

Computational Complexity and Information Asymmetry in Financial Products

By Sanjeev Arora, Boaz Barak, Markus Brunnermeier, and Rong Ge

1. INTRODUCTION

A *financial derivative* is a contract entered between two parties, in which they agree to exchange payments based on events or on the performance of one or more underlying assets—independently of whether they own or control the underlying assets (e.g., the DOW Jones index falling below 10,000). Securitization of cash flows using derivatives transformed the financial industry over the last three decades. However, mispricing of these derivatives is believed to have contributed to the financial crash of 2008 (see e.g. Brunnermeier⁸ and Coval et al.¹⁰).

There are also suggestions that derivatives were deliberately misused. In Spring 2010 there was a famous allegation about investment bank Goldman Sachs's Abacus derivative. The security and exchange commission alleged that Goldman Sachs collaborated with short-seller Paulson to select particularly bad mortgages as the underlying assets for this derivative. Tranches of Abacus were sold to ABN Amro and IKB Deutsche Industriebank, who were unaware of Goldman's selection methods and lost almost \$1 billion within a few months of buying these assets.

Hence, it is not surprising that derivatives have attracted criticism—Warren Buffett famously called them “financial weapons of mass destruction”—accompanied with calls for extensive regulation and even an outright ban. Others point out—with ample justification from economic theory—that derivatives are beneficial because they allow better risk sharing by “completing” markets, and also protect buyers from the effects of asymmetric information, ameliorating the so-called “lemon” problem (which arises whenever one party in the transaction has more information about the asset than the other; cf. Section 2). According to this viewpoint, problems with derivatives would disappear with use of more accurate financial models, more vigilance by buyers and better governmental oversight.

In our paper,⁵ which the current write-up is trying to describe at a simplified level, we injected a new aspect into this debate. We show that *even when the underlying financial model used by buyers and sellers is correct* there is an inherent obstacle to accurate pricing due to *computational complexity*. Formally, even in industry-standard models, the pricing problem can be as difficult as solving the *planted dense subgraph* problem, which has been proposed as a basis for cryptosystems. The practical implication is that though derivatives such as collateralized debt obligations (CDOs) can theoretically ameliorate the effects of asymmetric

information in the market, in practice these effects will persist—or even get worse—because market participants are not computationally sophisticated enough to solve cryptographic problems. This suggests that better regulation and more informed buyers are not sufficient for derivatives market to work correctly, or at least that regulators and buyers should take computational complexity into account. Such issues are further discussed at the end of the paper in Section 5.

The bite of computational complexity: Computational complexity studies *intractable* problems, such as NP-complete problems, which are conjectured to require more computational resources than can be provided by the fastest computers. (For an introduction see the text.⁴) The key reason it comes naturally into the study of financial derivatives is that it implies an asymmetry between the *ease* of creating problems and *solving* them. A simple example is the problem of *factoring* integers. It is easy to take two random prime numbers—say 7019 and 5683—and multiply them—in this case, to obtain 39888977. However, given 39888977, it is not that easy to factor it to get the two numbers 7019 and 5683. Algorithms that search over potential factors take very long time. This difficulty becomes more pronounced as the numbers have more and more digits. Computer scientists believe that factoring an n -digit number requires roughly $\exp(n^{1/3})$ time to solve,^a a quantity that becomes astronomical even for a moderate n like 10,000. The intractability of this problem leads to a concrete realization of *information asymmetry*. Anybody who knows how to multiply can randomly generate (using a few coin flips and a pen and paper) a large integer by multiplying two smaller factors. This integer could have say 1000 digits, and hence can fit in a paragraph of text. The person who generated this integer knows its (prime) factors, but no computational device in the universe can find a nontrivial factor in any plausible amount of time.^b This informational asymmetry underlies modern cryptosystems, which allow (for

^a The precise function is more complicated, but in particular the security of most electronic commerce depends on the infeasibility of factoring integers with roughly 800 digits.

^b Experts in computational complexity should note that we use factoring merely as a simple illustrative example. For this reason we ignore the issue of *quantum computers*, whose possible existence is relevant to the factoring problem, but does not seem to have any bearing on the computational problems used in this paper.

A previous version of this paper appeared in *The First Symposium on Innovations in Computer Science (ICS 2010)*. Tsinghua University Press, Beijing, China; 49–65.

example) two parties to exchange information over an open channel in a way that an eavesdropper can extract *no* information from it—not even distinguish it from a randomly generated sequence of symbols. More generally, in computational complexity we consider a computational task *infeasible* if the resources needed to solve it grow *exponentially* in the length of the input, and consider it *feasible* if these resources only grow polynomially in the input length.

