Essays in Household Finance and Corporate Finance

Author: Viktar Fedaseyeu

Boston College Electronic Thesis or Dissertation, 2011

Boston College
The Carroll Graduate School of Management
Department of Finance

ESSAYS IN HOUSEHOLD FINANCE AND CORPORATE FINANCE

a dissertation
by

VIKTAR FEDASEYEU

submitted in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy in Finance

May 2011

ESSAYS IN HOUSEHOLD FINANCE AND CORPORATE FINANCE

by

Viktar Fedaseyeu

In the fist two essays of this dissertation, I examine the role of third-party debt collectors in consumer credit markets. First, using law enforcement as an instrument, I find that higher density of debt collectors increases the supply of unsecured credit. The estimated elasticity of the average credit card balance with respect to the number of debt collectors per capita is 0.49, the elasticity of the average balance on non-credit card unsecured loans with respect to the number of debt collectors per capita is 1.32.

In the second essay I investigate the economics of the debt collection industry. The existence of third-party debt collection agencies cannot be explained by the benefits of specialization and economies of scale alone. Rather, the debt collection industry can serve as a coordination mechanism between creditors. If a debt collection agency collects on behalf of several creditors, the practices it uses will be associated will all creditors that hired it. Hence, consumers will be unable to punish individual creditors for using harsh practices. As a result, the third-party agency may use harsher debt collection practices than individual creditors collecting on their own. As long as the costs of hiring third-party debt collectors are below the benefits from using harsh debt collection practices, the debt collection industry will create economic value for creditors.

The last essay, written jointly with Thomas Chemmanur, develops a theory of corporate boards and their role in forcing CEO turnover. We show that in general the board faces a coordination problem, leading it to retain an incompetent CEO even when a majority of board members receive private signals indicating that she is of poor quality. We solve for the optimal board size, and show that it depends on various board and firm characteristics: one size does not fit all firms. We develop extensions to our basic model to analyze the optimal composition of the board between firm insiders and outsiders and the effect of board members observing imprecise public signals in addition to their private signals on board decision-making.

Contents

Chapter 1

Debt Collection Agencies and the Supply of Consumer Credit

Abstract

I examine the role of third-party debt collectors in consumer credit markets. Using law enforcement as an instrument for the number of debt collectors, I find that higher density of debt collectors increases the supply of unsecured credit. The estimated elasticity of the average credit card balance with respect to the number of debt collectors per capita is 0.49, the elasticity of the average balance on non-credit card unsecured loans with respect to the number of debt collectors per capita is 1.32. There is also some evidence that creditors substitute unsecured credit for secured credit when the number of debt collectors increases. Higher density of debt collectors improves recoveries, which enables lenders to extend more credit. Finally, creditors charge higher interest rates and lend to a larger pool of borrowers when the density of debt collectors increases, presumably because better collections enable them to extend credit to riskier applicants.

1.1 Introduction

Despite their large size, retail credit markets have received relatively little attention in the academic literature.[1] Even less effort has been devoted to studying the role of creditor rights in those markets. Aghion and Bolton (1992), Bolton and Scharfstein (1990), and Hart and Moore (1998) show that debt contracts are robust financial instruments if investors are assigned control rights contingent on debtors' payments. In retail credit markets, however, consumer protection laws restrict the range of options available to creditors. Providers of consumer credit never have full access to debtors' assets, and especially to their most valuable asset — human capital.[2] Even the threat of withdrawal of future financing from defaulting borrowers seems weak, as Cohen-Cole, Duygan-Bump, and Montoriol-Garriga (2009) document that consumers regain access to unsecured credit remarkably soon after filing for bankruptcy. In this paper, I examine a mechanism of creditor protection endemic to retail credit markets: third-party debt collectors. They ensure that defaulted debts will not go away easily, in effect enforcing creditor rights after default.

Consumer defaults have now reached historically high levels. The number of borrowers 120 days or more late on their payments approached 7 million people in 2009. While bankruptcy has been a topic of much debate among academics and regulators, consumer experience outside bankruptcy is also highly relevant. In 1999, the number of consumers with accounts in collections exceeded the number of consumers filing

[1] In the second quarter of 2009, the amount of consumer debt outstanding in the U.S., excluding loans secured by real estate, stood at $2.527 trillion, compared to $7.243 trillion in total nonfinancial corporate debt. Mortgage debt stood at $10.392 trillion. Source: http://www.federalreserve.gov/Releases/z1/20090917/z1.pdf, table D.3.

[2] This, however, has not always been the case. Debt prisons were common in the 19^{th} century: one of English literature's finest authors, Charles Dickens, immortalized this institution in his novel Little Dorrit (Charles Dickens' father and his entire family were held in a debt prison during the writer's childhood). In Ancient Rome and other slavery-based civilizations the borrower who defaulted could be sold into slavery, thus literally giving creditors full control over debtors after default.

for bankruptcy by the factor of 6. This ratio rose to 14:1 by 2009, likely due to the recession.

Stronger creditor protection should lead to more consumer credit, which is the primary hypothesis that I test. I find that a higher number of debt collectors per capita leads to an increase in balances on unsecured loans, but has no effect on secured loan balances. The estimated elasticity of the average credit card balance with respect to the number of debt collectors per capita is 0.49, the elasticity of the average balance on non-credit card unsecured loans with respect to the number of debt collectors per capita is 1.32. In addition, the pool of borrowers expands in response to higher debt collectors density, suggesting that creditors are willing to lend to riskier applicants. Consistent with this possibility, I document that creditors charge higher interest rates on unsecured non-credit card loans when debt collectors density increases.[3] I also show that higher debt collectors density is associated with higher recovery rates on delinquent credit card loans, which provides a direct mechanism behind my results on credit supply: better recoveries enable lenders to extend more credit in the first place.

Several empirical concerns arise in my analysis. First, it is difficult to separate supply from demand. Second, reverse causality can bias my estimates because the expansion in credit supply may lead to an increase in the number of debt collectors and not the other way around. Third, my variables are subject to the measurement error to the extent that I am unable to match debt collectors with the debts they are likely to collect. I use plausibly exogenous within state variation in the strength of law enforcement to instrument for the number of debt collectors in order to determine their

[3] Although I report results for credit card APRs as well, the effect on credit card interest rates cannot be identified because credit union call reports do not separate credit card fees from other types of fees. As Furletti (2003) documents, fees have become an increasingly important component of credit card pricing and can account for up to 50% of interest.

causal impact on the supply of consumer credit. I also use a falsification argument to strengthen my identification. Debt collectors are primarily engaged in collecting unsecured debts.[4] Hence, they should either have a negative effect on the amount of secured consumer credit (via the substitution effect) or no effect at all. If my results are attributed to spurious correlation, this spurious correlation is likely to affect both secured and unsecured credit similarly.

I use violent crime rate, the number of judicial employees per capita, and the median time interval from filing to disposition in federal civil cases as instruments. Debt collectors' compensation is customarily tied to the amount they collect and they are motivated to use all legal means available to them in order to collect the debts.[5] Thus, the quality of law enforcement should be inversely related to the effectiveness of debt collection. Crime rate and the number of judicial employees are direct metrics of law enforcement quality and quantity, respectively.[6] In accordance with the federal law that regulates debt collection, consumers subject to unfair and deceptive practices by debt collectors can file a lawsuit in a federal district court. The median time interval from filing to disposition is a measure of how quickly federal courts deal with incoming civil cases.[7] Most contract litigation is handled by state courts, and hence federal courts' statistics should not be contaminated with cases that directly affect the supply of credit. The strength of law enforcement matters only if the borrower defaults, at which point creditors turn to debt collectors. It is, therefore, unlikely

[4]Since secured creditors can repossess collateral, they do not rely on debt collectors. One rare exception to this rule is the situation when the value of collateral falls below the amount outstanding on the loan *and* the creditor decides to collect the difference.

[5]Debt collectors in Spain, for example, use public humiliation to extract payments from defaulting consumers. Such tactics are illegal in the U.S. See Thomas Catan, "Spain's showy debt collectors wear a tux, collect the bucks — their goal: Publicly humiliate non-payers." *The Wall Street Journal*, page A1, October 11, 2008.

[6]Since property crime can directly affect the demand for consumer credit, only violent crime is used as an instrument.

[7]Lawsuits bring uncertainty, and debt collection agencies prefer to dispose of them quickly.

that the above metrics can directly affect credit supply or demand in ways other than via debt collectors.

State-by-state data on the amount of consumer credit are unavailable from commercial banks that do business nationwide. To solve this problem, I use credit union Call Reports. By law, credit unions are allowed to lend only to their members, who must have a well-defined common bond (employer, location, or profession). Hence, credit unions are likely to be local credit providers. As of July 2009, credit unions provided 10% as much revolving credit as commercial banks and 42% as much unsecured non-revolving credit. Thus, they represent a significant share of all unsecured lending in the United States. Since credit unions are membership-owned organizations, they are likely to retain close contact with their borrowers.[8] Hence, debt collection effectiveness should matter less for them than for nationwide financial institutions. In order to provide evidence of the general applicability of my findings, I obtained qualitatively similar, although weaker, results by using call report data for small banks.

The rest of this paper is organized as follows. Section 1.2 reviews related literature. Section 1.3 provides some institutional details about the debt collection industry. Section 1.4 describes the data, estimation strategy, and empirical results. Section 2.6 concludes.

