hat{c}_{lcse} denotes the highest screening cost the originator is willing to pay ex ante holding the conjecture that LCSE will be played at the interim-stage. The first bracketed term in the expression for \hat{c}_{lcse} is the cutoff cost that would obtain if there were symmetric information regarding asset type. The second bracketed term is a number less than one, implying that interim-stage asymmetric information diminishes screening incentives relative to first-best whenever the LCSE is played. Intuitively, in the LCSE asymmetric information reduces the incentive to screen since the high type underinvests to signal quality.

Consider next the incentive to screen if the originator anticipates interim-stage pooling at OTD cum transparent securities. Screening is incentive compatible if the screening cost is less than β times the expected revenue differential between the high and low types. From equations (6) and (7) it follows that screening is incentive compatible iff:

$$\begin{aligned}
c &\leq \beta[E(R|\tau = \bar{q}) - E(R|\tau = \underline{q})] & (19) \\
&\Downarrow \\
c &\leq \hat{c}_{otdt} \equiv \beta[Z(\sigma^{eq}(\rho), \rho) - z(\sigma^{eq}(\rho), \rho)](\bar{q} - \underline{q})(H - L).
\end{aligned}$$

Condition (19) is central to our analysis. Here \hat{c}_{otdt} denotes the highest screening cost the originator is willing to pay given a conjecture that interim-stage play entails pooling at OTD cum transparent securities. The product of the last two terms in the expression for \hat{c}_{otdt} captures the pure gain to screening in terms of expected cash flow. Under OTD, this entire benefit is scaled up by the factor β , since the entire asset is sold. This total screening benefit is scaled by the difference in expected market-maker beliefs for the two types ($Z - z$). The difference is always less than one in our noisy rational expectations equilibrium, so once again interim-stage asymmetric information reduces ex ante incentives. However, increases in speculator effort (σ) drive prices closer to fundamentals (increase $Z - z$) and strengthen screening incentives.

With opaque securities, the speculator cannot acquire information and market-makers use the uninformed prior ($Z = z = \rho$) in setting prices. Following the same steps as in the transparent OTD case, it now follows that originators will not pay screening costs if they anticipate interim-stage pooling at OTD cum opacity. Further, from condition (17) it follows that endogenously low effort makes pooling at opaque OTD less viable as an equilibrium since the high type probability is low, making pooling unattractive for a high type.

The next proposition follows directly from a comparison of the screening cost cutoffs \hat{c}_{lcese} and \hat{c}_{otdt} . Critically, the proposition shows that OTD actually provides stronger screening incentives if the originator places sufficiently high value on funding and the speculator exerts sufficient effort.

Proposition 3 *The originator is willing to incur positive screening costs under the originate-to-distribute business model if types pool at transparent securities, with willingness-to-pay strictly increasing in speculator effort. If $\beta \in (1, 2 - \underline{q}/\bar{q}]$, the originator is willing to pay higher screening costs in the least-cost separating equilibrium. If $\beta > 2 - \underline{q}/\bar{q}$, the originator is willing to pay higher screening costs under originate-to-distribute if speculator effort exceeds $\hat{\sigma}$, where $\hat{\sigma}$ solves:*

$$Z(\hat{\sigma}, \pi_0 + \pi_c) - z(\hat{\sigma}, \pi_0 + \pi_c) = \frac{\bar{q} - \underline{q}}{\beta\bar{q} - \underline{q}}.$$

The empirical evidence presented by Keys, Seru and Vig (2011) is consistent with Proposition 3. In particular, it is reasonable to assume that markets for ABS backed by low documentation loans had low informational efficiency. Here one would expect a sharp reduction in screening incentives for securitized loans. In contrast, the market for ABS backed by full documentation loans was likely to achieve a higher degree of informational efficiency, since outside investors had more hard information available to them. Here one would expect originators to have stronger screening incentives for securitized loans given stronger market discipline.

We turn now to discuss potential equilibria in the Full Game. This analysis allows us to examine two interesting questions. A first natural question is whether the OTD business model could possibly be a rational expectations equilibrium—especially if market prices for CDOs are weakly correlated with true value. Second, we are interested in whether there can be self-fulfilling prophecies in securitization markets, with the expectation of (no) screening inducing (no) screening.

We have the following proposition.

Proposition 4 *There is a least-cost separating equilibrium in which the bank screens with probability π_0 if $c > \hat{c}_{lcs}$ and with probability $\pi_0 + \pi_c$ if $c \leq \hat{c}_{lcs}$. If*

$$\beta \geq 1 + \left[1 - \frac{\underline{q}}{\bar{q}} \right] \left[\frac{1 - Z[\sigma^{eq}(\pi_0 + \pi_c), \pi_0 + \pi_c]}{Z[\sigma^{eq}(\pi_0 + \pi_c), \pi_0 + \pi_c]} \right]$$

and

$$c \leq \beta [Z(\sigma^{eq}(\pi_0 + \pi_c), \pi_0 + \pi_c) - z(\sigma^{eq}(\pi_0 + \pi_c), \pi_0 + \pi_c)] (\bar{q} - \underline{q}) (H - L)$$

there is a pooling equilibrium in which the bank originates-to-distribute using transparent securities and screens with probability $\pi_0 + \pi_c$. If

$$\beta \geq 1 + \left[1 - \frac{\underline{q}}{\bar{q}}\right] \left[\frac{1 - Z[\sigma^{eq}(\pi_0), \pi_0]}{Z[\sigma^{eq}(\pi_0), \pi_0]}\right]$$

and

$$c > \beta[Z(\sigma^{eq}(\pi_0), \pi_0) - z(\sigma^{eq}(\pi_0), \pi_0)](\bar{q} - \underline{q})(H - L)$$

there is an equilibrium in which the bank originates-to-distribute using transparent securities and screens with probability π_0 . If

$$\beta \geq 1 + \left[1 - \frac{\underline{q}}{\bar{q}}\right] \left[\frac{1 - \pi_0}{\pi_0}\right],$$

there is an equilibrium in which the bank uses the originate-to-distribute business model cum opaque securities and screens with probability π_0 .

Proposition 4 can be understood as follows. The first equilibrium described follows immediately from the fact that the LCSE is always a interim-stage equilibrium, so it is always possible to base an equilibrium of the Full Game on the LCSE. Next, the proposition shows that screening and no screening can be self-fulfilling prophecies. To see this, consider the second and third equilibria described. In the second case, if the market expects originators to screen, then the informational efficiency of prices will be high and originators will indeed be justified in screening. Conversely, in the third case, if the market does not expect lenders to screen, then the informational efficiency of prices will be low and originators will indeed be justified in not screening. Therefore, another interpretation of the apparent collapse in lending standards is that the market was simply “trapped in the bad equilibrium.” Finally, the proposition shows that if funding value is extremely high, the OTD business model can be supported as an equilibrium outcome even when there is no speculator effort and market prices are completely uninformative ($\sigma = 1/2$ and $Z - z = 0$). This is the key argument made by Plantin (2010).