Computational complexity immediately implies the existence of hard-to-price derivatives, albeit unnatural ones. Consider for example a derivative whose contract contains a 10,000 digit integer n and has a nonzero payoff iff the unemployment rate next January, when rounded to the nearest integer, is the last digit of a factor of n . A relatively unsophisticated seller can generate such a derivative together with a fairly accurate estimate of its yield (to the extent that unemployment rate is predictable), yet even a sophisticated investor like Goldman Sachs would have no idea what to pay for it. This example shows both the difficulty of pricing arbitrary derivatives and the possible increase in asymmetry of information via derivatives.

While this “factoring derivative” is obviously far removed from anything used in current markets, in this work we show that similar effects can be obtained in simpler and more popular classes of derivatives that are essentially the ones used in real life in securitization of mortgages and other forms of debt. The person selling the derivative can structure (“rig”) the derivative in a way such that it has low yield, but distinguishing it from a normal (“unrigged”) higher yield derivative is computationally intractable. Thus any efficient pricing mechanism would either overvalue the rigged derivative or undervalue the unrigged one, hence creating an inefficiency in the market.

Densest subgraph problem: Our result relies on the conjecture that there does not exist a tractable algorithm to detect large *dense subgraphs* in random graphs. This is a more specialized assumption than the familiar $P \neq NP$ conjecture. We needed this assumption because we needed to exhibit the intractability of “real-life” derivatives, and the setting there naturally leads to random graphs, as will be clear in the description in Section 4.

Computational complexity and “bounded rationality”: Computational complexity can be related to the *bounded rationality* concept in economics. Simon¹⁴ proposed the notion of bounded rationality to recognize that in decision making, real-life agents are limited by their cognitive ability to process information and the finite amount of time they have. Simon postulates that agents use heuristics instead of time-consuming and complex optimizing behavior. Experimental evidence on behavioral biases supports this notion (e.g. Kahneman,¹³ etc.). On the other hand, economic experiments also suggest that as the stakes rise and people face similar situations repeatedly, they behave more deliberately in a way that approaches rationality. In particular this is the case in the setting of finance, where stakes are high and traders have access to cutting edge technology. However, even the most sophisticated traders cannot escape the limitations of computational complexity, since no physically realizable computational device can solve intractable problems.

And we show in this note that pricing certain financial derivatives may require solving problems that are believed to be intractable, hence placing it beyond the reach of any real-life agent.

2. THE “LEMONS PROBLEM” IN ECONOMICS

To understand the theoretical benefits of financial derivatives it is useful to recall the *lemons problem*, introduced in Akerlof’s classic 1970 paper.¹ The simplest setting is as follows. You are in the market for a used car. A used car in working condition is worth \$1000. However, 20% of the used cars are *lemons* (i.e., are useless, even though they look fine on the outside) and their true worth is \$0. Thus if you could pick a used car at random then its expected worth would be only \$800 and not \$1000. Now consider the seller’s perspective. Suppose sellers know whether or not they own a lemon. A seller who knows he has a non-lemon would be unwilling to sell for \$800, and would therefore withdraw from the market. The market would be left only with lemons, and knowing this, buyers would refuse to buy any car. Thus the market grounds to a halt. Akerlof’s paper goes on to analyze reasons why used cars do sell in real life. We will be interested in one of the reasons, namely, that there could be a *difference* between what a car is worth to a buyer versus a seller. In the above example, the seller’s value for a working car might be \$200 less than the buyer’s—perhaps because the seller is moving across the country and needs the cash—thus allowing trade to occur. In this case we say that the “lemon cost” of this market is \$200. Some authors refer to the lemon cost as a *wedge*. Generally, the higher this cost, the less efficient the market.

The lemons problem can potentially arise in almost every area of the economy, and a large body of work in *information economics* describes how it can be ameliorated. Akerlof’s original paper already described signaling mechanisms by which a seller can reliably communicate his private information—namely, prove that his car is not a lemon—to the buyer. For example, a used car dealer can show the buyer repair records, or provide a warranty for 6 months, or point out to his stellar reputation and rating from the Better Business Bureau.

3. FINANCIAL DERIVATIVES AND CDOs

The lemons problem also arises in the financial industry and the usual mechanisms for dealing with lemons problems (as identified by Akerlof) have also flowered: borrower FICA ratings (a.k.a. credit scores), ratings of individual securities by agencies such as Moody’s, etc. Financial derivatives provide another mechanism for dealing with the lemons problem. Below we will illustrate this with a common derivative called the *collateralized debt obligation* or *CDO*.