1.2 Relation to existing literature

In contrast to the large corporate finance literature on investor and creditor rights that followed La Porta, Lopez-de Silanes, Shleifer, and Vishny (1998), there has been little work on lender rights in retail credit markets. Hunt (2007) gives an overview of

[8]Credit union borrowers are a subset of credit union members.

the debt collection industry and provides details about its institutional structure and regulatory environment. Hynes (2008) examines the process of debt collection in state courts and finds that consumers who are sued by creditors or debt collectors are drawn from areas with lower socio-economic characteristics. Moreover, he finds that these consumers are not likely to file for bankruptcy. Hynes, Dawsey, and Ausubel (2009) show that states with anti-harassment statutes that apply to creditors collecting their own debts have lower bankruptcy filing rates, but borrowers living in these states are more likely to default without filing for bankruptcy.

This paper belongs to the growing literature on household finance. Campbell (2006) delineates the field. He finds that many households make effective investment decisions while a less educated minority make significant mistakes. Tufano (2009) gives a recent overview of this area and proposes its functional definition. There exists a public policy concern that lower income less educated households are being underserved by the banking system and have to resort to alternative financial services providers to meet their credit and transaction demand. Caskey (1991) is the first academic study of pawnbroking in the United States. That work, as well as Caskey (1994) and Caskey (2005) gave rise to a substantial literature whose current primary focus is the study of payday lending. The issue of whether short-term high-interest loans are welfare enhancing or not is one of the central topics in the household finance literature. Melzer (2009) finds that access to payday loans does not seem to alleviate financial hardship, while Morse (2009) provides evidence that payday lending mitigates individual financial distress. Flannery and Samolyk (2005) study the payday loans industry by using proprietary store-level data and find that high interest rates are generally justified by high fixed costs; they find no evidence that loans from frequent borrowers are more profitable than other loans per se. Morgan and Strain (2007) find that bans on payday lending in Georgia and North Carolina led to a a deterioration in households' financial situation in those states. Similarly, Zinman (2009) documents a deterioration in the overall financial condition of Oregon households after this state capped interest rates on payday loans. Karlan and Zinman (2009) use a powerful field experiment to demonstrate that expanding credit access resulted in significant net benefits for borrowers across a range of outcomes. The current paper complements this literature by studying a mechanism that enables traditional financial services providers to extend credit to risky borrowers.

Another strain of active literature in household finance studies personal bankruptcy, with the emphasis on explaining the rising rates of personal bankruptcy filings over the last two decades and on the effect of bankruptcy law on credit availability. Fay, Hurst, and White (2002) and Domowitz and Sartain (1999) find support for the strategic model of bankruptcy, which predicts that households are likely to file when their financial benefit from doing so is high. Gross and Souleles (2002) document that propensity to file for bankruptcy significantly increased from 1995 to 1997, even after controlling for a variety of personal risk characteristics and interpret this result as an increase in the borrowers' willingness to default. Dick and Lehnert (2010) show that the expansion of credit supply over time is responsible for rising personal bankruptcy rates, an explanation that was suggested by White (2007). Scott and Smith (1986) document that the Bankruptcy Reform Act of 1978, which made personal bankruptcy more pro-debtor, led to an increase in the contract interest rates on small business loans. Gropp, Scholz, and White (1997) find that generous state-level personal bankruptcy exemptions increase the amount of credit held by high-asset households and reduce the availability of credit for low-asset households. Debt collectors, the focus of this paper, provide a creditor protection mechanism, which complements bankruptcy as a consumer protection mechanism (at least in the U.S.). It is interesting to contrast the ways these mechanisms affect credit availability. While bankruptcy shifts credit supply toward more affluent households, it is strong creditor rights that enable lenders to provide loans to risky borrowers.

1.3 Industry overview

The size of the debt collection industry is significant. ACA International, an industry association of third-party debt collection agencies, commissions PriceWaterhouseCoopers to conduct annual surveys of the industry. According to the latest survey available, the total amount collected in 2007 was $40.4 billion, which represented nearly 21% of private sector bad debt for that year.[9] This compares with a total of $44 billion in payday loans extended in 2007 and around $75 billion in student loans for 2008-2009.[10] There are nearly 6,500 collection agencies in the U.S.[11] As of May 2009, they employed 107,340 debt collectors.[12]

Debt collectors play an active role in retail credit markets. They contact millions of American consumers every year. According to the Federal Trade Commission (FTC hereafter), which tracks consumer complaints, third-party debt collectors generate more complaints than any other industry. In 2008 the FTC received 78,838 complaints about third-party debt collectors, which represents 18.9% of all complaints received directly from consumers in 2008.[13] Thus, debt collectors are a very visible presence in the lives of American households.[14] Creditors turn to collectors after

[9] Source: http://www.acainternational.org.

[10] Source: Economic impact of the payday lending industry. IHS Global Insight (USA) Inc., 2009, and the Department of Education.

[11] Source: http://www.acainternational.org/publications-collections-information-5431.aspx

[12] According to Occupational Employment Statistics, the total number of bill and account collectors stood at 403,111 in May 2009, but this number includes collectors employed by creditors directly (in their in-house collection departments). The figure reported above only includes debt collectors working in the business support services industries, i.e., third-party debt collectors. Source: http://www.bls.gov/oes/current/oes433011.htm

[13] Source: Annual Report 2009: Fair Debt Collection Practices Act. Federal Trade Commission, Washington, DC, February 2009.

[14] According to InsideARM.com, there are several movie projects under way that feature debt

a loan has been in default for a certain period of time (usually after 180 days for credit card loans). Most debt collection agencies work on commission. According to ACA International's benchmarking survey, this commission usually constitutes 25-30% of the total amount they collect for the creditor.[15] The collection process is a human-intensive effort that requires debt collectors to constantly communicate with consumers. This communication is usually established over the telephone and by mail. Sometimes collection may require personal face-to-face contact but such cases are rare.

Debt collection in the United States is regulated by the federal law, the Fair Debt Collection Practices Act of 1977 (FDCPA hereafter). It supersedes state laws if those laws provide weaker protection for consumers. Forty-three states have their own laws that regulate collection practices. The majority of these statutes pre-date the FDCPA and provide very similar protection. As a result, most cases brought against debt collectors are tried in federal courts in accordance with the FDCPA. However, some states offer stronger protection to their residents than others.[16] In addition, state laws are important because most of them contain licensing requirements: debts owed in a particular state can usually be collected only by an agency licensed in that state.

The industry is geographically disperse and segmented at the state level, which provides the basis for my identifying assumption that collectors in a particular state are more likely to pursue debtors who reside in that state. In any given year between

collectors.

[15]This estimate is indirectly supported by data from the Census Bureau. Net revenues of the debt collection industry equaled $11.4 billion in 2005, which does not include the amount collected and returned to creditors. Net revenues were roughly 28% of the amount collected as reported in the PriceWaterhouseCoopers survey for 2007 (the data for 2005 are unavailable).

[16]These laws change infrequently: between 1987 and 2008 there appear to have been no more than four significant changes in state law concerning debt collection. Hence, controlling for the state fixed effect should account for the cross-sectional variation in state laws.

1988 and 2007, there has been no state without a debt collection establishment. The highest concentration of debt collection establishments was in California in 1989, with 11% of all U.S. debt collection establishments (it was 10.5% in 2007). In 2007, thirty-two states had the number of debt collectors of at least 1% of the total number of debt collectors in the U.S. Although some collection agencies are large public corporations with nationwide operations,[17] 99% of collection agencies have fewer than 250 employees. Top 17 firms in this industry receive less than 20% of all industry revenues.[18]

My identifying assumption may be invalid if debt collection is outsourced, either to several large call centers in the U.S. or abroad. However, outsourcing in the collection industry has had limited success.[19] Since negotiation is an important part of the collection process, collectors living in the U.S. are generally more effective than those from abroad. They are likely to be even more effective if they are familiar with the economic situation in the region where the debtor resides. As a result, most collection efforts that concern American consumers are performed by U.S. based collectors. In addition, state licensing requirements limit inter-state collection efforts. Within states, however, debt collection agencies are concentrated in low-cost metropolitan areas. Buffalo, NY, for example, is a hub for collection efforts in the state of New York. This suggests that individual states are the relevant unit of analysis.

---

[17] Asset Acceptance Capital Corporation, for instance, had market capitalization of $207.5 million on December 31, 2009. To the best of my knowledge, large national operators still maintain offices in several states.

[18] Source: the author's calculations based on Dun and Bradstreet adjustment and collections services industry report, http://www2.zapdata.com/IndustryReports. Accessed on January 11, 2010.

[19] Outsourcing certain information technology functions such as procedures for locating debtors, however, has been generally successful. Source: Operational Efficiency in the ARM Industry, Kaulkin Ginsberg Whitepaper, October 2006.

1.4 Debt collectors and consumer credit

1.4.1 Data

Data on payroll and employment of debt collection establishments from 1988 to 2007 are available from the Census Bureau's County Business Patterns Survey.[20] When for privacy reasons the survey contains ranges rather than point estimates, I replace ranges with midpoints.[21]

Debt collectors density is calculated by dividing the number of debt collectors in a particular state in a given year by the state's population. Since collection is a human-intensive process, a higher number of collectors per capita translates into a higher probability that a consumer will be contacted by a debt collector. In addition, debt collectors will have more time to negotiate with the consumer. Both of these factors should enhance collection efforts.

