

# The Limits to Compensation in the Financial Sector

Thomas Noe

Balliol College/Saïd Business School, Oxford

H. Peyton Young

Nuffield College/Economics Department, Oxford

10<sup>th</sup> December, 2012

## Abstract

In recent years bonuses tied to performance have become commonplace in banks and other financial institutions; indeed they now constitute a major part of employee compensation. The practice was originally justified by academic work on principal-agent contracts, which argued that performance bonuses would better align the interests of managers and shareholders. In this article we argue that such schemes are not well-suited to aligning these interests in the financial sector. There are two reasons for this failure. First, new financial products make it easy to create the *appearance* of superior performance over long periods of time even though the outsize returns are merely being driven by hidden tail risk. We show that it is virtually impossible to create performance contracts that get around this problem. Second, the complexity of new products and the size of modern financial institutions make it extremely difficult (and costly) to monitor risky activities directly. As in the first case, compensation schemes, including deferred compensation, are inefficient substitutes because it is easy to escape detection for long periods of time. This opens the door for outright fraud. We argue that a greater emphasis on ethical values, e.g., a duty of care to customers and shareholders, is more likely to produce effective reforms.

## 1 Background

It is widely believed that monetary incentives are an effective way to motivate employees' performance and thereby increase the profitability of the enterprises that employ them. Among economists this view has been in the ascendant since the advent of modern contract theory (Jensen and Meckling (1976), Hölmstrom (1979), and Hart and Moore (1990)). Moreover, this theory has had a major impact on how companies actually do structure their compensation. Bonuses tied to performance—usually paid in the form of options or shares—are now commonplace in the corporate world.

In this article we argue that such schemes in fact are not well-suited to inducing honest and trustworthy behavior in the financial sector. Not only are compensation schemes ineffective at deterring “cheats,” they can also be gamed by people who are not delivering above-average performance. We argue that monetary incentives are ill-suited to resolving the agency issues surrounding three functions which are key to efficient and stable financial markets:

- *Asset management*: Creating wealth through asset management
- *Risk management*: Vigilant monitoring of low probability dire consequence “black swan” events

- *Asset Trusteeship*: Protecting firm assets from diversion

The specific reasons for the failure of incentive compensation vary. For fund management, the ability of managers to use financial engineering to manipulate the timing of returns makes performance-based compensation problematic. For risk management, the inherent infrequency of black swan events makes the relation between diligent risk management and firm performance very noisy, thus making incentive compensation for diligence very expensive. In the case of asset trusteeship, the ability of fiduciary agents to accelerate the maturity of asset payoffs prevents principals from using the threat of future liquidation to discourage asset diversion. In each of these cases, the *transmutability* of financial assets—the ability of financial managers to use efficient financial markets to engineer the timing and volatility of cash flows—combined with the high sensitivity of financial asset values to low-probability but highly adverse “black-swan” events makes applying principal-agent contracting problematic.

## 2 Asset management and paying for performance

The key problem that compensation contracts in the financial industry must tackle is tail risk. This term refers to small probability events that have large adverse consequences. Such events are popularly known as “black swans”(see Taleb, 2005). The term suggests that such events occur naturally and they are extremely unlikely. We claim, however, that standard compensation contracts in the financial sector often encourage people to *engineer* the swans, that is, they undertake investment strategies that have a certain probability of blowing up. Furthermore the probability of a blow-up can be quite sizable and still be rational from the standpoint of the agent (though not from the standpoint of the shareholders).

### 2.1 Swans: natural and engineered

Here is an illustrative example. A trader in an investment bank is promised a bonus equal to 20% of the amount by which his investments beat the market in a given year. Let’s say that the market index is the S&P 500 and that the trader, Mr. McCoy, controls \$ 500 million in funds. McCoy’s strategy is the following: he invests everything in the S&P 500 collecting dividends along the way. At some point he creates (“writes”) a bundle of put options that he sells to third parties who wish to hedge against changes in market value. The options are designed so that the option-holder gets all the shares in McCoy’s fund if the S&P 500 price falls below a certain level by a specified nearby date, say at the end of the month. (Such an option is called an asset-or-nothing put and is a standard instrument in financial markets.) Prices in the futures market suggest that the chances of this event are about 1 in 20. This means that Mr. McCoy can sell the options for about .05 times the current number of shares in his fund. He uses the proceeds to buy still more shares of the S&P 500 and waits until the end of the month while keeping his fingers crossed. The chances are 95% that the options are not called. In this case he can sit back and