It is not commonly known, but the humble home mortgage is actually a very complex instrument that presents great risks for lenders. The payoff structure is complicated; risk of default is fairly high (compared say to a U.S. treasury bond); and there is high risk of prepayment at times of low interest rates, which is precisely when the mortgage’s locked-in higher rate is most valuable to the lender. CDOs are financial devices that allow many mortgages to be aggregated into a security that has a supposedly much more

predictable yield, and hence more attractive to risk-averse lenders such as retirement funds.

Consider the following simplistic example: suppose a bank holds a portfolio of 100 mortgages, and that each mortgage yields \$0 if the borrower defaults and \$1 million otherwise. For now assume the probability of default is 10%, which implies that the expected yield (and hence fair price) for the entire portfolio is \$90 million. The bank would like to get all or most of these mortgages off its books because this is favorable for regulatory reasons. Since each individual mortgage may be unacceptably risky for a risk-averse buyer, the bank holding the mortgages can do the following. Create two new assets by combining the above 100 mortgages. Set a *threshold*, say \$80 million. The first asset, called the *senior tranche*, has claim to the first 80 million of the yield; and the second, called the *junior tranche*, has claim to the rest. These assets are known as collateralized debt obligations (CDOs). The bank offers the senior tranche to the risk-averse buyer and holds on to the junior tranche (Figure 1).

Now a risk-averse buyer may reason as follows: *if* the mortgage defaults are independent events (which is a big if, though in practice justified by pooling mortgages made to a geographically diverse group of homeowners, whose defaults are presumably independent) then the senior tranche is extremely unlikely to ever yield less than its maximum total value of \$80 million. In fact for this to happen, more than 20 mortgages need to default, which happens only with probability $\sum_{k>20} \binom{100}{k} (0.1)^k (1-0.1)^{100-k} \cong 0.08\%$. Thus the senior tranche is a very safe asset with a highly predictable yield, and such derivatives were often rated by credit-rating agencies to be as safe as a U.S. treasury bond. In real-life CDOs the mortgage yields, payment streams, and tranching are all much more complex, but the basic idea is the same.

The lemons problem enters: There is an obvious lemons problem in the above scenario because of asymmetric information: The bank that issued the mortgages has the most accurate

estimate of the rate of default, whereas the buyer may have a less precise idea—say that the rate lies between 10% and 15%. Economic theory says that in addition to transforming the risk profile of the asset, the CDO also protects the investor from this lemons problem. The crucial observation is that even if the probability of default was 15%, the probability of the senior tranche yielding its maximum of \$80 million would still be roughly 99.9%, and hence the tranching *insulates* the buyer from the information-sensitive part of the mortgage pool.

In fact, as was shown by DeMarzo,¹¹ the choice of the threshold can be used as a *signaling* mechanism that allows the bank to transmit in a trustworthy way the true default value. We now illustrate the idea behind this result. Consider the problem from the bank's viewpoint. It is interested in getting rid of as many mortgages—specifically, the largest possible portion of the entire portfolio—from its book as possible, so it wants to set the threshold as high as possible. Suppose it knows the default rate is 15%. This is the highest possible default rate, so it can simply sell the whole bundle of mortgages (i.e. set the threshold to 100%). The buyers will use their most pessimistic evaluation (that the default rate is 15%) and pay the price of \$85 million. Both the seller and the buyer are satisfied because the price is just equal to the estimated yield. Suppose on the other hand that the bank knows the default rate is actually only 10%, the lowest possible rate. Now setting the threshold to 100% is no longer its best strategy, since the buyers will just pay \$85 million while the bundle is now worth \$90 million. To signal its confidence in the quality of mortgages, the bank will tranche the pool, set the threshold to 80%,^c and offer to hold the riskier part—the junior tranche. Knowing that the bank will not offer this lower threshold when the default rate is high (indeed, the best threshold for default rate 15% is 100%), the rational buyers should correctly interpret this as a signal to the true default rate and pay close to \$80 million for the senior tranche. Again, both the seller and the buyer are satisfied because the price is almost equal to the estimated yield of the senior tranche.

Making the above intuitive argument precise in the usual rational expectations framework of economic theory takes some work, and was done in DeMarzo,¹¹ where it is shown that the CDO is the *optimum* solution to the lemons problem in a setting somewhat more general than the above simplistic one. Specifically, the CDO allows the lemon cost—i.e., the difference in valuation of the security by buyer and bank required for the sale to occur—to approach 0.^d That is, the bank's secret information does not lead to large market inefficiencies. Henceforth we will refer to this as *DeMarzo's Theorem*.