I use credit union Call Reports to obtain data on consumer credit in each state.[22] In contrast to national banks, credit unions are likely to be local lenders because credit unions are allowed to lend only to their members who must have a well-defined common bond (residence, employment or profession).[23] Call Reports are available from the National Credit Union Administration and cover the years from 1989 to 2008, although not all measures are available for all years.[24]

[20] A single debt collection agency can have several establishments in one or several states but the survey does not aggregate information at the agency (firm) level.

[21] When I drop observations that report ranges, the results remain qualitatively similar.

[22] I exclude Delaware and South Dakota from my analysis since these states provide incentives for credit card banks to operate on their territory. However, my results are not sensitive to the exclusion of these states. I also exclude data from credit unions who are likely to provide credit nationwide: the Navy Credit Union (and all credit unions of naval bases) and the Pentagon Federal Credit Union.

[23] In addition, credit unions can provide up to 12.5% of their assets in business loans. Those loans also tend to be local. However, the focus of this paper is consumer credit.

[24] I report results using second-quarter Call Reports to establish a correspondence between credit union data and County Business Patters surveys, which report data as of March of each year. I obtained similar results by using first- or third-quarter call reports instead.

Data on law enforcement and crime in the United States are obtained from the Survey of Public Employment and the Department of Justice's crime statistics. Data on federal courts' caseload statistics, which include median time intervals from filing to disposition in federal district courts, are available from annual reports of the Director of the Administrative Office of the United States Courts. Each state and the District of Columbia has at least one district court, with more populous states having a larger number of districts.[25] When there are several district courts in a single state, I compute an aggregate measure for the state by weighting the median time interval in each district by the corresponding number of cases.

Table 1.1 provides basic summary statistics for my sample.

[INSERT TABLE 1.1 ABOUT HERE]

1.4.2 OLS

The purpose of this section is to establish basic correlations using simple OLS regressions. For comparison, it may also be interesting to contrast OLS coefficients with the coefficients from instrumental variables regressions. OLS may raise understandable concerns due to reverse causality because debt collection agencies may anticipate the rise in consumer credit and start hiring to build additional capacity. I will address these concerns with instrumental variables estimation in the next section.

I estimate the following constant elasticity model:

$$\ln Y_{i,t} = \alpha_i + \gamma_t + \beta \ln X_{i,t-1} + \eta' \ln \text{Controls}_{i,t} + \varepsilon_{i,t}, \tag{1.1}$$

where $Y_{i,t}$ is a measure of the amount of credit in state i in year t, and $X_{i,t-1}$ is the debt collectors density, the number of debt collectors per million capita in state

[25] California, New York and Texas each have four federal judicial districts.

i in year $t-1$. I use lags of the main explanatory variable to rule out a possibly mechanical relationship: the number of debt collectors may increase as a response to more credit offered in the current period. I use income per capita to control for general economic conditions and a non-performing loans rate as a proxy for the riskiness of the pool of borrowers.[26] I also include credit union assets per member (defined as total credit union assets divided by the number of credit union members) to control for the relative affluence of credit unions: this is a measure of the total amount of credit that can potentially be extended. Three lags of per capita income growth in each state are also included. I hope that lags of personal income will absorb the demand-side variation and will account for the local business cycle. Time fixed effects are included to remove macro-level trends while state fixed effects eliminate all unobserved state heterogeneity.[27] In all specifications, standard errors are clustered at the state level. All variables are expressed in real 1982 dollars using CPI.

Debt collectors pressure consumers to pay back their debts and to the extent that this pressure is inconvenient for consumers, which it must be based on the number of consumer complaints, a higher number of debt collectors should reduce the demand for credit. Hence, any positive effect of debt collectors on the amount of credit in the simple OLS framework should be attributed to credit supply (strong enough to overcome the negative impact of lower demand). I will use several measures of the amount of consumer credit in my analysis. The first of them is the amount of loans extended per credit union member. Since credit unions are allowed to lend only to their members, credit union membership is the relevant demographic by which total

[26]Excluding non-performing loans does not affect the results in a significant way. Using state-level GDP instead of personal income does not affect the results. I prefer using personal income, however, because there is a discontinuity in the GDP-by-state time series at 1997, where the data change from SIC industry definitions to NAICS industry definitions.

[27]As I mentioned before, this should also account for the difference in consumer protection laws across states since those laws remained virtually unchanged during my sample period.

amount of credit should be normalized. My other measures help me explore whether changes in the amount of credit are attributable to the loan size (the intensive margin), the number of loans (the extensive margin), or both. To do this I look at loan balances (amount of credit divided by the number of loans) and at loans per member and per capita (number of loans divided by credit union membership and by the state's population, respectively). I consider the following four types of loans: credit cards, other unsecured consumer loans, auto loans (secured by old or new vehicles), and mortgages (first-lien).

Table 1.2 presents the results of estimating the effect of debt collectors on the amount of unsecured credit per credit union member.

[INSERT TABLE 1.2 ABOUT HERE]

As expected, debt collectors density has a significant positive effect on the amount of total unsecured credit and on non-credit cards loans. A one-percent increase in the debt collectors density in year $t-1$ leads to a 0.77% increase in the amount of credit card loans per credit union member and a 0.087% increase in the amount of non-credit card unsecured loans per credit union member in year t. This effect will be larger in the instrumental variables estimation. I attribute this fact to the following two consideration. First, it is difficult to separate supply from demand in the simple OLS framework, and since debt collectors should decrease demand, the coefficients on the total amount of credit should be smaller than the coefficients on the supply of credit. Second, the number of debt collectors should rise in accordance with *delinquent* credit, not the total amount of consumer credit. It is precisely in these circumstances that creditors should be unwilling to provide credit, ceteris paribus.

[INSERT TABLE 1.3 ABOUT HERE]

Debt collectors are usually employed to collect unsecured debt. In the case of a secured loan the creditor can repossess the underlying collateral.[28] As a falsification test, I regress the amount of secured credit on debt collectors density. These results are presented in Table 1.3. Changes in debt collectors density do not seem to affect the supply of auto loans and have a negative impact on the supply of mortgage loans. This is evidence of a substitution effect: when the amount of unsecured credit increases the amount of secured credit falls (notice that my regressions include total assets per member, with the intention to control for the total amount of credit that can potentially be extended). This is evidence that demand-side omitted variables are not driving my results since those variables are likely to affect all types of loans similarly.[29]

1.4.3 Instrumental variables estimation

Debt collectors' compensation is usually tied to the amount of collections they generate. Therefore, they have strong incentives to be persistent. Sometimes the methods they use are on the borderline of legality. In many cases brought by the FTC and state Attorney Generals against debt collection agencies, the latter were found guilty of using abusive practices prohibited by the federal law. In testimony before the Consumer Affairs Subcommittee of the U.S. House Committee on Banking and Housing during 1992 oversight hearings on the Fair Debt Collection Practices Act, Richard W. Bell, a former collector, testified that abusive strategies were routine among the more

[28] In the case of auto loans the collateral can be relocated by the consumer and its repossession by the creditor may be complicated. In those instances they use repossession agencies (repo men as they are known colloquially). Those agencies are separate from debt collectors that are the focus of this paper. County Business Patterns surveys track these two types of establishments in separate categories.

[29] It may be the case, however, that recessions shift demand from secured to unsecured credit. This concern is mitigated to the extent that time fixed effects control for the nationwide business cycle and lags of personal income in each state account for business cycles at the state level.

than nineteen collection companies for which he worked over a ten-year period in Texas.[30] Bell testified that common abusive (and illegal) collection tactics included:

- Phoning a debtor's parent, impersonating a government prosecutor, and requesting the parent to get the debtor to call about a criminal investigation regarding the debtor.
- Threatening the debtor and his parent with criminal charges for capital gains tax fraud unless the balance of the debt was put on the parent's credit card.
- Calling 5-15 neighbors in a brief period of time, informing them that the debtor was suspected of receiving stolen goods, and asking them to go to the debtor's home and request the debtor to call the collector. This was called a "block party." A variant was to hold an "office party" by calling fellow employees.
- Soliciting postdated checks in order to later threaten criminal bad check prosecution.
- Threatening to report Latinos to immigration authorities and posing as an immigration officer.
- Encouraging women to engage in prostitution and men to sell drugs to pay a debt.

Collection efforts are more effective when law enforcement is less effective, although I am confident that most collectors use lawful means to collect consumer debts. Two facts support this statement. First, the FTC receives more complaints about third-party debt collectors than about any other industry. Second, the amount of civil litigation against debt collectors is significant. In the first five months of 2010,

[30]Source: Fair Debt Collection.

there were 4,808 lawsuits filed by consumers against debt collection agencies,[31] which compares with 185,900 original civil cases filed in the U.S. District Courts in 2009.[32] The threat of consumer litigation is so serious that there exists a specialized agency, WebRecon LLC, which tracks consumers and lawsuits in order to determine who the most litigious consumers are. This information is then used by collection agencies to determine what course of action is most appropriate should one of their debtors be on the list of repeat filers.