V. Privately Optimal Retentions and Tranching

To simplify the analysis, we have thus far confined attention to a specific pooling equilibrium, one in which the entire asset is sold in the form of a pass-through security. A natural question to ask is whether more informative prices, and hence stronger screening incentives, can be generated by different structurings. For example, is tranching privately optimal? Further, one might also be interested in identifying the high type's preferred retention policy in a pooling equilibrium. After all, pooling results in underpricing for the high type, so he might prefer a less extreme policy than full asset sales.

To address these questions we break the analysis into two steps. First, we consider the optimal packaging, from the perspective of the high type, of an arbitrary payoff pair (l, h) , with l (h) measuring the total amount paid to outside investors if the final cash flow is L (H). Suppose the originator splits the rights to the payoff pair (l, h) into two claims, A and B , with respective cash flow-contingent payoffs (A_L, A_H) and (B_L, B_H) . We conjecture, and then verify, that all UI demand will be concentrated in one security, say security A . In this case, there is zero UI demand for B , and the speculator will not trade in the market for B given the lack of any cover provided by UI. Therefore, Table 1 continues to be the relevant table depicting aggregate demand (for security A). Since there is no market segmentation, the aggregate demand for security A is also used by the market-makers in setting prices for security B . That is, the market-makers will set prices for all securities based on aggregate demand for A as follows:

$$\begin{aligned}
 P_A(D) &= \underline{q}A_H + (1 - \underline{q})A_L + (\bar{q} - \underline{q})(A_H - A_L) \Pr[\tau = \bar{q}|D] \quad \forall \quad D_A \in \{0, X_A, 2X_A\} \quad (20) \\
 P_B(D) &= \underline{q}B_H + (1 - \underline{q})B_L + (\bar{q} - \underline{q})(B_H - B_L) \Pr[\tau = \bar{q}|D] \quad \forall \quad D_A \in \{0, X_A, 2X_A\}.
 \end{aligned}$$

Using Table 1 one arrives at the following expressions for the expected prices computed by UI when vulnerable to negative endowment shocks:

$$\begin{aligned}
 E[P_A|\text{Vulnerable}] &= [\rho\sigma + (1 - \rho)(1 - \sigma)]P_A(2X) + [\rho(1 - \sigma) + \sigma(1 - \rho)]P_A(X) \quad (21) \\
 &= A_L + (A_H - A_L)[\rho\bar{q} + (1 - \rho)\underline{q}] + \rho(1 - \rho)(2\sigma - 1)(\bar{q} - \underline{q})(A_H - A_L).
 \end{aligned}$$

$$\begin{aligned}
E[P_B|\text{Vulnerable}] &= [\rho\sigma + (1-\rho)(1-\sigma)]P_B(2X) + [\rho(1-\sigma) + \sigma(1-\rho)]P_B(X) \\
&= B_L + (B_H - B_L)[\rho\bar{q} + (1-\rho)\underline{q}] + \rho(1-\rho)(2\sigma-1)(\bar{q}-\underline{q})(B_H - B_L).
\end{aligned} \tag{22}$$

In order for the UI to prefer security A to B , as has been conjectured, it must be the case that it provides higher expected marginal utility per unit spent. Recalling that UI only have positive marginal utility if the realized cash flow is H , they prefer security A to B iff:

$$\frac{A_H}{E[P_A|\text{Vulnerable}]} \geq \frac{B_H}{E[P_B|\text{Vulnerable}]}.$$
 \tag{23}

The objective of the high type is to maximize his expected revenue (and hence $Z(\sigma)$) by maximizing the incentive compatible signal precision for the speculator, which here is the unique solution to:

$$e'(\sigma_{ic}) = \rho(1-\rho)(\bar{q}-\underline{q})(A_H - A_L)X_A.$$
 \tag{24}

In order to determine the demand of vulnerable UI for security A we repeat the same steps as those performed for the pass-through security and arrive at directly analogous demand expressions:

$$\begin{aligned}
\theta &\in [1, \hat{\theta}_A) \Rightarrow x_A^*(\theta) = 0 \\
\theta &\in [\hat{\theta}_A, \infty) \Rightarrow x_A^*(\theta) = \frac{\phi}{A_H} \\
X_A &= \frac{\phi}{A_H} \left[1 - F(\hat{\theta}_A) \right]
\end{aligned} \tag{25}$$

where

$$\begin{aligned}
\hat{\theta}_A &= 1 + \frac{\rho(1-\rho)(2\sigma-1)(\bar{q}-\underline{q})}{\rho\bar{q} + (1-\rho)\underline{q}} \\
&\quad + \frac{A_L}{A_H} \left[\frac{1 - \rho(1-\rho)(2\sigma-1)(\bar{q}-\underline{q})}{\rho\bar{q} + (1-\rho)\underline{q}} - 1 \right].
\end{aligned} \tag{26}$$

Since the high type's objective is to maximize σ_{ic} , it follows from convexity of e that his objective is to maximize the right side of equation (24) taking as fixed the value of σ entering into the UI demand function X_A (since the speculator treats UI demand as invariant to his private choice of σ). Ignoring constants, we have the following program:

$$\begin{aligned}
&\max_{A_L, A_H} \left[1 - \frac{A_L}{A_H} \right] \left[1 - F(\hat{\theta}_A) \right] \\
&s.t. \quad A_L \leq l, \quad A_H \leq h, \quad \text{and} \quad (23).
\end{aligned}$$

To characterize the optimal structuring we solve a relaxed program ignoring the first and last constraints. Since the objective function is strictly decreasing in A_L it is optimal to set $A_L = 0$. With $A_L = 0$, any positive A_H suffices, say $A_H = h - l$. This implies claim B is a safe claim with face value l . Substituting this optimal structuring back into the maximand we see that the value attained is actually invariant to (l, h) . That is, the same σ_{ic} and Z value, call it Z^* , are attained for all (l, h) pairs provided they are bifurcated into safe debt and an Arrow claim paying off only if H is realized.

Thus, an interesting feature of the model is that it predicts originators have an incentive to increase asset span, as argued by Allen and Gale (1988). However, the motive here is an indirect one. In our model, originators cater to the preferences of UI since uninformed trading provides a subsidy to the speculator. In turn, informed speculation drives prices closer to fundamentals, as desired by high types.