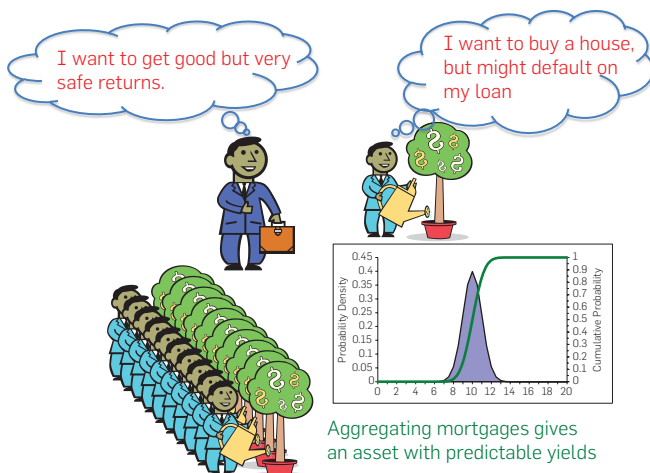
4. WHY COMPLEXITY MATTERS

We presented above the traditional justification for CDOs from economic theory. Now we explain at an intuitive level

^c The exact threshold here depends on a number of factors, including default rate and *discount factor*. The discount factor shows how much the seller prefers cash to assets. The threshold can be computed exactly using methods in DeMarzo.¹¹

^d A nonzero difference in valuation or *wedge* between the bank and buyer arises because the buyer holds cash and the bank holds the mortgages, and the bank prefers cash to mortgages because of regulatory or other reasons.

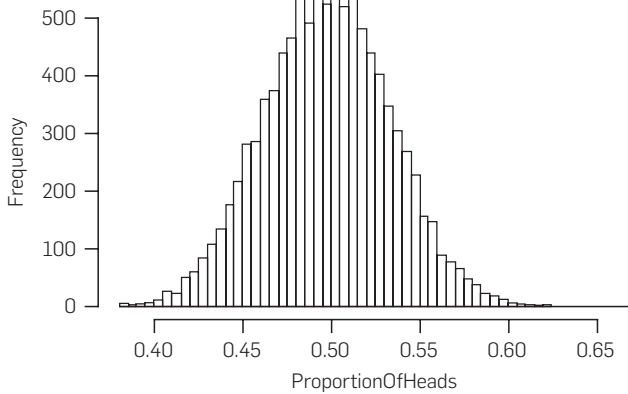
Figure 1. Aggregating many mortgages into a single asset makes the yield more predictable due to the law of large numbers (central limit theorem). This assumes that the yields of the different mortgages are independent random variables.



GAUSSIANS AND THE CENTRAL LIMIT THEOREM

The ability of CDOs to ameliorate effects of asymmetric information relies on the *law of large numbers*, which informally speaking states that the total payoff of a bundle of mortgages is close to the mean (or expected) value. More precisely, we have the central limit theorem: The sum of sufficiently many bounded and independent random variables will be approximately distributed like a Gaussian with same mean and variance. The Gaussian is a good approximation even when the number of variables is only in the hundreds. Therefore, if there are 100 mortgages, each paying 1 with probability 1/2 and paying 0 otherwise, the distribution of the total payoff is like a Gaussian with mean 50 and standard deviation 5. The probability that the payoff is outside [35, 65] (three standard deviations) is less than 0.2%.

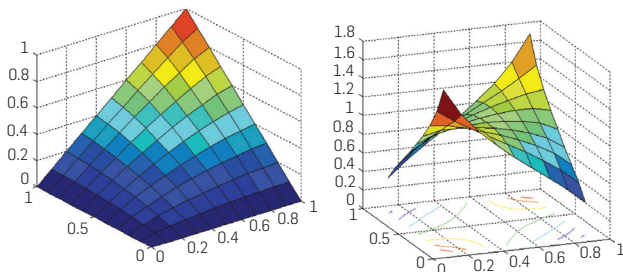
Histogram of ProportionOfHeads



In addition to their value as limiting distributions for the sum of independent random variables, Gaussians arise in one other way in finance: often the payoffs of assets themselves are assumed to be Gaussians. The joint distribution of these Gaussian valued assets is the well-known *Gaussian copula*:

Gaussian copula

Density of Gaussian copula



Although in the illustrative example we assumed binary payoffs for a single asset, similar results hold for asset yields that form a Gaussian copula with the same mean, variance, and covariance.

why introducing computational complexity into the picture greatly complicates the picture, and even takes away some of the theoretical benefits of CDOs. The full analysis appears in our longer paper.