do nothing more until the end of the year. In other words, he can take on one big bet during the course of the year, and if it works out he will have 5% more shares of the S&P 500 than he did at the start. He has apparently beaten the market by 5%. If the market went up by 10% his fund is now worth  $1.10 \times 1.05 = 1.155$ , that is, it went up by 15.5%. If the market went down by 10% his fund is now worth  $.90 \times 1.05 = .945$ , that is, it went down by only 5.5%. Whether the market goes up or down his bonus is 20% of the gain relative to the market, which comes to approximately  $1\% \times \$500$  million, or \$5 million per year. Of course this is the outcome provided that the black swan does not materialize (the options are not called). If it does his fund will be cleaned out and he will have to find employment elsewhere. But the probability of this outcome is small (5%), and it does not place him in legal jeopardy provided that he does not try to cover up what he is doing. After the fact he can always claim that the outcome was unfortunate and that in his judgment it had only a “tiny” chance of occurring.

This example is merely illustrative and the specific numbers are unimportant. With his arsenal of derivatives McCoy could have manufactured swans with different characteristics. For instance he could have written options that had a 10% probability of being called. In this case he would have earned twice as much money in bonuses by taking on twice as much risk. From McCoy’s point of view the optimally-engineered swan would depend on his degree of risk aversion, his impatience to get the money, and so on.

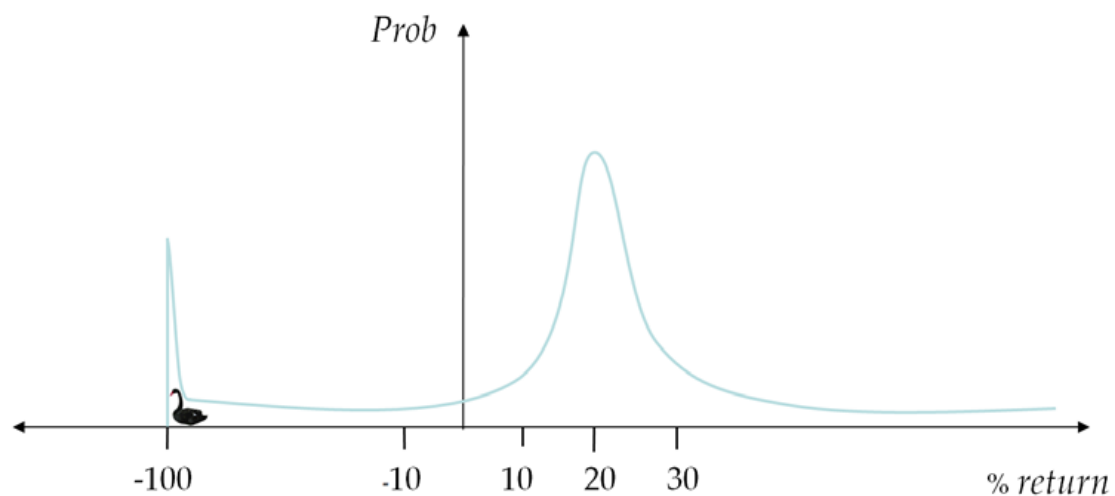


Figure 1: A financially-engineered black swan designed to produce apparent excess returns.

The optimal size swan also depends on another factor: the ability of the fund to attract more money. If a fund manager beats the market by a sizable amount year after year, people will think he is a genius and give him still more money to play with. This is true of fund managers and traders within large investment banks and it is also true of managers who operate their own hedge funds. One of the brilliant strokes of Bernard Madoff

was to manufacture steady returns that were quite good but not so good that suspicions would be aroused. His returns were about 12% per year in good markets and bad, which in that era was about 8% over and above the rate available on safe assets such as Treasury bonds. Of course Madoff was a cheat: he achieved his apparently steady returns by cooking the books and paying off existing investors with funds received from new investors. McCoy's approach is more subtle: he is not cooking the books, he is just taking on risk. To produce apparent excess returns equal to 8% a year by this method, for example, it suffices to manufacture a black swan that shows up once a year with a probability of about 8%.