Having characterized optimal tranching of the payoff pair (l, h) , we turn next to the high type's preferred retention policy if issuers pool using such tranching policies. The optimal retention policy for the high type then solves:

$$\max_{l \leq L, h \leq H} \bar{q}(H - h) + (1 - \bar{q})(L - l) + \beta\{Z^*[\bar{q}H + (1 - \bar{q})L] + [1 - Z^*][\underline{q}H + (1 - \underline{q})L]\}. \quad (27)$$

The objective function is strictly increasing in both arguments provided the conditions stated in Proposition (5) for the existence of a pooling equilibrium are satisfied. With this in mind, we have the following proposition.

Proposition 5 *The privately optimal interim-stage pooling contract for the high type is to retain zero interest in the asset, with speculator incentives for information acquisition maximized via bifurcation of the asset into a safe senior claim with face L and a junior claim with face $H - L$.*

The preceding proposition shows that tranching could actually help to bring about OTD cum screening when it would not be possible if the originator were restricted to selling only a pass-through security. In particular, the proposition shows that tranching leads to strictly higher speculator

effort (σ) and informational efficiency ($Z - z$), making it more likely that screening will be incentive compatible for the originator since \hat{c}_{otdt} increases. The underlying causal mechanism is simple. Tranching creates Arrow securities demanded by uninformed hedgers. In turn, the increase in uninformed trading volume allows the speculator to place larger buy orders and make larger profits. This increases speculator effort and results in more informative prices. In turn, this encourages screening.

VI. Social Welfare Analysis and Optimal Mandatory Retentions

Up to now attention has been confined to privately optimal securitization structures, focusing on the relative screening incentives generated by pooling at OTD versus the LCSE with voluntary retentions. There are two questions of more general interest. First, in the absence of regulation, what securitization models are socially preferred. This question has garnered increasing interest recently, with Dang, Gorton and Holmström (2010) arguing that opacity is socially optimal. Pagano and Volpin (2010) instead argue there is a tradeoff in that primary market opacity can lead to secondary market illiquidity as investors may acquire and trade on private information in secondary markets. The second question we address is whether there is scope for government regulation to increase welfare by requiring banks maintain skin-in-the-game.

A. Welfare in Unregulated Market Equilibria

This section takes the perspective of a utilitarian social planner placing equal weight on each agent's expected utility. For brevity, market-makers are ignored since their expected utility is always the same. We begin by comparing the equilibria described in Proposition 4: the LCSE featuring voluntary retentions by high types; pooling at the high type's preferred OTD structure cum tranching; and pooling at the low type's preferred OTD structure featuring opacity. Recall, the key feature of opacity is that it prevents informed speculation. To place opacity on its strongest possible footing in terms of social welfare, we assume that under the opaque OTD business model

marketed cash flows are bifurcated into safe debt and junior equity, since it is here Pareto improving to carve out the Arrow claim demanded by UI.

We first evaluate the social costs and benefits of opacity relative to transparency. To this end, let $\rho_{otdt} \in \{\pi_0, \pi_0 + \pi_c\}$ denote the equilibrium high type probability under pooling at the transparent OTD model, with π_0 being the equilibrium high type probability under the opaque OTD model. Adding the expected utility of the originator, the speculator and UI, social welfare under the opaque OTD business model is:

$$W_{otdo} = \beta[L + (H - L)(\pi_0\bar{q} + (1 - \pi_0)\underline{q})] + y_2^{ui} - \frac{\phi}{2}[\pi_0\bar{q} + (1 - \pi_0)\underline{q}] + y_2^s. \quad (28)$$

Under the opaque OTD business model, perfect risk-sharing is achieved with all UI buying enough of the Arrow claim (junior tranche) to insure against losing $\theta\phi$ in the event of a negative endowment shock. Additionally, the speculator cannot exert costly effort.

Adding the expected utilities, and noting that trading gains and losses cancel, social welfare under the transparent OTD business model is:

$$\begin{aligned} W_{otdt} &= \beta[L + (\rho_{otdt}\bar{q} + (1 - \rho_{otdt})\underline{q})(H - L)] - (\rho_{otdt} - \pi_0)c + y_2^s - e(\sigma^{eq}) \\ &\quad + y_2^{ui} - \frac{\phi}{2}[\rho_{otdt}\bar{q} + (1 - \rho_{otdt})\underline{q}] \left[1 + \int_1^{\hat{\theta}_A^*} (\theta - 1)f(\theta)d\theta \right] \\ \hat{\theta}_A^* &\equiv 1 + \frac{\rho(1 - \rho)(2\sigma - 1)(\bar{q} - \underline{q})}{\rho\bar{q} + (1 - \rho)\underline{q}}. \end{aligned} \quad (29)$$

The implied social gain to opacity is:

$$\begin{aligned} W_{otdo} - W_{otdt} &= e(\sigma^{eq}) + \frac{\phi}{2}[\rho_{otdt}\bar{q} + (1 - \rho_{otdt})\underline{q}] \left[\int_1^{\hat{\theta}_A^*} (\theta - 1)f(\theta)d\theta \right] \\ &\quad - (\rho_{otdt} - \pi_0) \left[(\bar{q} - \underline{q}) \left(\beta(H - L) - \frac{\phi}{2} \right) - c \right]. \end{aligned} \quad (30)$$

The first line above captures two social costs of transparency: the cost of information production by the speculator and the distortion in risk-sharing arising from asymmetric information across investors. In our model, the latter effect manifests itself as a subset of UI foregoing insurance against negative endowment shocks, despite the existence of gains from trade given that $\theta > 1$. In

this way, our model captures the two costs of transparency stressed by Dang, Gorton and Holmström (2010). The second line, which is weakly negative, captures the social value of additional screening incentives created by transparency.

Consider next the LCSE. Here the speculator simply consumes his endowment, achieving utility y_2^s . As shown in Proposition 2, the vulnerable UI are insulated from adverse selection since the private information of the originator has been signaled. Thus, each vulnerable UI effectively purchases a fairly priced Arrow claim paying ϕ if the realized cash flow is H . This allows them to avoid incurring the cost θ of a consumption shortfall in the final period. Adding these terms to the ex ante utility of the originator, accounting for underinvestment by the high type, we find:

$$W_{lcse} = \beta L + \beta(H - L)[\rho_{lcse}\bar{q} + (1 - \rho_{lcse})\underline{q}] - \rho_{lcse}\bar{q}(\beta - 1) \left[\frac{\beta(\bar{q} - \underline{q})(H - L)}{\beta\bar{q} - \underline{q}} \right] - (\rho_{lcse} - \pi_0)c + y_2^s + y_2^{ui} - [\rho_{lcse}\bar{q} + (1 - \rho_{lcse})\underline{q}] \frac{\phi}{2}. \quad (31)$$

This implies the following welfare gain to opacity over the LCSE:

$$W_{otdo} - W_{lcse} = \rho_{lcse}\beta \left[\frac{\bar{q}(\beta - 1)(\bar{q} - \underline{q})(H - L)}{\beta\bar{q} - \underline{q}} \right] - (\rho_{lcse} - \pi_0) \left[(\bar{q} - \underline{q}) \left(\beta(H - L) - \frac{\phi}{2} \right) - c \right]. \quad (32)$$

As shown above, the benefit of opaque OTD relative to the LCSE is that there is no reduction in bank investment. However, the cost of opacity is that it potentially weakens screening incentives, as captured by the second term.