The important twist we introduce in the above scenario is that a large bank is selling many CDOs and not just a single one. DeMarzo's theorem does not generalize to this case, and indeed we will show that whether or not the CDOs ameliorate the lemons problem depends upon the computational ability of the buyer. The lemons problem gets ameliorated only if the buyer is capable of solving the densest subgraph problem, which is currently believed to be computationally difficult.

We will use the assumption—consistent with practice—that mortgages are grouped into *classes* depending upon factors such as the borrower's credit score, geographic location, etc., and that default rates within a class are the same. Consider for simplicity a bank with N asset classes, each of which contains C assets. Some asset classes are "lemons": assets in these classes will always *default* and have payoff 0. All other asset classes are good: assets in these classes pay $1/C$ with probability $1/2$ and default (i.e., have payoff 0) with probability $1/2$. The yields of assets from different asset classes are independent.

The buyer's prior is that the number of lemon classes is uniformly distributed in $[0, 2n]$ for some $n < N/2$, and the set of lemon classes is uniformly picked among all classes. However, the bank has additional information: it knows precisely which classes are lemons (this implies that it knows the number of lemons as well). This is the *asymmetric information*.

Since the expected number of lemon classes is n , each with payoff 0 and the remaining $N-n$ good classes have payoff $1/2$, a buyer purchasing the entire portfolio would be willing to pay the expected yield, which is $(N-n)/2$. Thus a *wedge* à la Akerlof arises for banks who discover that the number of lemons is lower than the expectation, and they would either exit the market, or would need to prefer cash by an amount that overcomes the wedge.

Of course, DeMarzo's theorem allows this lemons problem to be ameliorated, via securitization of the entire portfolio into a single CDO. As already mentioned, we are interested in the case where the number of assets held by the bank is *large*, and so, rather than using a single CDO, the bank partitions them into multiple CDOs. Now how does the bank's extra information affect the sale? Clearly, it has new *cherry-picking possibilities*, involving which asset to package into which CDO. We will assume that all transactions are public and visible to all buyers, which means that seller must do any such cherry picking in full public view.^e

Now let us show that in principle derivatives should still allow buyers to rule out any significant cherry picking, thus ameliorating the lemon wedge. Consider the following: the seller creates M new financial products, each of them

^e This assumption of transparency only makes our negative results stronger. It also may be a reasonable approximation if buyers are well-informed, and recent financial regulation has mandated more transparency into the market.

a CDO depending on a pool of D of the underlying assets. We assume $MD=NC$ and that every asset occurs in exactly one pool. Clearly, assets in the same class should be distributed to different products in order to ensure the diversity of the products. Each CDO will have assets from distinct classes (and hence have stochastically independent yields) so that sum of their yields follows the central limit theorem. Suppose each one of the M products has the following design: it pays off $N/(3M)$ units as long as the number of assets in its pool that defaulted is at most $D/2+t\sqrt{D}$ for some parameter t (set to be about $\sqrt{\log D}$), and otherwise it pays 0. Henceforth we call such a product a “binary CDO”; it can be viewed as the senior tranche of a simple CDO.^f Assets contained in the same product come from different classes, the yields of good assets are uniformly iid, so the expected number of defaults among D good assets is $D/2$ and the standard deviation is $\sqrt{D}/2$. Now the central limit theorem applies and the total number of defaults may be assumed to be distributed like a Gaussian. Thus so long as the fraction of lemon classes is much smaller than the safety margin of t standard deviations, the probability of default for an individual CDO is tiny. Thus, if V denotes the combined expected yield of these M products, then $V \approx M \times N/3M = N/3$. (The exact value of V is unimportant below.)

If the bank were to pick the pools truly randomly—i.e., the entire portfolio of assets is randomly partitioned into the M pools—then the portfolio’s expected yield is only mildly affected by the presence of lemons. Specifically, if V is the expected yield when there are no lemon classes, then it can be shown that the yield is still $V - o(n)$ (i.e. larger than $V - \epsilon n$ for any $\epsilon > 0$) when the number of lemon classes is $2n$, the maximum possible. In this sense derivatives can help significantly reduce the lemon wedge from n to $o(n)$, thus performing their task of allowing a party to sell off the least information-sensitive portion of the risk.