I use the following measures of law enforcement effectiveness as instruments: violent crime rate, the number of law enforcement personnel, and the median time from filing to disposition in federal civil cases. If law enforcement is pre-occupied with crime, collectors are more likely to get away with shady tactics. Thus, the number of debt collectors should be positively related to the crime rate and negatively related to the number of law enforcement employees. Total crime rate, although a significant predictor of the number of debt collectors, may also be correlated with the demand for credit. For example, people who have had something stolen may need to increase spending on their credit cards or obtain an additional consumer loan. In order to address this concern I distinguish between violent and property crime,[33] with the idea that property crime should pick up the demand side variation. I believe that violent crime per se should be unrelated to the demand for credit.[34]

[31]Source: WebRecon LLC, published by InsideArm.com. Of the 4,808 lawsuits, 4,099 were filed under the FDCPA, 419 – under the Fair Credit Reporting Act, 16 – under the Telephone Consumer Protection Act, 172 – under the Truth in Lending Act, with the remaining suits filed under various other federal acts and state consumer statutes.

[32]Source: Judicial Business of the United States Courts, 2009. The total number of civil filings in 2009 was 276,397, which also includes removals from state courts, remands from courts of appeals, reopens, and transfers.

[33]The following offenses are classified as violent: murder and non-negligent manslaughter, forcible rape, robbery, aggravated assault. Property offenses are: burglary – breaking or entering, larceny-theft, motor vehicle theft.

[34]The only possible channel is medical bills if a consumer gets injured. However, these bills are either covered by health insurance or are a separate form of unsecured credit, which is not the focus of this study.

Judicial employment indicates the likelihood that legal action will be brought against debt collectors. Actions by federal and state regulators are a major concern for the debt collection community.[35] Collection agencies are sued regularly by state Attorney Generals and consumers.[36] Conditional on being sued, however, debt collectors prefer to settle the matter as soon as possible.[37] Lawsuits bring uncertainty, and the extent of this uncertainty is high due to the potentially large penalties that can be imposed. On May 28, 2010, a jury in Texas awarded $1.5 million in punitive damages against a debt collection agency, in addition to $50,000 in mental anguish damages. The initial debt the agency was trying to collect was only $200.[38] In order to avoid such ruining outcomes, collection agencies prefer to settle cases. In addition, collection efforts concerning disputed accounts may have to be halted, which reduces revenues, at least temporarily. Most cases against debt collectors are filed in federal courts under the FDCPA, which includes provisions for civil action. Therefore, the length of federal civil proceedings is relevant for debt collectors. On the other hand, most contract cases are tried in state courts, so that federal civil proceedings are not contaminated by actions brought by creditors against consumers. I present the results of the first-stage instrumental variables estimation in Table 1.4.

[INSERT TABLE 1.4 ABOUT HERE]

[35] InsideARM.com, a leading on-line resource for debt collectors, regularly sends newsletters to its subscribers. In the first quarter of 2010, 59 newsletters were distributed, 30 of which discussed issues related to regulation, lawsuits involving collectors, and law enforcement matters.

[36] New York Attorney General Andrew M. Cuomo, for example, started a statewide initiative in May 2009 to clean up the debt collection industry. As of May 2010, his office shut down 14 debt collection companies and required others to reform their deceptive practices. 10 collectors were criminally prosecuted. Other recent actions against debt collectors were initiated by Attorney Generals in West Virginia and Colorado.

[37] As one collector from Florida put it on a discussion forum during Expo 3.0, an on-line conference of debt collectors, "we have to settle out of necessity."

[38] Allen Jones v. Advanced Call Center Technologies. Source: InsideArm.com.

As expected, violent crime rate is a significant predictor of the number of debt collectors. In addition, higher capacity of the state's judicial system measured by the per capita judicial employment leads to a lower number of collectors. The median time interval from filing to disposition in federal civil cases has a negative effect on the number of debt collectors, confirming the intuition that conditional on going to court debt collection agencies prefer to finish proceedings quickly.

[INSERT TABLE 1.5 ABOUT HERE]

I use instrumental variables to estimate the effect of debt collectors density on the total amount of credit (measured by the amount of loans per credit union member). Table 1.5 reports the results of this estimation. Again, I attribute the results that indicate credit expansion to credit supply: first, my instrumental variables are intended to pick up the supply side variation, second, as I mentioned before, debt collectors should reduce demand. Debt collectors density has a statistically significant effect on the amount of unsecured credit and a negative effect on secured credit, consistent with the substitution hypothesis. The coefficient on the variable of interest is now an order of magnitude larger than it was in simple OLS regressions. This finding is not surprising since the instrumental variable estimation is meant to purge other influences that contaminate OLS coefficients. For the reasons outlined above, I expect those influences to bias OLS coefficients downwards. Instrumental variables results are significant for both credit cards and other unsecured loans, unlike simple OLS results above. A one-percent rise in the debt collection capacity leads to an increase of 0.529% for credit card loans per credit union member and 1.67% for other unsecured loans per credit union member. It also leads to a 1.37% decrease in the amount of auto loans extended, per credit union member, and a 0.156% decrease in the amount of mortgage loans per credit union member. On the one hand, this is evidence that creditors substitute unsecured credit for secured credit when the collection process is more effective (the provision of unsecured credit becomes less expensive). On the other hand, secured creditor are also concerned about the value of their collateral. If law enforcement is weak and crime is rampant, property may be damaged or destroyed, thus directly affecting secured creditors. My instruments are unable to separate these effects and since both of them should reduce the supply of secured credit, the above results should be interpreted with care.

[INSERT TABLE 1.6 ABOUT HERE]

There are two ways in which credit expansion can occur. Credit unions can increase their exposure to the current customers in terms of loan balances or the number of loans, or they may extend credit to a larger number of people by attracting new members. I explore these mechanisms below.

Table 1.6 presents estimation results for loan balances (loan amounts divided by the number of loans). Debt collectors have a positive effect on loan balances for unsecured loans, with the effect on non-credit card debt being the strongest. There is no effect on secured loan balances, which is not surprising given that debt collectors should be irrelevant for secured credit other than via a possible substitution effect. A one-percent increase in the number of debt collectors per capita leads to an increase of 0.493% for credit cards balances and 1.322% for balances on other unsecured loans. These results indicate that credit unions increase their exposure to current customers in terms of the size of the loans they are willing to offer.

I also investigate whether credit unions increase the number of unsecured loans per member, with the results reported in Table 1.7. I find no significant effect on the number of unsecured loans per member, both for credit cards and other unsecured loans. Although credit unions are willing to increase loan sizes, they do not seem to raise the number of loans each member receives.

[INSERT TABLE 1.7 ABOUT HERE]

To study whether credit unions expand membership I look at the number of members per capita and the number of loans per capita in Table 1.8. I find that the number of credit union members per capita grows in response to higher debt collectors density. The number of unsecured non-credit card loans per capita also increases. While credit unions seem to keep the number of loans per member stable, the fact that their membership grows results in the increase in the number of loans per capita. From the policy standpoint, robust collections enable credit unions to offer credit to a larger number of people: debt collectors may help traditional credit providers expand their network.

[INSERT TABLE 1.8 ABOUT HERE]

The expansion of credit attributed to debt collectors should benefit riskier borrowers. Debt collectors provide an enforcement mechanism in case the borrower defaults. This mechanism should be more important if the borrower is risky since in this case the creditor is more likely to turn to debt collectors. It is likely that a larger pool of credit union borrowers associated with higher debt collectors density means that borrowers are becoming riskier. The following considerations speak in favor of this hypothesis. First, assuming that credit unions behave rationally, they should start by attracting the safest applicants. Second, since credit unions are membership organizations and are likely to maintain a close relationship with their members, they should be able to assess the relative riskiness of their potential members. Although I cannot observe the riskiness of the pool directly, I can look at interest rates charged on unsecured loans. If credit unions are willing to lend to riskier borrowers, they should charge higher interest rates. Table 1.9 reports coefficients from regressions of interest rates on debt collectors density.

[INSERT TABLE 1.9 ABOUT HERE]

Although my results on pricing are weak, they indicate that higher debt collectors density is associated with higher interest rates on unsecured non-credit card debt, consistent with the hypothesis that gains in credit union membership are coming from riskier borrowers. Effective collection alleviates credit rationing and ensures entrance into retail credit markets for borrowers who would otherwise be unable to participate. These results complement Gropp, Scholz, and White (1997) who show that generous bankruptcy exemptions are beneficial for the wealthiest consumers. The effect on credit card pricing in Table 1.9 is positive but insignificant. This latter coefficient should be interpreted with care, however, because of data availability issues. Fee income is an increasingly important element of credit card pricing and can account for up to 50% of interest. Furletti (2003) describes the trends in credit card pricing and shows that lenders have dramatically changed their pricing strategies since the 1990s. In particular, they reduced APRs and shifted to charging various fees to credit card borrowers. Late fee revenue, for example, quadrupled between 1996 and 2001. Credit unions report APRs but do not separate their fee income into credit card related and other fees. As a result, I am unable to develop a good measure of credit card interest rates in my sample.

1.4.4 Debt collectors, loan recoveries and charge-offs

The results presented above indicate that effective third-party debt collection increases the supply of unsecured credit. In this section I intend to delineate the direct mechanism behind this finding. Higher debt collectors density should be associated with higher recovery rates. Higher recovery rates, on the other hand, decrease creditors' losses conditional on default and make them more willing to lend in the first place. Naturally, this mechanism should be more important for borrowers who are ex ante more likely to default. Charge-off rates, on the other hand, should be lower when debt collectors density increases.