Notice above that the only cost to the opaque OTD structure in terms of total social welfare is that it fails to provide any incentive for screening. This implies the following general result.

Lemma 6 *Originate-to-distribute cum opacity is socially optimal amongst all mechanisms that fail to induce originators to engage in costly screening.*

Lemma 6 is consistent with the arguments in Dang, Gorton and Holmström (2010). However, it also illustrates that opacity is optimal within a relatively narrow set—the set of structures that fail to provide screening incentives. The intuition is simple. Separating equilibria lead to investment distortions while transparent pooling equilibria lead to distortions in risk-sharing across investors

and costly information acquisition. Absent some benefit in terms of ex ante incentives, these interim-stage distortions are socially useless.

From Lemma 6 it follows that inducing costly screening is a *necessary* condition for some configuration other than opacity to be socially optimal. Therefore, the remainder of our analysis is devoted to finding socially optimal methods for inducing originators to engage in costly screening. Consistent with this argument, curbing lender laxity appears to be the primary purpose of mandatory retention legislation, such as the Dodd-Frank bill.

B. Motivating Screening: The Socially Optimal Separating Mechanism

Consider first the optimal mandatory retention scheme aimed at inducing both separation at the interim stage and screening ex ante by the originator facing cost c . From a social perspective, all such schemes yield the same expected utility for the speculator, who consumes her endowment, and the UI provided a suitable Arrow claim is carved out of marketed securities. We shall ignore this latter condition and verify it is slack at any proposed optimum.

We begin by noting that if screening is incentive compatible (IC below) under the LCSE described in Proposition 2, a regulator cannot possibly find a socially preferable separating scheme. After all, the LCSE achieves perfect risk-sharing across investors and does not lead to socially costly speculator effort. Further, the low type gets first-best and the high type gets the highest payoff consistent with nonmimicry in the LCSE. Also, we recall the LCSE is always in the set of unregulated market equilibria. However, Lemma 4 shows the LCSE is not necessarily a unique equilibrium, so a regulator would still need to constrain originators to choose between the LCSE allocations in order to be sure they do not implement some other equilibrium that dominates from an originator perspective.

We need only determine the socially optimal separating mechanism when the IC constraint is violated at the LCSE, with $c > \hat{c}_{lcse}$. To this end, let (b_L, b_H) and (B_L, B_H) denote the bundle of cash flows sold to outside investors by the low and high types, respectively. The planner's problem is to maximize expected originator utility subject to IC, self-selection by the low type (with the high type's self-selection constraint being slack), and limited liability.

We solve the following relaxed program which ignores some limited liability constraints and then verify the neglected constraints are slack:

$$\begin{aligned}
& \max_{b_L, b_H, B_L, B_H} \rho \{ \bar{q}(H - B_H) + (1 - \bar{q})(L - B_L) + \beta[\bar{q}B_H + (1 - \bar{q})B_L] \} \\
& + (1 - \rho) \{ \underline{q}(H - b_H) + (1 - \underline{q})(L - b_L) + \beta[\underline{q}b_H + (1 - \underline{q})b_L] \} \\
& s.t. \\
IC & : \bar{q}(H - B_H) + (1 - \bar{q})(L - B_L) + \beta[\bar{q}B_H + (1 - \bar{q})B_L] - c = \\
& \quad \underline{q}(H - b_H) + (1 - \underline{q})(L - b_L) + \beta[\underline{q}b_H + (1 - \underline{q})b_L] \\
NM & : \underline{q}(H - b_H) + (1 - \underline{q})(L - b_L) + \beta[\underline{q}b_H + (1 - \underline{q})b_L] \geq \\
& \quad \bar{q}(H - B_H) + (1 - \bar{q})(L - B_L) + \beta[\bar{q}B_H + (1 - \bar{q})B_L] \\
LL & : B_L \leq L, \quad b_L \leq L
\end{aligned}$$

We first observe that NM must bind in the relaxed program otherwise the objective function could be increased by raising B_H by an infinitesimal amount while still meeting all constraints. Next, we substitute the binding NM into both the objective function and IC constraint to rewrite the relaxed program as:

$$\begin{aligned}
& \max_{B_L, B_H} \rho \{ \bar{q}(H - B_H) + (1 - \bar{q})(L - B_L) + \beta[\bar{q}B_H + (1 - \bar{q})B_L] \} \quad (33) \\
& + (1 - \rho) \{ \underline{q}(H - B_H) + (1 - \underline{q})(L - B_L) + \beta[\bar{q}B_H + (1 - \bar{q})B_L] \} \\
& s.t. \\
IC' & : B_H = H - L - B_L - \frac{c}{\bar{q} - \underline{q}}. \\
LL & : B_L \leq L
\end{aligned}$$

Substituting the right side of IC' into the new objective function we find it is increasing in B_L from which it follows the socially optimal separating contract entails:

$$(B_L^*, B_H^*) = \left(L, H - \frac{c}{\bar{q} - \underline{q}} \right) = \left(L, L + \frac{(\beta - 1)\underline{q}(H - L)}{\beta\bar{q} - \underline{q}} - \frac{c - \hat{c}_{lcs\epsilon}}{\bar{q} - \underline{q}} \right) \quad (34)$$

The optimality of forcing the originator to take a junior position, with $B_L^* = L$, is a consequence of a standard single-crossing condition, as distinct from the traditional moral hazard argument that

calls for agents to be residual claimants (see e.g. Innes (1990)). To see this, consider first the following argument demonstrating $B_L^* = L$. If instead one set $B_L < L$, it would be possible to increase B_L and decrease B_H in a way that left the NM constraint unaffected, while relaxing IC and increasing the objective function. Alternatively, note that for any (b_L, b_H) pair intended for the low type, the optimal choice of (B_L, B_H) deters mimicry in the most efficient way. Now note that one may express the primitive utility of an originator having success probability q as follows:

$$u(B_L, B_H, REVENUE; q) = q(H - B_H) + (1 - q)(L - B_L) + \beta \times REVENUE.$$

Computing the utility differential du over the first two arguments and setting this to zero, the slope of the iso-utility curve for type- q is:

$$\frac{\Delta B_H}{\Delta B_L} = -\frac{1 - q}{q}. \quad (35)$$

Since the low type has a steeper indifference curve, separation is achieved by having the high type take a bundle with high B_L and low B_H .