However, the above description assumed that the seller creates the pools *disinterestedly* using pure randomness. But this may be against his self-interest given his secret information! Given that the seller’s interest is to give out the minimum yield possible, as long as this is undetected by the buyer, it turns out that his optimum strategy is to pick some subset of m of the financial products, and ensure that the lemon assets are *overrepresented* in the pools of these m products—to an extent about $t\sqrt{D}$, which is just enough to significantly skew the probability of default. We call this subset of n CDOs the “boobytrap.” Thus the CDOs in the boobytrap have a much higher probability of default than buyers expect, causing the expected yield of the entire portfolio of CDOs to be smaller by an amount proportional to m (roughly $mN/(3M)$).^g

With some settings of m the tampered derivative can have much smaller yield than the yield of $V - o(n)$ obtained

^f This is a so-called *synthetic binary option*. The more popular CDO derivative described above behaves in a similar way, except that if there are defaults above the threshold (in this case $D/2+t\sqrt{D}$) then the payoff is not 0 but the defaults are just deducted from the total payoff. We call this a “tranching CDO” to distinguish it from the binary CDO.

^g The non-booby trapped CDOs will have a slightly smaller probability of default than in the untampered (i.e., random) case, but a simple calculation shows that this will only contribute a negligible amount to the yield.

by random pooling. The question is whether buyers can be fooled by this cherry picking. We have to consider two cases, based on the buyer’s computational powers.

Fully rational (computationally unbounded) buyer: He will not be fooled. Even though he does not know the set of lemon classes, he knows thanks to *random graph theory* (see the excellent references of Alon and Spencer² and Bollobás⁷) that in a randomly chosen portfolio of CDOs the possibility of accidentally setting up such a boobytrap is vanishingly remote. Therefore it suffices for him to rule out the existence of any boobytrap in the presented portfolio: he enumerates over *all* possible $2n$ -sized subsets of the N classes and verifies that *none* of them are over-represented in *any* subset of m products. The same calculations as above guarantee him that in this case the yield of the derivative is at least $V - o(n)$, even though he does not know the identity of the lemon classes. Thus a seller has no incentive to plant a boobytrap for a fully rational buyer, and the lemon wedge is indeed ameliorated greatly if buyers are fully rational.

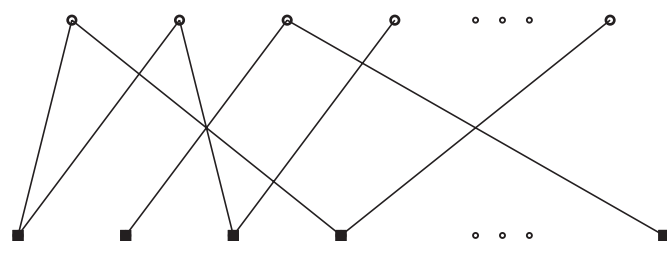
Real-life buyer, who is feasibly rational (computationally bounded): For him the above computation for detecting boobytraps is infeasible even for moderate parameter values.

To get an appreciation of the infeasible problem lurking here, it helps to take a graph-theoretic view of the problem. Recall that a *bipartite* graph consists of two disjoint sets of vertices A, B such that each edge has an endpoint in both A and B . We can use a bipartite graph to represent the portfolio of CDOs: A is the set of asset classes and B is the set of CDOs, and an edge (a, b) indicates that the CDO numbered b contains an asset from the asset class numbered a (see Figure 2).

Of course the buyer will also try other possible algorithms to detect the boobytrap. If the bank randomly throws assets into CDOs, then this graph that represents the portfolio is some kind of random graph. If the bank creates a boobytrap as described above, then the boobytrap corresponds to a *dense subgraph* in this bipartite graph: it is a subset of asset classes (the lemons) and a subset of CDOs (the boobytrapped ones) where the number of edges lying between them is substantially higher than it would be in a random graph.

The problem of detecting a boobytrap is equivalent to the so-called *hidden dense subgraph* problem, which is widely believed to be intractable. In fact the conjecture is that there is no efficient way to distinguish the truly random bipartite

Figure 2. Using a bipartite graph to represent asset classes and derivatives. There are M vertices on top corresponding to the derivatives and N vertices at the bottom corresponding to asset classes. Each derivative references D assets in different classes.



graph from one in which the bank has “planted” the dense subgraph (a.k.a. boobytrap). Formally, the two kinds of graphs are believed to be *computationally indistinguishable* for polynomial-time algorithms and this was the basis of a recent cryptosystem proposed by Applebaum et al.³

The conjecture is that even quite large boobytraps may be undetectable: the expected yield of the entire portfolio could be much less than say $V - n^{1.1}$ and yet the buyer may not be able to distinguish it from a truly random (i.e., honestly constructed) portfolio, whose yield is $V - o(n)$.

We conclude that if buyers are computationally bounded, then introducing derivatives into the picture not only *fails to reduce* the lemon wedge, but paradoxically, *amplifies* it even beyond the total value $2n$ of all lemon assets. Though the above example is highly simplified, it can be embedded in settings that are closer to real life and similar results are obtained.