[INSERT TABLE 1.10 ABOUT HERE]

Table 1.10 presents estimation results from regressions of credit card charge-off and recovery rates on debt collectors density. I use credit card recoveries and charge-offs because debt collectors are relevant for unsecured credit (recoveries and charge-offs for other unsecured loans are unavailable). Debt collectors have a positive effect on credit card recoveries. These results indicate that a larger number of debt collectors enhances creditors' ability to collect delinquent debt and leads to the expansion of credit supply.

1.4.5 Robustness checks

Although credit unions are not insignificant, most consumer lending in the United States is provided by banks. In order to address the issue of external validity of my results I test whether debt collectors density has an impact on the amount of credit provided by small banks.[39] Since call reports do not disaggregate data by geographic location, it is difficult to obtain good measures of banks' credit availability at the state level. It seems, however, reasonable to assume that small banks are more likely than large banks to restrict their activities to a particular state, which is crucial in my setting since I use state-level data to obtain debt collectors density. Bank call

[39] I define small banks as banks with total assets below $1 billion, expressed in year 2000 dollars. The cut-off is recalculated each year, so that the same bank may be regarded as small in some years and as big in others.

reports distinguish between two types of non-mortgage consumer loans: credit cards (and related plans), and other loans. Since other loans include secured along with unsecured loans, I use credit card loans as a measure of the amount of unsecured credit provided by small banks. Auto loans are not tracked as a separate category (they are included in other consumer loans), which is why the only measure of secured credit I can use is real estate loans. I use home equity line of credit loans in my falsification tests. Since banks do not report the number of loans, I normalize the total amount of loans by the state's population.

[INSERT TABLE 1.11 ABOUT HERE]

Table 1.11 presents regressions that use data on small banks. The results are weaker than for credit unions. They indicate a positive impact of debt collectors density on the supply of unsecured credit and no effect on secured credit. The fact that the coefficient of interest is only marginally significant is attributed to the measurement error in the dependent variable, which in this case should only increase standard errors.[40] First, it is impossible to obtain a clean measure of state-by-state amount of credit from bank call reports, and it is especially difficult to do for credit card loans. Second, unlike for credit unions, the relevant demographic to which small banks provide credit is difficult to determine. Small banks are unlikely to be able to serve the entire state's population, so that the denominator of my measure of the amount of credit introduces an additional measurement error.

My next test looks at state laws that regulate debt collection activities. Unfortunately, those laws change very infrequently (there have been no more than four significant changes during my sample period). Hence, I am only able to exploit the cross-sectional variation between states in terms of how many restrictions they place

[40]This measurement error should not bias the coefficients, however, unlike a measurement error in the regressors, which I attempted to deal with via instrumental variables estimation.

on third-party debt collectors. Table 1.12 reports the results. The index of state laws is the total number of restrictions that a state puts on third-party debt collection (licensing requirements, branching restrictions, record keeping requirement, the list of prohibited practices etc.). Consistent with the overall results of this paper, more stringent regulation of third-party debt collectors reduces the supply of unsecured credit and has no effect on secured credit.

[INSERT TABLE 1.12 ABOUT HERE]

As an additional robustness check, I look at the role of repossession professional in the supply of consumer credit. Repossession agencies are involved in locating vehicles that have been pledged as collateral for secured auto loans and returning them to creditors. Since those agencies are involved in enforcing secured credit rights, they should have no effect on unsecured credit. In fact, the correlation between the number of debt collectors per capita and the number of repossession professionals per capita is -5% (when time and state fixed effects are eliminated). Adding the number of repossession professionals to the regressions presented above does not change the results. Table 1.13 reports the results of adding the number of repossession professionals to the regressions of loan balances on the number of debt collectors.

[INSERT TABLE 1.13 ABOUT HERE]

However, repossession professionals may have an independent effect on the supply of secured credit. Even though I have not been able to devise a convincing instrumental variables strategy to test this hypothesis, Table 1.14 reports results from OLS regressions of loan balances on the number of repossession professionals.

[INSERT TABLE 1.14 ABOUT HERE]

Lastly, I can use a different measure of debt collectors' harshness. I have obtained from Don Morgan of the Federal Reserve Bank of New York some data on the number of complaints against third-party debt collectors. This is a direct measure of how aggressive debt collectors are in recovering the debts they have been assigned. Estimation results from adding this new variable to the regressions are reported in Table 1.15

[INSERT TABLE 1.15 ABOUT HERE]

1.5 Conclusion

Using plausibly exogenous within state variation in the strength of law enforcement as an instrument for the number of debt collectors, I find that higher debt collectors density increases the supply of unsecured consumer credit, both in terms of the amount of credit per credit union member and the size of the loan. A one-percent change in the number of debt collectors per capita leads to a 0.49% change in the average credit card balance and a 1.32% change in the average balance on non-credit card unsecured loans. Consistent with the fact that debt collectors collect unsecured debts, there is no effect on the size of secured loans. In terms of secured loans per credit union member, however, I find weak evidence that creditors substitute unsecured credit for secured credit when the number of debt collectors increases. The number of credit union members grows with higher debt collectors density. Accordingly, the number of unsecured non-credit card loans per capita also increases.

Increased membership indicates that creditors lend to a larger pool of applicants when debt collectors density rises. These additional borrowers are likely to be riskier consumers with lower incomes. Consistent with this, creditors charge higher interest rates on unsecured non-credit card loans. I am unable to identify the effect on credit card interest rates because I cannot account for the fee income, which constitutes an increasingly important fraction of interest. I also show that higher debt collectors density is associated with higher credit card recovery rates, which provides a direct mechanism behind my results on credit supply.

My findings indicate the importance of lender protection in retail credit markets. While generous bankruptcy exemptions benefit more affluent consumers, effective debt collection enables creditors to lend to riskier, presumably lower income borrowers. Financial regulation aimed at consumer protection must be balanced with strong creditor rights to achieve the goal of expanding credit supply to the underserved populations.

Table 1.1: Summary statistics

Variable	Mean	Median	St. dev.
Debt collectors per million capita	331.01	285.55	203.44
Total amount of unsecured credit per member	$280.81	$271.20	$98.96
Amount of credit card loans extended, per member	$99.42	$96.98	$63.55
Average balance on credit card loans	$804.39	$780.57	$185.11
Average interest rate on credit card loans	12.00%	12.12%	1.36%
Amount of other unsecured loans extended, per member	$137.18	$127.73	$87.91
Average balance on other unsecured loans	$1,251.54	$1,210.58	$388.49
Average interest rate on other unsecured loans	12.62%	12.64%	1.04%
Amount of secured auto loans extended, per member	$764.14	$742.69	$258.28
Average balance on secured auto loans	$5,196.72	$5,188.78	$797.30
Average interest rate on secured auto loans	7.40%	7.61%	1.20%
Amount of mortgage loans extended, per member	$551.19	$434.39	$371.27
Average balance on mortgage loans	$39,317.24	$35,227.83	$17,731.44
Average interest rate on mortgage loans	7.29%	7.27%	1.90%

Summary statistics for the entire sample period, 1988-2009, in real 1982 dollars. Not all variables have observations in every year. The number of debt collectors is for the 1988-2007 period, various loan statistics are for the 1989-2009 period. The amount of credit per member is obtained by dividing the dollar amount of a particular type of loan by the number of credit union members. Other unsecured loans include all non-credit card unsecured loans (big ticket purchases, unsecured home improvement loans, etc.). Prior to 1992, call reports did not distinguish between credit cards and other unsecured loans. All variables are obtained by aggregating credit union data at the state level.

Table 1.2: OLS regressions of unsecured credit supply on debt collectors density, 1992-2008

Variable	*ln*(Amount of credit card loans extended, per member), t	*ln*(Amount of other unsecured loans extended, per member), t
ln(Debt collectors per million capita), $t-1$	0.077* (1.66)	0.087** (2.21)
ln(Total non-performing loans rate), t	-0.028 (-0.61)	0.069** (2.32)
ln(Total non-performing loans rate), $t-1$	0.060 (0.53)	-0.062 (-1.21)
ln(Assets per member), t	0.030 (0.69)	0.004 (0.13)
ln(Income per capita), t	-0.083 (-0.02)	4.334 (1.03)
ln(Income growth), from $t-1$ to t	-0.001 (-0.18)	0.002 (0.40)
ln(Income growth), from $t-2$ to $t-1$	-0.001 (-0.29)	0.0002 (0.05)
ln(Income growth), from $t-3$ to $t-2$	0.001 (0.37)	-0.006 (-1.56)
Year fixed effects	YES	YES
State fixed effects	YES	YES
No. of obs.	816	816
Adjusted R^2	0.87	0.94

All regressions use data aggregated at the state level. Amounts of loans per member are obtained by dividing total real amount of a particular loan type from credit union call reports by the total number of credit union members. Standard errors are clustered at the state level, t-statistics are reported in parentheses below the coefficients.