How long will it take before McCoy is "discovered"? In expectation it will take about 13 years. More generally, if  $p$  is the annual probability of the black swan materializing it will take on average  $1/p$  years for it to show up. But the deeper problem is that McCoy may not be "discovered" even if the black swan does show up. The difficulty is that McCoy is doing nothing illegal: he is merely taking on risks in the hope of making profits. This is precisely what traders and investment managers are supposed to do. The preceding example might suggest that it is easy to spot strategies whose sole purpose is to game the bonus system, but this is not the case. Firstly, there are many other investment strategies that have similar characteristics and look more normal. For example one could take leveraged positions in the foreign exchange markets, or in sovereign debt, or write credit default swaps, and duplicate the essential features of the above example. Moreover the investment manager may genuinely believe that he has a special talent for discerning arbitrage opportunities, and that his strategies really are going to make money, both for himself and the bank. In other words he may not be dishonest but simply delusional about his own investment ability.

Indeed it could be that he actually does have investment ability, and that his risky strategies will pay a handsome excess return on average. In this case the bank certainly would like to employ him. The difficulty is that there is no practical way of telling, within a reasonable length of time, whether or not he is the real McCoy. The problem arises from the fact that what distinguishes the real from the fake is the probability that the black swan will appear. However the probability of the black swan cannot be estimated without having a great deal of data.

Consider the following example. A hedge fund manager produces returns that are 8% over and above the market's returns and he does this for several years in a row. The investor asks himself how likely it is that the manager really has talent for spotting arbitrage opportunities, or is merely lucky in the bets he undertakes. His prior belief is 50-50 that the manager is truly talented. How many years of returns must he see in order to be 95% sure that the manager is talented? Let  $A$  be the event that the manager produces 8% excess returns for  $t$  periods in a row. Let  $B$  be the event that the manager is truly talented and that he can do this every year with probability one. (This is a somewhat extreme assumption but it illustrates our point with even greater force.) As we have seen, a manager with no talent (not  $B$ ) can generate these returns for  $t$  years in a row with probability  $\frac{1}{(1.08)^t} = 0.926^t$ . The investor's prior odds are 1: 1 that the manager is talented versus not talented. Therefore the posterior odds are 1:  $0.926^t$ . To be at least 95% sure that he is investing with a talented manager, the investor needs to see  $t$  years of

returns where  $t$  satisfies

$$0.926^t = \frac{0.05}{0.95} \quad (1)$$

This means collecting about thirty-eight years of data!

In practice therefore it is virtually impossible to distinguish in a statistical sense between the really talented managers and those who are merely imitating the talented managers with high probability. However, it might still be the case that one can distinguish between them by designing their compensation in an appropriate way. Unfortunately this also turns out to be impossible in any practical sense. The arguments require a formal analysis that would take us too far afield here; for details see Foster and Young (2008), Foster and Young (2010), and Foster and Young (2012). We can illustrate the extent of the difficulty by focusing on an apparently plausible way to deal with the problem, namely deferred compensation. These schemes take various forms: the essential idea is to pay bonuses in a form that puts them at risk if future results are poor. One way to do this is to pay the bonus only after  $T$  years of returns have been observed, where  $T$  is typically 3-5 years. Another version is to pay the bonus in restricted stock which only vests after  $T$  years have elapsed. If the firm does badly in the interim, the value of the bonus goes down, which presumably acts as a deterrent for excessive risk-taking. A third, and more recent variant, is to pay the bonuses in contingent convertible bonds (“cocos”). This is a financial instrument that pays interest like a bond, but that converts into common stock if the bank’s capital falls below a target level. In other words if the bank gets into serious trouble, the coco holders are in the same boat as the common stockholders.

We shall make two simple points about such schemes. First, a deferral of 3-5 years does not amount to much of a deterrent. Consider McCoy again, and suppose that his bonus is computed on the basis of returns over a five-year period. If the annual probability of a black swan is 5%, the probability that no black swan materializes in five years is  $0.95^5 = 0.77$ . Meanwhile his excess returns have been compounding: he makes 5% above the market in every year that the black swan does not appear. After five years he has made  $(1.05)^5 = 1.28$ , which is a 28% higher return than the stock market. Hence at the end of five years his bonus will be 20% of 28%, which is 5.6%. Notice that this is actually larger than if he were paid a bonus each year in which the black swan does not appear. (The latter is 20% of 5% = 1% for five years running, which comes to 5%.) Thus postponement for 3-5 years has relatively little effect on the incentive to take on risk; it merely forces the trader to wait longer for his money.