4.1. Can the cost of complexity be mitigated?

In Akerlof’s classic analysis, the no-trade outcome dictated by lemon costs can be mitigated by appropriate signaling mechanism—e.g., car dealers offering warranties to increase confidence that the car being sold is not a lemon. In the above setting, however, there seems to be no direct way for seller to prove that the financial product is *untampered* i.e., free of boobytraps. (It is believed that there is no simple way to prove the absence of a dense subgraph; this is related to the $NP \neq coNP$ conjecture.) Furthermore, we can show that for suitable parameter choices the tampering is *undetectable* by the buyer even *ex post*. The buyer realizes at the end that the financial products had a much lower yield than expected, but would be unable to prove that this was due to the seller’s tampering. Nevertheless, we do show in our paper⁵ that one could use ideas from computer science in designing derivatives that are tamperproof in our simple setting.

4.2. Complexity ranking

Recently, Brunnermeier and Oehmke⁹ suggested that traders have an intuitive notion of complexity for derivatives. Real-life markets tend to view derivatives such as CDO^2 (a CDO whose underlying assets are CDOs like the one described earlier) as complex and derivatives like CDO^3 (a CDO whose underlying assets are CDO^2) as even more so. One might think that the number of layers of removal from a simple underlying real asset could be a natural measure of complexity. However, as Brunnermeier and Oehmke⁹ point out, such a definition might not be appropriate, since it would rank e.g. highly liquid stocks of investment banks, which hold CDO^2 s and other complex assets, as one of the most complex securities. Our paper⁵ proposes an alternative complexity ranking which is based on the above discussed notion of *lemon cost due to complexity*. This ranking also confirms the standard intuition that CDO^2 s are more complex than CDO s. Roughly speaking, the cherry-picking possibilities for sellers of CDOs described in this paper become even more serious for derivatives such as CDO^2 and CDO^3 .

5. DISCUSSION

The notion that derivatives need careful handling has been extensively discussed before. Coval et al.¹⁰ show that pricing (or rating) a structured finance product like a CDO is extremely fragile to modest imprecision in evaluating underlying risks, including systematic risks. The high level idea is that these everyday derivatives are based upon the *threshold* function, which is *highly sensitive* to small perturbations of the input distribution. Indeed, empirical studies suggest that valuations for a given financial product by different sophisticated investment banks can be easily 17% apart⁶ and that even a single bank’s evaluations of different “tranches” of the same derivative may be mutually inconsistent.¹² Thus one imagines that banks are using different models and assumptions in evaluating derivatives.

The question studied in our work is: *Is there a problem with derivatives even if one assumes away the above possibilities, in other words the yield of the underlying asset exactly fits the stochastic model assumed by the buyer?* Economic theory suggests the answer is “No”: informed and rational buyers need not fear derivatives. (Recall our discussion of DeMarzo’s theorem.)

The main contribution of our work has been to formalize settings in which this prediction of economic theory may fall short (or even be falsified), and manipulation is possible and undetectable by all real-life (i.e., computationally bounded) buyers. We have worked within existing conceptual frameworks for asymmetric information. It turns out that the seller can benefit from his secret information (viz., which assets are lemons) by using the well-known fact that a random election involving n voters can be swung with significant probability by making \sqrt{n} voters vote the same way; this was the basis of the boobytrap described earlier. The surprising fact is that a computationally limited buyer may not have any way to distinguish such a *tampered* CDO from untampered CDOs. Formally, the indistinguishability relies upon the conjectured intractability of the *planted dense subgraph* problem.^h


The model in our more detailed paper has several notable features:

1. The largeness of the market—specifically, the fact that sellers are constructing thousands of financial products rather than a single product as was the case in the model of DeMarzo¹¹—allows sellers to cherry pick in such a way that cannot be detected by feasible rational (computationally bounded) buyers—i.e., all real-world buyers—while it can be detected by fully rational (computationally unbounded) buyers.
2. The possibility of cherry picking by sellers creates an Akerlof-like *wedge* between buyer’s and seller’s valuations of the financial product. We call this the *lemon cost due to computational complexity*. In our detailed paper we can quantify this wedge for several classes of derivatives popular in securitization. This allows a par-

^h Note that debt-rating agencies such as Moody’s or S&P currently use simple simulation-based approaches to evaluate derivatives, which certainly do not attempt to solve something as complicated as densest subgraph.

tial ranking of these classes, which can be seen as a quantification of more familiar heuristic notions of “complexity.” This answers the open question of Brunnermeier and Oehmke.⁹

3. It can be difficult for regulatory bodies to control the above-mentioned cherry picking because the cherry picking can be difficult to detect *ex ante*. In some models the cherry picking seems undetectable even *ex post*. Both these remain true even in a fully transparent market where all transactions occur on a public exchange. It also implies that verifying the existence of the lemon cost due to computational complexity in historical data (in other words, an empirical test of our paper) may prove difficult, especially given that the market has not been fully transparent.