Table 1.3: OLS regressions of secured credit supply on debt collectors density, 1989-2008

Variable	ln(Amount of secured auto loans extended, per member), t	ln(Amount of mortgage loans extended, per member), t
ln(Debt collectors per million capita), $t-1$	-0.029 (-0.74)	-0.137*** (-2.82)
ln(Total non-performing loans rate), t	0.026 (0.56)	0.138*** (3.13)
ln(Total non-performing loans rate), $t-1$	-0.189*** (-4.26)	-0.014 (-0.30)
ln(Assets per member), t	0.036 (0.99)	0.169* (1.71)
ln(Income per capita), t	2.261 (0.63)	3.307 (0.72)
ln(Income growth), from $t-1$ to t	0.003 (0.68)	-0.009* (-1.65)
ln(Income growth), from $t-2$ to $t-1$	0.004 (0.90)	-0.009 (-1.54)
ln(Income growth), from $t-3$ to $t-2$	0.009** (2.34)	0.001 (0.27)
Year fixed effects	YES	YES
State fixed effects	YES	YES
No. of obs.	912	912
Adjusted R^2	0.91	0.93

All regressions use data aggregated at the state level. Amounts of loans per member are obtained by dividing total real amount of a particular loan type from credit union call reports by the total number of credit union members. Standard errors are clustered at the state level, t-statistics are reported in parentheses below the coefficients.

Table 1.4: First stage of instrumental variables estimation

Variable	ln(Debt collectors per million capita), $t-1$
ln(Judicial employment per million capita), $t-1$	-0.065** (-2.03)
ln(Median time from filing to disposition), $t-1$	-0.009* (-1.82)
ln(Violent crime rate per 100,000 capita), $t-1$	0.207*** (3.21)
Other controls	YES
Year fixed effects	YES
State fixed effects	YES
No. of obs.	910
Adjusted R^2	0.86

All regressions use data aggregated at the state level. Median time from filing to disposition was obtained from caseload statistics for federal district courts. If a state had more than one district court, the median for the state was calculated as a weighted average of district courts' medians using the number of cases as weights. Standard errors are clustered at the state level, t-statistics are reported in parentheses below the coefficients.

Table 1.5: IV regressions of unsecured and secured credit supply on debt collectors density, 1989-2008

Variable	ln(Amount of credit card loans extended, per member), t	ln(Amount of other unsecured loans extended, per member), t	ln(Amount of secured auto loans extended, per member), t	ln(Amount of mortgage loans extended, per member), t
ln(Debt collectors per million capita), $t-1$	0.529**	1.670***	-1.370**	-0.156*
	(1.97)	(3.07)	(-2.04)	(-1.88)
ln(Assets per member), t	0.014	-0.102	-0.008	0.158***
	(0.45)	(-0.22)	(-0.15)	(3.79)
ln(Income per capita), t	-2.099	-1.643	6.202	4.610
	(-0.58)	(-0.22)	(1.14)	(1.52)
ln(Property crime rate) per 100,000 capita, $t-1$	-0.693**	-1.22*	0.812	-0.262
	(-1.97)	(-1.60)	(1.29)	(-0.86)
Other controls	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
State fixed effects	YES	YES	YES	YES
No. of obs.	814	814	910	910
Adjusted R^2	0.86	0.94	0.90	0.93
Hansen J statistic (p-value)	2.547 (0.28)	0.225 (0.89)	0.959 (0.62)	1.988 (0.37)

All regressions use data aggregated at the state level. Amounts of loans per member are obtained by dividing total real amount of a particular loan type from credit union call reports by the total number of credit union members. Standard errors are clustered at the state level, z-statistics are reported in parentheses below the coefficients.

Table 1.6: IV regressions of unsecured and secured loan balances on debt collectors density, 1989-2008

Variable	ln(Average balance on credit card loans), t	ln(Average balance on other unsecured loans), t	ln(Average balance on secured auto loans), t	ln(Average balance on mortgage loans), t
ln(Debt collectors per million capita), $t-1$	0.493* (1.65)	1.322** (2.16)	-0.500 (-1.15)	0.646 (1.03)
ln(Assets per member), t	-0.005 (-0.19)	0.079* (1.71)	0.004 (0.19)	0.097** (2.28)
ln(Income per capita), t	-2.631 (0.77)	-6.455 (-1.08)	4.990** (2.00)	0.476 (0.14)
ln(Property crime rate per 100,000 capita), $t-1$	0.331 (0.97)	-0.904 (-1.37)	0.233 (0.91)	-0.291 (-0.75)
Other controls	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
State fixed effects	YES	YES	YES	YES
Instrumental variables	YES	YES	YES	YES
No. of obs.	814	814	910	910
Adjusted R^2	0.81	0.82	0.68	0.94
Hansen J statistic (p-value)	2.312 (0.31)	0.495 (0.78)	0.530 (0.77)	3.321 (0.19)

All regressions use data aggregated at the state-year level. Amounts of loans per member are obtained by dividing total real amount of a particular loan type from credit union call reports by the total number of credit union members. Standard errors are clustered at the state level, z-statistics are reported in parentheses below the coefficients.

Table 1.7: IV regressions of the number of unsecured loans per credit union member on debt collectors density, 1992-2008

Variable	ln(Number of credit card loans per 1000 members), t	ln(Number of other unsecured loans per 1000 members), t
ln(Debt collectors per million capita), $t-1$	1.023 (1.34)	0.349 (0.81)
ln(Total non-performing loans rate), t	-0.020 (-0.20)	0.037 (0.97)
ln(Total non-performing loans rate), $t-1$	0.094 (0.79)	0.056 (1.15)
ln(Assets per member), t	0.019 (0.36)	0.065** (2.24)
ln(Income per capita), t	-4.730 (-0.85)	4.811* (1.74)
ln(Income growth), from $t-1$ to t	-0.002 (-0.23)	-0.002 (-0.38)
ln(Income growth), from $t-2$ to $t-1$	-0.004 (-0.39)	0.003 (0.59)
ln(Income growth), from $t-3$ to $t-2$	-0.012 (-1.12)	-0.009* (-1.64)
ln(Property crime rate per 100,000 capita), $t-1$	-1.023* (-1.83)	-0.317 (-1.02)
Year fixed effects	YES	YES
State fixed effects	YES	YES
No. of obs.	814	814
Adjusted R^2	0.52	0.72
Hansen J statistic (p-value)	2.259 (0.32)	1.347 (0.51)

All regressions use data aggregated at the state level. Number of loans per member is obtained by dividing total number of loans of a particular loan type from credit union call reports by the total number of credit union members. Standard errors are clustered at the state level, z-statistics are reported in parentheses below the coefficients.

Table 1.8: IV regressions of the number of unsecured loans and credit union membership per capita on debt collectors density, 1989-2008

Variable	ln(Number of credit card loans per million capita), t	ln(Number of other unsecured loans per million capita), t	ln(Credit union membership per 1000 capita), t
ln(Debt collectors per million capita), $t-1$	1.362 (1.34)	0.687** (2.09)	0.144** (2.25)
ln(Total non-preforming loans rate), t	0.0093 (0.07)	0.066 (1.11)	0.043 (1.30)
ln(Total non-preforming loans rate), $t-1$	0.062 (0.41)	0.023 (0.32)	-0.035 (-1.38)
ln(Assets per member), t	0.041 (0.59)	0.087** (2.43)	0.026 (1.15)
ln(Income per capita), t	-2.864 (-0.41)	6.677* (1.65)	3.994*** (3.68)
ln(Income growth), from $t-1$ to t	0.001 (-0.05)	-0.0002 (-0.03)	0.0003 (0.10)
ln(Income growth), from $t-2$ to $t-1$	-0.001 (-0.08)	0.005 (0.82)	0.001 (0.30)
ln(Income growth), from $t-3$ to $t-2$	-0.014 (-1.04)	-0.011 (-1.50)	-0.002 (-0.87)
ln(Property crime rate per 100,000 capita), $t-1$	-1.137 (-1.52)	-0.431 (-1.02)	0.035 (0.26)
Year fixed effects	YES	YES	YES
State fixed effects	YES	YES	YES
No. of obs.	814	814	910
Adjusted R^2	0.45	0.69	0.96
Hansen J statistic (p-value)	1.292 (0.52)	2.553 (0.28)	2.457 (0.29)

All regressions use data aggregated at the state level. Number of loans per million capita is obtained by dividing total number of loans of a particular loan type from credit union call reports by the state's total population. Standard errors are clustered at the state level, z-statistics are reported in parentheses below the coefficients.

Table 1.9: IV regressions of interest rates on debt collectors density, 1992-2008

Variable	ln(Average interest rate on credit card loans), t	ln(Average interest rate on other unsecured loans), t
ln(Debt collectors per million capita), $t-1$	0.036 (0.50)	0.109** (2.22)
ln(Total non-performing loans rate, t	0.008 (1.13)	0.009 (0.99)
ln(Total non-performing loans rate), $t-1$	-0.013 (-1.52)	0.009 (0.82)
ln(Assets per member), t	0.010** (2.10)	0.001 (0.09)
ln(Income per capita), t	0.287 (0.63)	-0.134 (-0.25)
ln(Income growth), from $t-1$ to t	-0.001* (-1.86)	0.0001 (0.14)
ln(Income growth), from $t-2$ to $t-1$	-0.002*** (-3.57)	-0.0003 (-0.33)
ln(Income growth), from $t-3$ to $t-2$	-0.002** (-2.19)	-0.002** (-2.06)
ln(Property crime rate per 100,000 capita), $t-1$	-0.040 (-0.82)	-0.067 (-1.13)
Year fixed effects	YES	YES
State fixed effects	YES	YES
Instrumental variables	YES	YES
No. of obs.	814	814
Adjusted R^2	0.92	0.85
Hansen J statistic (p-value)	3.494 (0.17)	3.311 (0.19)

All regressions use data aggregated at the state level. Average interest rates on a particular loan type were obtained by averaging interest rates reported by individual credit unions. Standard errors are clustered at the state level, z-statistics are reported in parentheses below the coefficients.