It is also instructive to look at the scheme whereby bankers are paid their bonuses in coco’s, which has recently been proposed by some investment banks. The first point to observe is that the rate of interest on a coco bond must be larger than the risk-free rate because investors need to be compensated for the possibility that the bond will convert to equity when the bank gets into trouble (in which case the equity could turn out to be worthless). Suppose that the risk-free rate is zero and the probability of conversion is  $p$  per annum. Assume for simplicity that conversion leads to a total loss. Then the appropriate interest rate on the coco’s is  $r = 1/(1 - p) - 1$ , since this implies that the expected excess return is zero  $((1 + r)(1 - p) - 1 = 0)$ .

Suppose now that the managers of an investment bank are paid their bonuses in cocos, and suppose for the sake of argument that payment is 10% of the excess return. Their actions create the risk level  $p$ . Assuming they have special investment talent the excess return generated by taking on this risk is precisely  $r$ . At the end of a year the cocos are worth exactly 1 in expectation, because the extra interest offsets the probability of loss. Hence in expectation the bonuses are worth  $.10r$ , which is increasing in  $r$ . In other words, under this scheme the managers are being paid to take on more risk. The difficulty is that the higher level of interest on the cocos offsets the increased risk, hence payment in cocos does nothing to alleviate the problem.

One can in fact establish a quite general impossibility result, which states that, in the framework of limited liability there exists no compensation scheme that separates the talented risk-takers from those who are merely lucky. Any scheme that rewards the former will also reward the latter to a significant extent. The only effective schemes are those that inflict severe personal penalties for bad performance (Foster and Young, 2010).

### **3 Risk management: Paying for diligence**

The preceding analysis shows that incentive compensation is likely to be ineffective for rewarding talented asset managers. Per se, lack of talent is not an ethical fault. Fund managers claiming to be talented when they know that they are not is an ethical fault. However, such managers might simply overestimate their own ability. We now consider another agency problem that is clearly tied to moral fault: negligence in monitoring the firm's environment with the aim of detecting small probability but large-magnitude risks, i.e., "black swan" events. For example, before the financial crisis, academic research identified an increase in the systemic risk of mortgage-backed securities. To the researchers, this raised an issue of mispricing. A diligent risk manager might have been able to tease out the implications of this research for correlated default in response to large negative shocks, and thus saved the bank a lot of money.

Unfortunately, monitoring such risks is usually not terribly exciting work. Most of the time, even if the risk manager is not diligent, potential risks will never materialize. Militaries face a similar problem of monitoring for black swan events: watching the parapets for surprise attack by the enemy. Most the time, there is nothing very interesting to watch and most of the time, neglecting watch duties will not have any adverse consequences. However, when the enemy is actually approaching in the dead of night, neglecting watch duties can lead to the death or capture of the entire unit. Armies have developed simple, direct incentive mechanisms to motivate sentry diligence, for example:

Any sentinel or look-out who is found drunk or sleeping upon his post, or leaves it before he is regularly relieved, shall be punished, if the offense is committed in time of war, by death.

Article 133, U. S. Uniform Code of Military Justice

The military solution to the problem of motivating sentry vigilance is simple and no doubt fairly effective. However, for it to work, negligence must be verifiable, i.e., third parties need to be able to verify whether the soldier is staring across the parapets or over the bar room table. Verification in the case of sentries seems straightforward. However, for financial professionals, third-party verification is much more problematic. Their tasks are complex and probably opaque even to well informed outsiders. Making a verifiable distinction between an analyst diligently trying to ferret out financial threats and one simply going through the motions is difficult and perhaps impossible.

In addition, firms cannot impose the sort of penalties militaries impose. The worst penalty the firm can impose in an incentive contract is dismissal and forfeiture of incentive payments. When verification is costly or impossible and penalties are limited, a natural question is whether monetary incentives can substitute for punishment and monitoring.