In sum, our approach of combining insights from computer science with economic questions allows one to formally study phenomena, such as complexity and bounded rationality, that are of first-order importance but were difficult to capture in formal economic models. These new insights should help shape future regulation and the post-2008 financial architecture. 

References

1. Akerlof, G.A. The market for “lemons”: Quality uncertainty and the market mechanism. *Q. J. Econ.* 84(3) (1970), 488–500.
2. Alon, N., Spencer, J.H. *The*

Probabilistic Method, 3rd edn. Wiley, Hoboken, NJ, 2008.

3. Applebaum, B., Barak, B., Wigderson, A. Public-key cryptography from different assumptions. In *Proceedings of STOC*, 2010, 171–180.

4. Arora, S. and Barak, B. *Computational Complexity: A Modern Approach*. Cambridge University Press, 2009.
5. Arora, S., Barak, B., Brunnermeier, M., Ge, R. Computational complexity and information asymmetry in financial products. In *The First Symposium on Innovations in Computer Science, ICS 2010*. Tsinghua University Press, Beijing, 2010, 49–65.
6. Bernardo, A.E., Cornell, B. The valuation of complex derivatives by major investment firms: Empirical evidence. *J. Finance* 52 (1997), 785–798.
7. Bollobás, B. *Random Graphs*, 2nd edn. Cambridge University Press, 2001.
8. Brunnermeier, M.K. Deciphering the liquidity and credit crunch 2007–08. *J. Econ. Perspect.* 23(1) (2009), 77–100.
9. Brunnermeier, M.K., Oehmke, M. Complexity in financial markets. *Working Paper*. Princeton University.
10. Coval, J., Jurek, J., Stafford, E. The economics of structured finance. *J. Econ. Perspect.* 23(1) (2009), 3–25.
11. DeMarzo, P.M. The pooling and tranching of securities: A model of informed intermediation. *Rev. Financ. Stud.* 18(1) (2005), 1–35.
12. Duffie, D. Innovations in credit risk transfer: Implications for financial stability. *BIS Working Papers*.
13. Kahneman, D. Maps of bounded rationality: Psychology for behavioral economics. *Am. Econ. Rev.* 93(5) (2003), 1449–1475.
14. Simon, H.A. Bounded rationality and organizational learning. *Organ. Sci.* 2(1) (1991), 125–134.

Sanjeev Arora (arora@cs.princeton.edu), Department of Computer Science, Center for Computational Intractability, Princeton University, Princeton, NJ.

Markus Brunnermeier (markus@princeton.edu), Department of Economics, Bendheim Center for Finance, Princeton University, Princeton, NJ.

Boaz Barak (boaz@microsoft.com), Microsoft Research New England, Princeton University, Princeton, NJ.

Rong Ge (rongge@cs.princeton.edu), Department of Computer Science, Center for Computational Intractability, Princeton University, Princeton, NJ.

© 2011 ACM 0001-0782/11/05 \$10.00



Association for
Computing Machinery

Advancing Computing as a Science & Profession



You've come a long way.
Share what you've learned.



ACM has partnered with MentorNet, the award-winning nonprofit e-mentoring network in engineering, science and mathematics. MentorNet's award-winning **One-on-One Mentoring Programs** pair ACM student members with mentors from industry, government, higher education, and other sectors.

- Communicate by email about career goals, course work, and many other topics.
- Spend just **20 minutes a week** - and make a huge difference in a student's life.
- Take part in a lively online community of professionals and students all over the world.



Make a difference to a student in your field.
Sign up today at: www.mentornet.net
Find out more at: www.acm.org/mentornet

MentorNet's sponsors include 3M Foundation, ACM, Alcoa Foundation, Agilent Technologies, Amylin Pharmaceuticals, Bechtel Group Foundation, Cisco Systems, Hewlett-Packard Company, IBM Corporation, Intel Foundation, Lockheed Martin Space Systems, National Science Foundation, Naval Research Laboratory, NVIDIA, Sandia National Laboratories, Schlumberger, S.D. Bechtel, Jr. Foundation, Texas Instruments, and The Henry Luce Foundation.