Inspection reveals $B_H^* < B_H^{lcse}$. That is, the high type is now forced to hold a larger junior equity tranche than in the LCSE. Assumption 2 is sufficient to guarantee both $B_H^* > B_L^*$ and that the high type can carve out a sufficient supply of the Arrow security to ensure market-makers are never called upon to short. For example, (B_L^*, B_H^*) can be bifurcated into a safe claim with face value L and a mezzanine tranche with face value $B_H^* - B_L^*$.

Finally, we substitute (B_L^*, B_H^*) into the NM constraint to compute the low type's payoff under the optimal contract:

$$\tau = \underline{q} \Rightarrow U_{sep}^O = \beta[\underline{q}H + (1 - \underline{q})L] - \frac{(\beta\bar{q} - \underline{q})(c - \hat{c}_{lcse})}{\bar{q} - \underline{q}}. \quad (36)$$

Here we see the clear contrast between the interim-stage LCSE and the separating mechanism that induces screening. Recall, in the LCSE the low type receives his first-best allocation, since doing so actually relaxes the nonmimicry constraint. However, it is apparent that giving the low type such a high payoff weakens ex ante effort incentives. Above, we see that motivating screening necessitates reducing the low type's equilibrium payoff.

Any pair (b_L, b_H) giving the low type the correct utility level suffices. For example it suffices to set:

$$(b_L^*, b_H^*) = \left(L, H - \frac{(\beta\bar{q} - \underline{q})(c - \widehat{c}_{l_{cse}})}{(\beta - 1)\underline{q}(\bar{q} - \underline{q})} \right). \quad (37)$$

It is readily verified that $L < b_H^* < B_H^*$. Intuitively, the high type signals positive information by retaining a larger junior claim. Further, the two neglected limited liability constraints are satisfied.

Finally, we return to implementation. Recall, to ensure efficient risk-sharing the marketed claims of both types must include the Arrow claims demanded by the UI. To that end, the high (low) type should bifurcate marketed cash flows into a senior claim with face value L and a mezzanine claim with face $B_H^* - L$ ($b_H^* - L$), with the originator retaining a junior claim with face $H - B_H^*$ ($H - b_H^*$). Notice, in each case the mezzanine claim is sufficient to meet the demand of the UI for ϕ units of state- H consumption.

And we have established the following proposition:

Proposition 6 *The socially optimal separating mechanism for inducing screening calls for both types to sell safe senior claims with face L . The high and low types should also sell mezzanine claims with respective face values:*

$$\left[\frac{(\beta - 1)\underline{q}}{\beta\bar{q} - \underline{q}} \right] (H - L) - \frac{(c - \widehat{c}_{l_{cse}})^+}{\bar{q} - \underline{q}}$$

and

$$H - L - \left[\frac{(\beta\bar{q} - \underline{q})}{(\beta - 1)\underline{q}(\bar{q} - \underline{q})} \right] (c - \widehat{c}_{l_{cse}})^+.$$

The originator retains the residual equity tranche.

Critically, Proposition 6 shows that a uniform mandate regarding retentions is not optimal if regulators intend to implement separating equilibria. Rather, in such cases originators should be offered a menu of retention options of varying scale. Intuitively, the separating mechanism described above is designed to accomplish three distinct tasks: increasing asset span via the creation of Arrow claims; provision of ex ante screening incentives by increasing the wedge between high and low-type

payoffs; and revelation of the originator's private information. This last effect is socially valuable since it obviates the need for speculators to acquire information and insulates uninformed investors from adverse selection, facilitating efficient risk sharing.

Consider now social welfare under the effort-inducing separating contract described in Proposition 6:

$$W_{sep} = \beta L + \beta(H - L)[(\pi_0 + \pi_c)\bar{q} + (1 - \pi_0 - \pi_c)\underline{q}] - \pi_c c + y_2^s + y_2^{ui} - [(\pi_0 + \pi_c)\bar{q} + (1 - \pi_0 - \pi_c)\underline{q}]\frac{\phi}{2} \\ - (\pi_0 + \pi_c)\bar{q}(\beta - 1) \left[\frac{\beta(\bar{q} - \underline{q})(H - L)}{\beta\bar{q} - \underline{q}} \right] - (c - \widehat{c}_{lcsce})^+ \left[\frac{(\pi_0 + \pi_c)(\beta - 1)\bar{q} + (1 - \pi_0 - \pi_c)(\beta\bar{q} - \underline{q})}{\bar{q} - \underline{q}} \right]$$

The social gain to opacity over the effort-inducing separating contract is:

$$W_{otdo} - W_{lsce} = (\pi_0 + \pi_c)\bar{q}(\beta - 1) \left[\frac{\beta(\bar{q} - \underline{q})(H - L)}{\beta\bar{q} - \underline{q}} \right] \\ + (c - \widehat{c}_{lcsce})^+ \left[\frac{(\pi_0 + \pi_c)(\beta - 1)\bar{q} + (1 - \pi_0 - \pi_c)(\beta\bar{q} - \underline{q})}{\bar{q} - \underline{q}} \right] \\ - \pi_c \left[(\bar{q} - \underline{q}) \left(\beta(H - L) - \frac{\phi}{2} \right) - c \right].$$

Above we see that opacity has costs and benefits relative to the separating mechanism. The separating mechanism induces screening effort, but also imposes underinvestment costs. Thus, regulators must ultimately decide whether social value of screening incentives merits the costs in terms of lower bank investment.

C. Motivating Screening: The Socially Optimal Pooling Mechanism

Suppose instead the utilitarian social planner would like to induce screening effort using a pooling mechanism, with the privately optimal pooling contract not sufficing for this purpose, i.e. suppose $\widehat{c}_{otdt} < c$. To characterize the optimal pooling mechanism, we note that for arbitrary fixed σ , a socially optimal contract maximizes the weighted average of originator utilities, here subject to the appropriate IC constraint. Intuitively, the sum of speculator and UI utilities is decreasing in σ , so the dual problem for the planner is to maximize expected originator utility for fixed σ , Z , and z .