We argue that, in these circumstances, it is very costly to use financial incentive contracts to motivate financial professionals to “do their duty.” The contract’s downside is limited, so upside rewards must do the heavy lifting. However, because the upside is fairly likely even when the agent is not diligent, the relation between upside rewards and diligence is very noisy. Thus, motivating diligence will require huge upside rewards: the rewards must be orders of magnitude greater than the risk manager’s cost of effort. If a black swan event is sufficiently disastrous, the firm may be willing to pay this sort of compensation. Otherwise, it will simply take its chances, hoping that the black swan does not arrive. The result, in a world of amoral agents—i.e., agents who place no inherent value on diligence—is either (a) some agents will receive compensation packages offering huge rewards with very low performance targets or (b) financial firms will skimp on watchmen, leading to an increased probability of disaster. To illustrate these ideas, we will first develop a very simple example which illustrates both the efficacy and efficiency of monetary compensation for inducing some kinds of agent effort, and how costly and inefficient it can be for inducing agents to exert agents to watch for and mitigate black swan events.

### **3.1 Black swans and white swans**

First, consider the problem of motivating a scientist working in a research lab to produce a new invention, the “white swan.” The scientist maximizes his expected payoffs which depend on how much effort he devotes to producing he swan and how he is compensated. If the scientist produces the white swan, the firm’s value increases by £1 billion. If the scientist exerts high effort the chance of producing the swan is 1%; if the scientist does not exert high effort, there is no chance of producing the white swan. High effort imposes a personal cost on the scientist with monetary value equal to £10,000. Whether the scientist is exerting high effort cannot be observed or verified by anyone but the scientist.

Note that, in the white-swan case, high effort is a necessary but not sufficient condition for success. Thus,

incentive compensation payments can be targeted precisely at effort. In fact, if the firm pays the scientist based on a simple bonus scheme—£1 million if and only if the scientist produces the white swan—the scientist will be willing to exert high effort. To see this, note that the gain from exerting high effort is earning the £1 million pound bonus whenever the scientist succeeds. With high effort, the scientist will succeed with probability 0.01. Thus, the expected gain from high effort  $0.01 \times 1\text{million} = 10,000$  which is exactly the cost of high effort to the scientist. Note also that the expected reward to the scientist for high effort exactly equals the cost of high effort, 10,000. The firm's gain from offering the incentive contract is simply the expected value of the invention given high effort,  $1\text{Billion} \times 0.01 = 10\text{million}$ , less the cost of offering incentive compensation, 10,000. This is the best the firm could possibly do. Even if the firm could contract on the level of effort, the firm would have to pay the scientist a sufficient increment to induce him to sign the contract. That increment is the cost of high effort, 10,000, and exactly equals the expected value of the incentive payment.

Now let us turn to the a problem which seems quite similar. The firm is a bank with a value of £1 billion. Its value will remain £1 billion so long as a black-swan event does not occur. However, a black swan will lead to immediate collapse. If the bank's risk manager is diligent, he can detect and avoid the black swan. However, if the risk manager is not diligent, in the rare situations in which black-swan event occurs, the bank's value collapses to 0. Diligence is costly to the risk manager, who would prefer tending his Facebook account to diligent risk management. Expressed in monetary terms, the cost of diligence to the risk manager is 10,000. The chance that diligence will be required to avoid collapse is 1%.

Note that in many ways this problem is similar to the scientist motivation problem discussed above: in both cases, the gain to the firm from the agent (i.e., the scientist or risk manager) exerting high effort is 10 million. In the white-swan case, the gain is produced by a 1% chance of producing a 1 billion gain and, in the black-swan case, it is produced by reducing the probability of a 1 billion loss by 1%. Moreover, the cost of high effort to the scientist equals the cost of diligence to the risk manager.

There is one difference between the two cases, a difference which will have dramatic consequences. In the white-swan case, high effort is a necessary but not sufficient condition for success. In the black-swan case, high effort is a sufficient but not necessary condition for success. This makes targeting rewards at effort very difficult.<sup>1</sup> In the black-swan case, there are only two possible outcomes: collapse and non-collapse. The least the risk manager can receive if a collapse occurs is 0. Paying him any more than this when the bank collapses will clearly not improve his incentives. Thus, the problem of incentive design reduces to how much to pay the risk manager when the bank does *not* collapse. Let's call this payment,  $p$ . If the risk manager is diligent he receives  $p$  100% of the time and incurs the cost of high effort, 10,000. If the risk manager is not diligent, he will receives  $p$  99%