Table 1.10: IV regressions of credit card charge-off and recovery rates on debt collectors density, 1998-2008

Variable	ln(Average charge-off rate on credit card loans), t	ln(Average recovery rate on credit card loans), t
ln(Debt collectors per million capita), $t-1$	0.031 (0.10)	0.571* (1.84)
ln(Total non-performing loans rate), t	0.183*** (3.65)	-0.489*** (-4.04)
ln(Total non-performing loans rate), $t-1$	0.281*** (4.89)	-0.022 (-0.17)
ln(Assets per member), t	-0.479** (-1.96)	0.558 (1.13)
ln(Income per capita), t	-11.609** (-2.16)	9.607 (-3.16)
ln(Income growth), from $t-1$ to t	0.012 (1.63)	0.004 (0.31)
ln(Income growth), from $t-2$ to $t-1$	-0.008 (-1.04)	0.011 (0.66)
ln(Income growth), from $t-3$ to $t-2$	-0.005 (-0.87)	0.009 (0.78)
ln(Property crime rate per 100,000 capita), $t-1$	-0.154 (-0.66)	0.629 (1.36)
Year fixed effects	YES	YES
State fixed effects	YES	YES
Instrumental variables	YES	YES
No. of obs.	526	526
Adjusted R^2	0.73	0.63
Hansen J statistic (p-value)	0.385 (0.82)	2.668 (0.26)

All regressions use data aggregated at the state level. Standard errors are clustered at the state level, z-statistics are reported in parentheses below the coefficients.

Table 1.11: OLS and IV regressions of small banks' credit supply on debt collectors density, 1989-2008

Variable	ln(Amount of credit card loans extended, per capita), t		ln(Amount of HELOC loans extended, per capita), t	
ln(Debt collectors per million capita), $t-1$	0.273*	0.585*	-0.092	5.381
	(1.67)	(1.80)	(-1.05)	(1.15)
ln(Bank assets per capita), t	2.261***	2.299***	1.276***	1.398***
	(5.29)	(3.97)	(8.31)	(4.13)
ln(Income per capita), t	-28.807*	-30.511	23.361	5.854
	(-1.64)	(-1.03)	(1.49)	(0.25)
ln(Income growth), from $t-1$ to t	0.016	0.013	-0.031**	-0.087
	(0.76)	(0.37)	(-2.33)	(-1.39)
ln(Income growth), from $t-2$ to $t-1$	0.016	0.017	-0.023*	-0.037
	(0.78)	(0.72)	(-1.82)	(-0.90)
ln(Income growth), from $t-3$ to $t-2$	0.009	0.010	0.005	-0.014
	(0.49)	(0.53)	(0.41)	(-0.34)
Year fixed effects	YES	YES	YES	YES
State fixed effects	YES	YES	YES	YES
Instrumental variables	NO	YES	NO	YES
No. of obs.	372	370	960	958
Adjusted R^2	0.92	0.92	0.90	0.84

HELOC stands for Home Equity Lines of Credit, loans extended with equity in the house as the collateral. Data for HELOC loans are available since 1989, data for credit card loans are available since 2001 (prior to that year banks reported all installment loans to individuals arising from bank check credit or other bank revolving credit plans together with credit card loans). All regressions use data aggregated at the state level. Each year, small banks are defined as banks with domestic offices only, whose total assets in that year are below \$1 billion, expressed in real 2000 dollars. The same bank can be included in some years and excluded from others. Standard errors are clustered at the state level, z-statistics are reported in parentheses below coefficients.

Table 1.12: OLS regressions of loan balances on index of state laws, 1989-2008

	UNSECURED		**SECURED**	
Variable	Credit cards, t	Other unsec. loans, t	Auto loans, t	Mortg. loans, t
Index of state laws	-10.507**	-75.898**	96.432	-182.085
	(1.98)	(2.09)	(1.16)	(-1.39)
ln(Income per capita), t	381.139***	1216.716***	602.270	7834.380***
	(4.12)	(5.08)	(1.45)	(4.14)
Other controls	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
State fixed effects	NO	NO	NO	NO
No. of obs.	867	867	1020	1020
Adjusted R^2	0.43	0.34	0.40	0.68

All regressions use data aggregated at the state level. Standard errors are clustered at the state level, t-statistics are reported in parentheses below the coefficients.

Table 1.13: Controlling for repossession professionals in IV regressions of unsecured and secured loan balances on debt collectors density, 1989-2008

Variable	ln(Average balance on credit card loans), t	ln(Average balance on other unsecured loans), t	ln(Average balance on secured auto loans), t	ln(Average balance on mortgage loans), t
ln(Debt collectors per million capita), $t-1$	0.319** (1.93)	0.897*** (2.89)	0.082 (0.96)	0.123 (1.21)
ln(Repo people per million capita), $t-1$	0.002 (0.21)	0.024 (1.48)	0.011 (2.24)	0.017* (1.80)
ln(Assets per member), t	0.128 (1.00)	0.089 (0.92)	0.148 (1.58)	0.409*** (2.81)
ln(Income per capita), t	-6.670* (-1.67)	-1.973 (-0.80)	3.397*** (2.71)	4.773*** (2.50)
ln(Property crime rate per 100,000 capita), $t-1$	-0.155 (-1.25)	-0.073 (-0.83)	-0.081 (-1.01)	-0.057 (-0.49)
Other controls	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
State fixed effects	YES	YES	YES	YES
Instrumental variables	YES	YES	YES	YES
No. of obs.	504	504	504	504
Adjusted R^2	0.87	0.92	0.88	0.95
Hansen J statistic (p-value)	2.312 (0.32)	2.505 (0.28)	0.659 (0.74)	0.708 (0.70)

All regressions use data aggregated at the state-year level. Amounts of loans per member are obtained by dividing total real amount of a particular loan type from credit union call reports by the total number of credit union members. Standard errors are clustered at the state level, z-statistics are reported in parentheses below the coefficients.

Table 1.14: OLS regressions of unsecured and secured loan balances on density of repossession professionals, 1989-2008

Variable	ln(Average balance on credit card loans), t	ln(Average balance on other unsecured loans), t	ln(Average balance on secured auto loans), t	ln(Average balance on mortgage loans), t
ln(Repo people per million capita), $t-1$	-0.001 (-0.02)	0.023 (1.09)	0.012*** (3.04)	0.020** (2.25)
ln(Assets per member), t	0.178** (2.21)	0.098 (1.01)	0.178*** (3.14)	0.438*** (3.31)
ln(Income per capita), t	-6.270*** (4.04)	-1.873 (-0.77)	3.724*** (4.38)	5.152*** (3.49)
ln(Property crime rate per 100,000 capita), $t-1$	-0.044 (-0.83)	-0.047 (-1.74)	-0.006 (-0.17)	0.059 (1.08)
Other controls	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
State fixed effects	YES	YES	YES	YES
No. of obs.	506	506	506	506
Adjusted R^2	0.89	0.93	0.92	0.95

All regressions use data aggregated at the state-year level. Amounts of loans per member are obtained by dividing total real amount of a particular loan type from credit union call reports by the total number of credit union members. Standard errors are clustered at the state level, t-statistics are reported in parentheses below the coefficients.

Table 1.15: OLS regressions of unsecured and secured loan balances on the number of complaints against third-party debt collectors, 1998-2007

Variable	ln(Average balance on credit card loans), t	ln(Average balance on other unsecured loans), t	ln(Average balance on credit card loans), t	ln(Average balance on other unsecured loans), t
ln(Complaints per million capita), $t-1$	0.035** (2.03)	0.014 (0.36)		
ln(Complaints per debt collector), $t-1$			0.005 (0.72)	0.004 (0.30)
ln(Assets per member), t	0.187** (2.37)	0.102 (1.03)	0.185** (2.36)	0.101 (1.03)
ln(Income per capita), t	-5.521*** (-3.68)	-0.786 (-0.33)	-5.774*** (-3.87)	-0.887 (-0.38)
ln(Property crime rate per 100,000 capita), $t-1$	-0.062 (-1.17)	-0.050 (-0.74)	-0.048 (-0.89)	-0.043 (-0.70)
Other controls	YES	YES	YES	YES
Year fixed effects	YES	YES	YES	YES
State fixed effects	YES	YES	YES	YES
No. of obs.	509	509	509	509
Adjusted R^2	0.89	0.92	0.90	0.93

All regressions use data aggregated at the state-year level. Amounts of loans per member are obtained by dividing total real amount of a particular loan type from credit union call reports by the total number of credit union members. Standard errors are clustered at the state level, t-statistics are reported in parentheses below the coefficients.

Chapter 2

The Economics of Debt Collection: Ex-post Enforcement of Consumer Credit Contracts

Abstract

This paper is the first attempt in the literature to study the economics of debt collection in consumer credit markets. The existence of third-party debt collection agencies cannot be explained by the benefits of specialization and economies of scale alone. Rather, the debt collection industry can serve as a coordination mechanism between creditors who can effectively impose harsh debt collection practices on consumers. If a debt collection agency collects on behalf of several creditors, the practices it uses will be associated will all creditors that hired it. Hence, consumers will be unable to single out individual creditors and punish them for using harsh practices. As a result, the third-party agency may use harsher debt collection practices than individual creditors collecting on their own. As long as the costs of hiring third-party debt collectors are below the benefits from using harsh debt collection practices, the debt collection industry will create economic value for creditors. I show that effective debt collection expands the supply of credit because banks are willing to lend to relatively riskier borrowers who would otherwise be unable to obtain credit. As a result, the pool of borrowers becomes riskier and interest rates are higher when debt collection is effective.