We solve this dual problem as follows. Again let (l, h) denote the total cash flow accruing to outside investors as a function of the final realized cash flow, to be tranching in a socially optimal

manner after the optimal values for (l, h) have been determined. For fixed σ , Z , and z we may write the low and high type's respective expected payoffs as follows:

$$\begin{aligned}\underline{U} &= \underline{q}(H - h) + (1 - \underline{q})(L - l) + \beta[z(\bar{q}h + (1 - \bar{q})l) + (1 - z)(\underline{q}h + (1 - \underline{q})l)]. \\ \bar{U} &= \bar{q}(H - h) + (1 - \bar{q})(L - l) + \beta[Z(\bar{q}h + (1 - \bar{q})l) + (1 - Z)(\underline{q}h + (1 - \underline{q})l)].\end{aligned}\quad (38)$$

The program is as follows:

$$\begin{aligned}\max_{l, h} \quad & \rho\bar{U} + (1 - \rho)\underline{U} = L + (H - L)[\rho\bar{q} + (1 - \rho)\underline{q}] + (\beta - 1)[l + (h - l)(\rho\bar{q} + (1 - \rho)\underline{q})] \\ & \text{s.t.} \\ & IC : \bar{U} - \underline{U} = (\bar{q} - \underline{q})[(H - L) - (h - l)(1 - \beta(Z - z))] \geq c \\ & LL : l \in [0, L] \text{ and } h \in [0, H].\end{aligned}$$

Under the working hypothesis that $\widehat{c}_{otdt} < c$ it must be the case that IC is binding and it must also be the case that $1 > \beta(Z - z)$. Substituting the binding IC constraint into the objective function it follows immediately that the optimal pooling contract allows the originator to market the following bundle of cash flows:

$$\begin{aligned}l^* &= L \\ h^* &= L + \frac{(\bar{q} - \underline{q})(H - L) - c}{(\bar{q} - \underline{q})[1 - \beta(Z - z)]} = H - (H - L)\frac{(c - \widehat{c}_{otdt})^+}{(\bar{q} - \underline{q})(H - L) - \widehat{c}_{otdt}}.\end{aligned}\quad (39)$$

We see that if the planner wants to induce pooling, he mandates that all originators maintain the same skin-in-the-game, holding only the junior tranche. Intuitively, increases in l allow both types to invest more and simultaneously relax the IC constraint. However, increases in h tighten the IC constraint, and must be curtailed to provide screening incentives. Also, we see that mandatory retentions and information production by the speculator are substitutes. In particular, the originator should be allowed to market more of the asset payoff (increase h^*) for higher levels of σ (and hence higher $Z - z$). Finally, if the IC constraint is slack at full securitization, then full securitization solves the dual problem.

Let us return to implementation. To increase the efficiency of risk-sharing, the marketed claims sold by originators must include the Arrow claims demanded by UI. To that end, originators should bifurcate the marketed cash flows into a senior claim with face value L and a mezzanine claim with face $h^* - L$, with the originator retaining a junior claim with face $H - h^*$. It is readily verified that $h^* - L > 2\phi$ so there is sufficient supply of the Arrow claim such that market-makers need never take short positions regardless of the equilibrium order flow.

And we have established the following proposition:

Proposition 7 *The socially optimal pooling mechanism for inducing screening calls for originators to sell a senior claim with face L and a mezzanine claim with face value:*

$$(H - L) \left[1 - \frac{(c - \widehat{c}_{otdt})^+}{(\bar{q} - \underline{q})(H - L) - \widehat{c}_{otdt}} \right].$$

The originator retains the residual equity tranche.

We are now in position to compare the merits of using the separating scheme as opposed to the pooling scheme for inducing screening. The socially optimal pooling scheme now reflects mandatory retentions, with implied social welfare:

$$\begin{aligned} W_{pool} &= \beta L + \beta(H - L)[\rho\bar{q} + (1 - \rho)\underline{q}] - \pi_c c + y_2^s - e(\sigma^{eq}) + y_2^{ui} \\ &\quad - \frac{\phi}{2}[\rho\bar{q} + (1 - \rho)\underline{q}] \left[1 + \int_1^{\widehat{\theta}_A^*} (\theta - 1)f(\theta)d\theta \right] \\ &\quad - [\rho\bar{q} + (1 - \rho)\underline{q}](\beta - 1)(H - L) \frac{(c - \widehat{c}_{otdt})^+}{(\bar{q} - \underline{q})(H - L) - \widehat{c}_{otdt}}. \end{aligned}$$

The implied welfare differential between the pooling and separating regulatory regimes is:

$$\begin{aligned} W_{pool} - W_{sep} &= \rho\beta \left[\frac{\bar{q}(\beta - 1)(\bar{q} - \underline{q})(H - L)}{\beta\bar{q} - \underline{q}} \right] + \left[\frac{\rho(\beta - 1)\bar{q} + (1 - \rho)(\beta\bar{q} - \underline{q})}{\bar{q} - \underline{q}} \right] (c - \widehat{c}_{lcse})^+ \\ &\quad - [\rho\bar{q} + (1 - \rho)\underline{q}](\beta - 1)(H - L) \frac{(c - \widehat{c}_{otdt})^+}{(\bar{q} - \underline{q})(H - L) - \widehat{c}_{otdt}} \\ &\quad - e(\sigma^{eq}) - \frac{\phi}{2}[\rho\bar{q} + (1 - \rho)\underline{q}] \left[\int_1^{\widehat{\theta}_A^*} (\theta - 1)f(\theta)d\theta \right]. \end{aligned} \tag{40}$$

The first three terms in the equation above represent the difference in underinvestment costs between the two regulatory regimes. This term is strictly positive when the informational efficiency of markets is high, so that $c - \widehat{c}_{otdt}$ is small. Further, this term is strictly positive if both IC constraints were (nearly) slack, reflecting the fact that there is always a positive amount of underinvestment required under separating contracts.

However, it is not necessarily the case that the separating regime entails higher underinvestment costs. For example, suppose $c = \widehat{c}_{lcse}$ so that the IC constraint is just slack at the LCSE. At the same time suppose the market is inherently informationally inefficient with \widehat{c}_{otdt} close to zero. In such cases, the pooling contract converges to the high type's separating contract, so that the pooling contract imposes gratuitous underinvestment costs on the low type, implying the separating contract dominates on all counts, including bank investment levels. More generally, the pooling regime is preferred from an investment perspective when the informational efficiency of markets is high, so that little/no underinvestment is needed to provide screening incentives. With pooling contracts, the market becomes the motivator.

The final two terms represent the key cost of inducing screening via pooling contracts. First, the pooling regime entails socially costly effort by speculators. Second, the pooling regime results in informational asymmetries between speculators and uninformed investors. This distorts the portfolio choice of uninformed investors, resulting in socially inefficient risk sharing.

Finally, we can evaluate the merits of opacity relative to the screening-inducing pooling scheme:

$$\begin{aligned}
W_{otdo} - W_{pool} &= e(\sigma^{eq}) + \frac{\phi}{2}[(\pi_0 + \pi_c)\bar{q} + (1 - \pi_0 - \pi_c)\underline{q}] \left[\int_1^{\widehat{\theta}^A} (\theta - 1)f(\theta)d\theta \right] \\
&+ [(\pi_0 + \pi_c)\bar{q} + (1 - \pi_0 - \pi_c)\underline{q}](\beta - 1)(H - L) \frac{(c - \widehat{c}_{otdt})^+}{(\bar{q} - \underline{q})(H - L) - \widehat{c}_{otdt}} \\
&- \pi_c \left[(\bar{q} - \underline{q}) \left(\beta(H - L) - \frac{\phi}{2} \right) - c \right].
\end{aligned} \tag{41}$$

As shown above, there are three welfare costs to the pooling scheme: speculator effort costs, distorted risk-sharing across investors, and underinvestment stemming from mandated retentions. These costs

must be weighed against the social benefits of increasing screening effort, as captured by the final term.