<sup>1</sup>For the technically inclined reader, who desires to relate this example to the technical literature on the principal/agent problem, we note that the likelihood ratio for the favorable event (discovery in the white swan case and solvency in the black swan case) in the white-swan case is infinite while in the black-swan case it is only slightly higher than one. See Grossman and Hart (1983) for a discussion of likelihood ratios and the principal/agent problem



of the time but does not incur the cost of high effort. To induce diligence, the expected gain from diligence must exceed its cost, in other words, it must be the case that

$$1.0 \times p - 10,000 \geq 0.99 \times p + 0.01 \times 0 \quad (2)$$

Thus, the incentive payment must at least equal 1 million. This is exactly the same incentive payment required in the white-swan case. However, in the white-swan case, the payment was only received when the invention is produced, which occurred with probability 0.01. In the black-swan case, if the payment works to produce diligence, the bank will never collapse and thus the payment will be made with probability 1. Thus, the expected cost of the incentive payment is at least 1 million. In other words, the bank must pay the risk manager £1 million in order to induce the risk manager to absorb a £10,000 personal cost of diligence. Nice work if you can get it!

### 3.2 More generally

The example given above is highly simplified and the simplifications might lead a reader to conjecture that the problem of motivating diligence through compensation is not nearly as hard as we assert. First, in the black-swan case, the collapse can only occur if the risk manager is negligent. Thus, collapse is proof of negligence, in which case, tort-liability for the risk manager might be an alternative incentive mechanism. There are two problems with this argument. First, even if the parties know that negligence is the only possible cause of collapse, enforcing tort-liability claims on the agent would require the bank to be able to *prove* to third parties that negligence and collapse are perfectly connected. Second, we can simply augment our analysis by positing a 0.10% chance of collapse even if the manager is diligent and thus break the perfect relation between collapse and negligence. Under this assumption, it is not possible to infer with a high degree of certainty that negligence was the cause of collapse. This modification only affects the level of managerial compensation by a small amount, but rules out enforcement through tort liability.

Another reasonable objection to the relevance of this example is risk aversion. If agents are risk averse then they will be willing to exert significant effort in diligence to avoid large losses. It is indeed the case that, as risk aversion increases, the premium required to ensure diligence falls. It is certainly possible to posit levels of risk aversion at which the diligence premium is quite small. The relevant practical question however is whether the premium will be reduced to reasonable size when agents have “typical” levels of risk aversion. Answering this question is not easy because instruments that measure risk aversion are plagued with problems of internal and external validity. Putting these questions aside, most of the literature that has attempted to measure the average risk aversion of human subjects using the simplest possible functional form for risk aversion—expected utility of wealth maximization under constant relative risk aversion—estimates the relative risk aversion coefficient to be around 2/3 (see Harrison and Rutström, 2008). Using this estimate, and solving for the premium required

to induce diligence in the example, yields an expected payment to the risk manager of approximately £330,000 which is still 33 times the cost of diligence to the agent. This is less the cost under risk neutrality but it is still huge relative to the cost of effort. Firms might try to reduce compensation by screening agents for risk aversion but this would be rather difficult. The less risk averse a potential risk-manager is, the more attractive the strategy of accepting the risk management position and then being negligent. Thus, potential risk managers would have an incentive to hide their level of risk aversion. Because risk aversion is hard to measure even where there are no incentives do dissimulate, screening for risk aversion would be very difficult to put into practice. Moreover, even if firms could divine the risk aversion of potential employees, restricting hiring to the most risk averse applicants might have an adverse effect on employee quality.

## **4 Paying for honesty**

Descending ever downward toward the depths of depravity, consider now the problem of using compensation to deter stealing, i.e., diversion of firm, client, or trust-fund assets. Consider the case of an amoral manager who is perhaps capable of superior performance but who is also able to divert corporate assets for his personal consumption. Examples are fairly easy to identify, e.g., Dennis Kozlowski CEO of Tyco in 2002. When diversion cannot be verified or monitored by outsiders, is it possible to design contractual arrangement which deter diversion? At first glance, deterrence might seem impossible. However, contract theory demonstrates that contractual mechanisms do exist which can deter diversion even in this seeming hopeless situation. In the end, we will argue that such mechanisms will not be effective in for financial firms. To understand why, we need to understand how these mechanisms work. Because the mechanisms are much more complex than simple cash flow sharing rules and because we aim for understanding rather than mathematical rigor, we will explain the mechanism through a fable and refer the reader to Clementi and Hopenhayn (2006) for a rigorous development.