Conclusions and Policy Recommendations

This paper developed a tractable framework for analyzing social welfare in both regulated and unregulated ABS markets, accounting for ex ante moral hazard, interim asymmetric information between the originator and investors, and ex post asymmetric information between investors. We show that in unregulated markets, interim adverse selection reduces screening incentives, so mandated retentions have the potential to increase welfare.

The first important policy implication to emerge from the model is that originators should be required to hold the most *junior* tranche, with the underlying logic for this policy depending on the nature of the mechanism. In a pooling mechanism, marketing more of the low-state payoff increases the spread between the interim payoffs of high and low types. In a separating mechanism, a natural single-crossing condition implies it is most efficient to use the size of the junior tranche as the basis for sorting.

Second, the separating mechanism calls for giving originators discretion, but of a very specific form. In particular, originators should only have discretion regarding the size of the junior tranche retention. This is distinct from existing proposals for granting originators broad discretion over the form of skin-in-the-game. For example, our analysis indicates that it would be socially inefficient to grant originators the option of retaining a fraction of each tranche. Third, in contrast to standard signaling results, which call for low types to receive their first best allocations, our analysis shows it is optimal to impose costs on even the low type, since this serves to increase the spread between interim payoffs across types.

Fourth, the case for government intervention is weakest in informationally efficient markets. And further, in any pooling mechanism, the size of the socially optimal junior tranche retention is smaller the more informationally efficient the market. For example, this implies that securitized assets with

high documentation levels should carry smaller retention requirements than assets with low levels of documentation. Conversely, in informationally inefficient or inherently opaque markets, separating mechanisms are socially preferable.

It should also be stressed that our analysis shows that it may be both necessary and optimal to impose caps on retentions as well as floors. To see this, note that the objective of socially optimal mechanisms is to tilt the interim-stage equilibrium away from inefficient equilibria. For example, we showed that if markets are informationally inefficient, pooling equilibria impose higher underinvestment costs on both types than the optimal separating mechanism, and the separating mechanism dominates. However, absent the imposition of a cap on retentions, originators may end up playing the inefficient pooling equilibrium.

Finally, some natural conclusions, in addition to those discussed above, emerge if one moves beyond a literal interpretation of our model. In our view, many ABS markets will find their way to a semi-separating equilibrium if the number of potential types becomes large (e.g. a continuum of types). In such cases, more separation can be induced if the set of allowable junior tranche retention sizes is granular. This is preferable to a coarse menu when efficient risk-sharing is a priority and when informational efficiency is low.

Appendix: Proof of Z Increasing

From the definition of Z we have:

$$\begin{aligned}\frac{2Z}{\rho} &= 1 + \frac{\sigma^2}{(1-a)} + \frac{(1-\sigma)^2}{a} \\ a(\sigma) &\equiv \rho + \sigma - 2\sigma\rho.\end{aligned}$$

Differentiating we obtain:

$$\begin{aligned}\frac{2Z'(\sigma)}{\rho} &= \frac{2(1-a)\sigma + (1-2\rho)\sigma^2}{(1-a)^2} - \frac{2a(1-\sigma) + (1-2\rho)(1-\sigma)^2}{a^2} \\ &= \frac{[2(1-a) + (1-2\rho)\sigma]\sigma a^2 - (1-a)^2(1-\sigma)[2a + (1-2\rho)(1-\sigma)]}{(1-a)^2 a^2}\end{aligned}\tag{42}$$

This is strictly positive if and only if.

$$\begin{aligned}[2(1-a) + \sigma(1-2\rho)]\sigma a^2 &> (1-a)^2(1-\sigma)[2a + (1-2\rho) - \sigma(1-2\rho)] \\ &\Downarrow \\ [(1-a) + (1-\rho)]\sigma a^2 &> (1-a)^2(1-\sigma)[a + (1-\rho)] \\ &\Downarrow \\ (1-\rho)\sigma a^2 &> (1-a)[(1-\sigma)(1-a)a + (1-\sigma)(1-\rho)(1-a) - \sigma a^2] \\ &\Downarrow \\ (1-\rho)\sigma a^2 &> (1-a)[a(1-a) - a\sigma + (1-\sigma)(1-\rho)(1-a)] \\ &\Downarrow \\ [(1-\rho)a + 1-a]\sigma a &> (1-a)^2[a + (1-\sigma)(1-\rho)] \\ &\Downarrow\end{aligned}$$

$$\begin{aligned}
[1 - \rho a] \sigma a &> (1 - a)^2 (1 - \rho \sigma) \\
&\Downarrow \\
\sigma a - \sigma \rho a^2 &> (1 - a)^2 - \rho \sigma (1 - a)^2 \\
&\Downarrow \\
\rho \sigma [(1 - a)^2 - a^2] + \sigma a &> (1 - a)^2 \\
&\Downarrow \\
\rho \sigma + \sigma a (1 - 2\rho) &> (1 - a)^2 \\
&\Downarrow \\
a^2 + \rho(\sigma - a) &> (1 - a)^2 \\
&\Downarrow \\
\rho^2(2\sigma - 1) + 2[\sigma - \rho(2\sigma - 1)] &> 1 \\
&\Downarrow \\
(\rho - 1)^2(2\sigma - 1) - (2\sigma - 1) + 2\sigma &> 1 \\
&\Downarrow \\
(\rho - 1)^2(2\sigma - 1) &> 0. \blacksquare
\end{aligned}$$

References

- [1] Aghion, P., P. Bolton and J. Tirole, 2004, Exit Options in Corporate Finance: Liquidity versus Incentives, *Review of Finance* 8, 327-353.
- [2] Allen, F., and D. Gale, 1988, Optimal Security Design, *Review of Financial Studies*, 89, 1, 229-63.
- [3] Boot, A., and A. Thakor, 1993, Security Design, *Journal of Finance*, 48 , 1349-78.
- [4] Dang,T.V., Gorton, G., and B. Holmström, 2010, Opacity and the Optimality of Debt in Liquidity Provision, working paper, MIT.
- [5] DeMarzo, P., and D. Duffie, 1999, A Liquidity-Based Model of Security Design, *Econometrica* 67, 65-99.
- [6] Dow, J., 1998, Arbitrage, Hedging, and Financial Innovation, *Review of Financial Studies*, 11, 739-755.
- [7] Faure-Grimaud, A. and D. Gromb, 2004, Public Trading and Private Incentives, *Review of Financial Studies* 17, 985-1014.
- [8] Fulghieri, P., and D. Lukin, 2001, Information Production, Dilution Costs, and Optimal Security Design, *Journal of Financial Economics*, 61, 3-42.
- [9] Glosten, L., and P. Milgrom, 1985, Bid, Ask and Transactions Prices in a Specialist Market with Heterogeneously Informed Traders, *Journal of Financial Economics*, 14, 71-100.
- [10] Gorton, G., and G. Pennacchi, 1990, Financial Intermediaries and Liquidity Creation, *Journal of Finance*, 45, 49-71.
- [11] Gorton, G. and G. Pennacchi, 1995, Banks and Loan Sales: Marketing Nonmarketable Assets, *Journal of Monetary Economics*, 389-411.

- [12] Hanson, S.G. and A. Sunderam, Are There Too Many Safe Securities? Securitization and Incentives for Information Production, working paper, Harvard Business School.
- [13] Holmström, B. and J. Tirole, 1993, Market Liquidity and Performance Monitoring, *Journal of Political Economy*, 678-709.
- [14] Innes, R., 1990, Limited Liability and Incentive Contracting with Ex Ante Action Choices, *Journal of Economic Theory*, 45-67.
- [15] Keys, B.J., T. Mukherjee, A. Seru, and V. Vig, 2010, Did Securitization Lead to Lax Screening? Evidence from Subprime Loans, *Quarterly Journal of Economics*.
- [16] Keys, B.J., A. Seru and V. Vig, 2011, Lender Screening and the Role of Securitization: Evidence from Prime and Subprime Mortgage Markets, forthcoming in *Review of Financial Studies*.
- [17] Kyle, A. S., 1985, Continuous Auctions with Insider Trading, *Econometrica*, 1315-1335.
- [18] Leland, H.E. and D. Pyle, 1977, Information Asymmetries, Financial Structure, and Financial Intermediation, *Journal of Finance*, 371-387.
- [19] Maskin, E., and J. Tirole, 1992, The Principal-Agent Relationship with an Informed Principal, II: Common Values, *Econometrica*, 60, 1-42.
- [20] Maug, E., 1998, Large Shareholders as Monitors: Is There a Trade-off Between Liquidity and Control?, *Journal of Finance* 53, 65-98.
- [21] Myers, S., and N. Majluf, 1984, Corporate Financing and Investment When Firms Have Information Shareholders Do Not Have, *Journal of Financial Economics*, 187-221.
- [22] Pagano, M. and P. Volpin, 2010, Securitization, Transparency and Liquidity, working paper, University of Naples Federico II.
- [23] Parlour, C.A. and G. Plantin, 2008, Loan Sales and Relationship Banking, *Journal of Finance*, 1291-1314.

- [24] Plantin, G., 2010, Good Securitization, Bad Securitization, forthcoming in *Journal of Financial Economics*.
- [25] Rajan, U., A. Seru, and V. Vig, 2010, The Failure of Models that Predict Failure, working paper, Chicago-Booth GSB.
- [26] Shleifer, A. and R.W. Vishny, 2010, Unstable Banking, *Journal of Financial Economics*, 306-318.
- [27] Tirole J., 2005, *The Theory of Corporate Finance*, Princeton University Press.

Table 1: Aggregate Demand Outcomes

Type	Signal	Uninformed Vulnerable	Informed Demand	Uninformed Demand	Aggregate Demand	Probability
\bar{q}	\bar{s}	Y	X	X	$2X$	$\frac{\rho\sigma}{2}$
\bar{q}	\bar{s}	N	X	0	X	$\frac{\rho\sigma}{2}$
\bar{q}	\underline{s}	Y	0	X	X	$\frac{\rho(1-\sigma)}{2}$
\bar{q}	\underline{s}	N	0	0	0	$\frac{\rho(1-\sigma)}{2}$
\underline{q}	\underline{s}	Y	0	X	X	$\frac{(1-\rho)\sigma}{2}$
\underline{q}	\underline{s}	N	0	0	0	$\frac{(1-\rho)\sigma}{2}$
\underline{q}	\bar{s}	Y	X	X	$2X$	$\frac{(1-\rho)(1-\sigma)}{2}$
\underline{q}	\bar{s}	N	X	0	X	$\frac{(1-\rho)(1-\sigma)}{2}$

Figure 1: Timeline

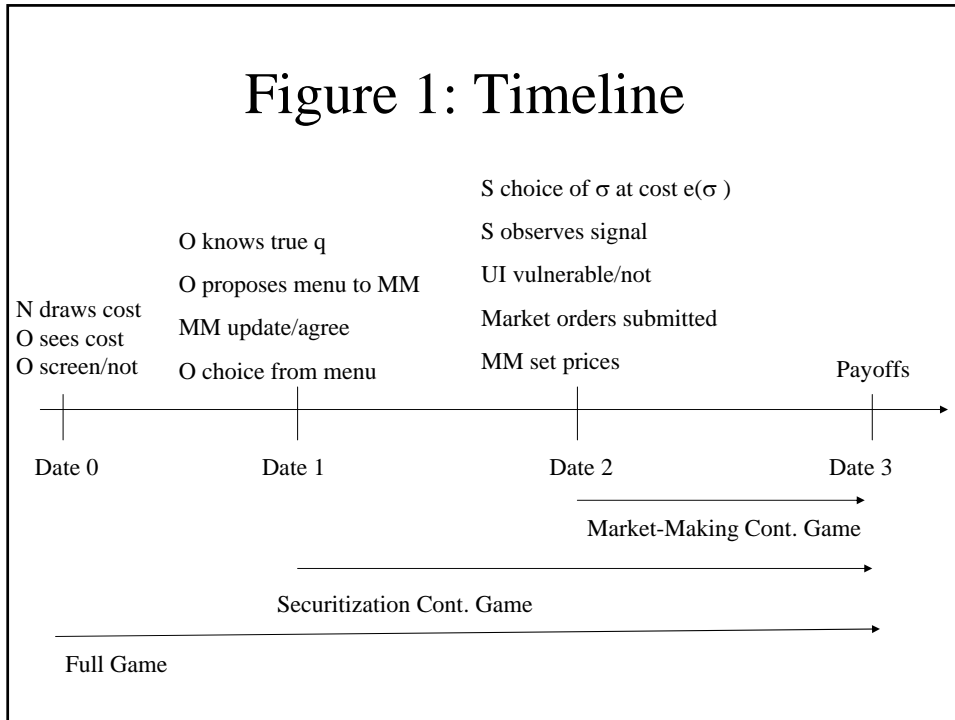


Figure 2: Market-Making Game Equilibrium