### **4.1 Mother Goose: a tale of asset hypothecation**

Consider the problem of Jack, an entrepreneur endowed with a high net present value investment opportunity. Jack has the rather unusual ability to induce a particular goose, henceforth called “Goose,” to lay golden eggs. Jack can never be sure whether his magic will work. In fact, each week that Goose roams with Jack, Goose has a 50% chance of laying a golden egg. Jack has one more magic ability—he can transform his golden egg into invisible strawberries which he can eat without being observed. Jack has an unlimited appetite for strawberries and in fact Jack desires to consume nothing but strawberries. Compensating for his two magic abilities, Jack has two handicaps. First he is penniless, owning nothing, not even Goose. Second, Jack is amoral and moreover his morality deficit is known to all.

Craving strawberries, Jack conceives a plan. He will approach an investment banker, henceforth called

“Banker,” to acquire funds with which to buy Goose. Jack proposes to Banker that, in exchange for financing, he will give Banker 50% of the proceeds from the sale of the golden eggs. Banker, however, is fairly shrewd, and rebuffs Jack, raising the objection that Jack will simply proclaim each week that a golden egg was not laid while transforming any egg that is actually laid into invisible strawberries. Jack, now crestfallen and about to leave Banker’s office remembers a paper he studied in graduate school on long-term contracting. Based on the paper, he proposes a new deal to Banker.

Under Jack’s scheme, Banker will take the goose as collateral. In addition to funding the goose purchase, Banker will establish a bank account for Jack and moreover deposit a starting balance in the account. Each week, Jack will report whether the goose laid an egg. If Jack reports that Goose laid an egg, Jack will make a fixed payment to Banker and deposit the remaining funds from the sale of the egg into the bank account. If Jack reports that the goose did not lay an egg, Jack will make the same fixed payment to Banker out of the funds in the bank account. If the funds are not sufficient to make the payment, Jack will hand over all the funds remaining in the account to Banker. Banker will then draw lots: if Banker draws a long straw, he kills Goose; if he draws a short straw, he does not kill Goose and also tops up Jack’s bank account, returning its balance to the starting level. If Jack ever refuses to make payments he is able to make using the bank balance, Banker will seize the bank account and kill Goose. If Jack’s bank balance becomes sufficiently large to discharge Jack’s debt to the banker, Jack pays off Banker, the account is closed, and Banker exits.

Jack then solves a rather difficult mathematical problem of determining the required payment to Banker, the starting bank balance, and the ratio of short to long straws. He convinces Banker that under this scheme, even he, amoral Jack, will have an incentive not to divert the riches produced by Goose to personal consumption. Why? Whenever Jack secretly transforms the eggs into strawberries, his bank balance falls because of the required payment to Banker. The reduced bank account lowers the life expectancy of Goose, Jack’s only meal ticket. If Jack and Banker have their sums right, diversion will not be in Jack’s selfish interest.

## **4.2 Moral of the story**

Our story illustrates both the power and the limitations of contracting to prevent diversion. On the one hand, even when agents are amoral and diversion is impossible to monitor or verify (and thus immune from legal sanction), contacts can prevent diversion. On the other hand, these sort of contractual solutions are problematic for three reasons. First, they are not efficient. In our example, there is a positive probability that the Banker will kill the goose that lays the golden eggs. Second, it is somewhat questionable whether this contract will be enforced. When the banker is about to draw the knife to Goose’s neck, he realizes that the scheme ensured that Jack was a victim of bad luck who never actually diverted funds and that killing the goose is pointless. One might argue that Jack could at this point propose to Banker that he simply restart the scheme, which would benefit both Jack and Banker. If a

restart were allowed, amoral Jack would anticipate this from the start and thus he would have no incentive not to divert. However, even if we are willing to bear the inefficiency and can find a mechanism to enforce the contracted slaughter, there is a third difficulty. We need a goose, that is, we need a fixed asset with a return pattern that Jack cannot transform. This fixed tangible asset is the hostage required to coerce honest behavior from an immoral agent. If Goose itself could be transformed, the problem of contractual control becomes hopeless. If, for example, Jack could dramatically accelerate Goose's egg current egg production at the cost of losing future production, Jack would simply speed up production, consume the eggs and let Banker seize the now worthless goose.

Even if compensation cannot play an important role when diversion of corporate wealth cannot be verified, perhaps compensation can play a role when opportunities for diversion are more limited. For example consider the case where revenue is verifiable and thus not divertible but investment is not verifiable and thus divertible. Under this assumption, investors can see the revenue from the project they finance but not the investment that went into producing the revenue. An example might be a farmer who invests in seeds and fertilizer to produce a harvest. Investors in the farm can see the harvest and thus can verify its size and value but they cannot divine how much of the farmer's budget was actually invested in the harvest and how much was diverted to the farmer's personal consumption. Another example would be limited partners providing financing for a motion picture. The revenue from the picture might be fairly easy to verify based on box office receipts, but discriminating between expenditures which contributed to making the picture and expenditures which were simply personal perk consumption by the directors and producers might be very difficult. In these situations, reasonably effective compensation mechanisms can be designed (Noe, 2009). However, while it seems quite plausible in the farm and movie examples that cash output is easier to measure than cash input, the input/output distinction makes little sense in a financial asset setting where managers transform one asset into another through trade and the assets traded do not differ to any appreciable degree with respect to tangibility, visibility, or verifiability.

In short, the defining characteristics of financial assets are mutability and intangibility. Financial asset managers are hired to manipulate the return structure of the assets they manage, hence, forbidding asset transformation would defeat the very purpose of asset management. Manipulating the return structure requires little more than keystrokes entered on a computer terminal. Any control system so weak to allow diversion is not likely to be able to exercise control over the fine details of asset transformation. We conclude that paying unethical financial fiduciary agents is an even more hopeless task than paying them for vigilance or extraordinary performance.

## **5 Implications**

We have argued that incentive contracts are ill-suited to resolving the agency problems in financial institutions. Since these institutions do function, albeit imperfectly, there must some mechanisms at work that mitigate agency conflicts. Some of these mechanisms involve monitoring, which, like incentive compensation, acts to change the

extrinsic rewards and costs of associated with ethical behavior. Monitoring can occur through internal accounting controls, regulator surveillance, legal liability, or financial product regulation. The other lever for mitigation is affecting the intrinsic value agents place on being ethical, e.g., instilling a duty of care, or affecting the culture of banking. The path to better financial system will be built by retuning the mix of external and internal controls on financial agent behavior. A mix that will not, in our opinion, rely heavily on incentive contracting.

## References

- Clementi, G. L., Hopenhayn, H. A., 2006. A theory of financing constraints and firm dynamics. *The Quarterly Journal of Economics* 121 (1), pp. 229–265.  
URL <http://www.jstor.org/stable/25098789>
- Foster, D. P., Young, H. P., 2008. The not-so-real McCoy, institutional investor's alpha, University of Oxford working paper.
- Foster, D. P., Young, H. P., 2010. Gaming performance fees by portfolio managers. *The Quarterly Journal of Economics* 125 (4), 1435–1458.  
URL <http://qje.oxfordjournals.org/content/125/4/1435.abstract>
- Foster, D. P., Young, H. P., 2012. A strategy-proof test of portfolio returns. *Quantitative Finance* 12, 671–693.
- Grossman, S. J., Hart, O. D., 1983. An analysis of the principal-agent problem. *Econometrica* 51 (1), pp. 7–45.  
URL <http://www.jstor.org/stable/1912246>
- Harrison, G. W., Rutström, E. E., 2008. Risk aversion in the laboratory. In: *Research in Experimental Economics*. No. 12. Emerald Group Publishing, pp. 41 – 196.
- Hart, O., Moore, J., 1990. Property rights and the nature of the firm. *Journal of Political Economy* 98 (6), pp. 1119–1158.  
URL <http://www.jstor.org/stable/2937753>
- Hölmstrom, B., 1979. Moral hazard and observability. *The Bell Journal of Economics* 10 (1), pp. 74–91.  
URL <http://www.jstor.org/stable/3003320>
- Jensen, M. C., Meckling, W. H., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3 (4), 305 – 360.  
URL <http://www.sciencedirect.com/science/article/pii/0304405X7690026X>
- Noe, T. H., 2009. Tunnel-proofing the executive suite: Transparency, temptation, and the design of executive compensation. *Review of Financial Studies* 22 (12), 4849–4880.  
URL <http://rfs.oxfordjournals.org/content/22/12/4849.abstract>
- Taleb, N., 2005. *The black swan: Why don't we learn that we don't learn*. NY: Random House.