2.1 Introduction

In the third quarter of 2010, the amount of consumer non-mortgage debt outstanding in the U.S. was \$2425.3 billion, compared with \$7350.9 billion in total non-financial corporate debt.[1] Credit card debt alone stood at \$806.1 billion (a 16% drop from the high of \$957.5 billion in 2008).[2] Yet, despite their large size, retail credit markets received relatively little attention in the academic literature.

One interesting feature of retail credit markets is the enforcement mechanisms that creditors use in those markets. In many respects, those mechanisms are different from what corporate lenders use. Just as corporate creditors, consumer creditors can collect the debts or repossess the collateral on their own, using their own facilities and personnel, possibly with legal assistance from outside law firms (such activity is termed "in-house collection" in the consumer finance industry). Yet, unsecured consumer creditors often use third-party debt collectors while secured consumer creditors use repossession agencies. This paper is the first to address the question of why there is demand for third-party services even when the creditor has the option to collect the debt in-house.

Debt collectors play an active role in retail credit markets. They contact millions of American consumers every year. According to the Federal Trade Commission (FTC hereafter), which tracks consumer complaints, third-party debt collectors generate more complaints than any other industry. In 2010, the FTC received 140,036 complaints about third-party debt collectors, which represents 27% of all complaints received directly from consumers in 2010.[3] Thus, debt collectors are a very visible

[1]Source: Flow of Funds Accounts of the United States, Flows and Outstandings, Third Quarter 2010, Table D.3, available at http://www.federalreserve.gov/releases/z1/20101209/z1.pdf.

[2]Source: Federal Reserve Statistical Release, G.19, Consumer Credit, December 2010, available at http://www.federalreserve.gov/releases/g19/20110207/g19.pdf

[3]Source: Annual Report 2011: Fair Debt Collection Practices Act. Federal Trade Commission, Washington, DC, March 2011.

presence in the lives of American households.[4] In addition, the amount of civil litigation against debt collectors is significant. In the first five months of 2010, there were 4,808 lawsuits filed by consumers against debt collection agencies,[5] which compares with 185,900 original civil cases filed in the U.S. District Courts in 2009.[6]

Certain features of consumer credit markets and the U.S. legal system make it costly to outsource debt collection to third-party agencies. First, when an account is transferred to a third party, a significant loss of information occurs. As the FTC notes, "When accounts are transferred to debt collectors, the accompanying information often is so deficient that the collectors seek payment from the wrong consumer or demand the wrong amount from the correct consumer."[7] A representative from the National Consumer Law Center stated in testimony to the FTC that "debt collectors often lack significant information about the debts they are attempting to collect, including the date the debt was incurred and a breakdown of the fees and charges added to the original debt."[8] Moreover, debt collectors rarely have any information about the consumer's assets, unlike the original creditor who may have other relationships with the same consumer or may infer the consumer's financial health from their credit application. Hence, original creditors are better equipped with information that may be imperative in collecting consumer debt. Second, debt collection laws in the United

[4]According to InsideARM.com, there are several movie projects under way that feature debt collectors.

[5]Source: WebRecon LLC, published by InsideARM.com. Of the 4,808 lawsuits, 4,099 were filed under the FDCPA, 419 – under the Fair Credit Reporting Act, 16 – under the Telephone Consumer Protection Act, 172 – under the Truth in Lending Act, with the remaining suits filed under various other federal acts and state consumer statutes.

[6]Source: Judicial Business of the United States Courts, 2009. The total number of civil filings in 2009 was 276,397, which also includes removals from state courts, remands from courts of appeals, reopens, and transfers.

[7]"Collecting Consumer Debts: The Challenges of Change," A Report by the Federal Trade Commission, February 2009, p. 21

[8]National Consumer Law Center & the National Association of Consumer Advocates (Saunders, Lauren), comments submitted in connection with the workshop "Collecting Consumer Debts: The Challenges of Change."

States usually exempt original creditors collecting their own debts. In particular, the federal law that regulates debt collection, the Fair Debt Collection Practices Act (FDCPA hereafter), explicitly excludes original creditors from its definition of debt collectors.[9] In addition, forty-four states have state laws that regulate debt collection practices but only twenty-four of them apply to original creditors as well as to third-party debt collectors.[10] Hence, at least in principle, original creditors may have more options in pursuing defaulting consumers and face less scrutiny from regulators.

Despite the above-mentioned costs, the debt collection industry is large. There are nearly 6,500 debt collection agencies in the U.S.[11] As of May 2009, they employed 107,340 debt collectors.[12] Its estimated revenues in 2007 totalled $40.4 billion, which represented nearly 21% of private sector bad debt for that year.[13] This compares with a total of $44 billion in payday loans extended in 2007 and around $75 billion in student loans for 2008-2009.[14] Evidently, there must be advantages to hiring third-party agencies to collect consumer debts, despite the associated costs.

Benefits of specialization and economies of scale alone cannot explain the existence of the debt collection industry. First, debt collection is a human intensive process that requires little education and limited training, which makes potential benefits of specialization small. Second, the debt collection industry is large and yet very

[9]The Fair Debt Collection Practices Act, § 803(6).

[10]In most instances, state laws provide similar or weaker protection for consumer than the federal law, which is why most litigation brought by consumers against debt collectors occurs in federal courts.

[11]Source: http://www.acainternational.org/publications-collections-information-5431.aspx

[12]According to Occupational Employment Statistics, the total number of bill and account collectors stood at 403,111 in May 2009, but this number includes collectors employed by creditors directly (in their in-house collection departments). The figure reported above only includes debt collectors working in the business support services industries, i.e., third-party debt collectors. Source: http://www.bls.gov/oes/current/oes433011.htm

[13]Source: PriceWaterhouseCoopers Survey of the Debt Collection Industry, commissioned by ACA International, available at http://www.acainternational.org. More recent estimates are not available.

[14]Source: Economic impact of the payday lending industry. IHS Global Insight (USA) Inc., 2009, and the Department of Education.

decentralized. I will argue that the debt collection industry serves as a reputation protection mechanism that shields creditors from damage done by harsh collection practices. In performing this function, the debt collection industry increases the aggregate recovery rates for the creditors.

If a debt collection agency collects debts owed to several creditors, consumers cannot directly associate the practices that the agency uses with any individual creditor. Rather, they will infer that all creditors are using the same practices. Hence, if in equilibrium all creditors assign debt to third-party debt collectors, then all creditors can in effect use harsh practices. The debt collection industry is therefore a coordination mechanism that helps creditors to avoid the Prisoner's Dilemma. This argument will be made precise in a model below. The model developed in this paper generates empirical predictions that match the behavior of creditors and debt collectors and produces implications that can be addressed by future research. In particular, I show that outsourcing debt collection to a third-party agency can be beneficial for creditors even when there are no benefits from specialization or economies of scale. In addition, the model implies that each bank will employ several debt collection agencies and debt collection agencies will collect on behalf of several creditors, which is consistent with industry practice. Effective debt collection alleviates credit rationing and I show that banks will be willing to expand lending to risky borrowers when debt collection is effective. As a result, effective debt collection expands the supply of credit. I show that the average observed equilibrium interest rate is higher and the pool of borrowers is riskier when debt collection is effective.

The rest of this paper is organized as follows. Section 2.2 relates this paper to the existing literature. Section 2.3 provides evidence that specialization or economies of scale cannot suffice to explain outsourcing of debt collection to third-party agencies. In that section I develop a theory of the debt collection industry as a coordination device among creditors. In Section 2.5, I provide policy and empirical implications that follow from the model developed in this paper. Section 2.6 concludes. All proofs are confined to the appendix.

2.2 Relation to existing literature

There has been little work on creditor protection mechanisms in retail credit markets. Krumbein (1924) provides the first historical account of third-party debt collection. Hunt (2007) gives an overview of the debt collection industry and provides details about its institutional structure and regulatory environment. Hynes (2008) examines the process of debt collection in state courts and finds that consumers who are sued by creditors or debt collectors are drawn from areas with lower socio-economic characteristics. Moreover, he finds that these consumers are not likely to file for bankruptcy. Hynes, Dawsey, and Ausubel (2009) show that states with anti-harassment statutes that apply to creditors collecting their own debts have lower bankruptcy filing rates, but borrowers living in these states are more likely to default without filing for bankruptcy. This paper is most closely related to Fedaseyeu (2011) who studies the role of third-party debt collectors in expanding the supply of consumer credit. He shows that effective debt collection increases the supply of unsecured credit and is associated with higher interest rates.

This paper is also related to the theoretical literature on delegation. Fershtman, Judd, and Kalai (1991) show that principals can achieve commitments in non-cooperative games by using agents to play the game on their behalf. It happens when contracts between principals and agents are common knowledge, so that in effect principals can contract on Pareto optimal outcomes. This paper uses a much simpler setup by letting principals to simultaneously hire the same agent and thus avoiding the coordination problem altogether.[15]

2.3 Why creditors outsource debt collection

In practice, many creditors hire third-party debt collectors to collect consumer debts. In this section, I will present evidence that the benefits from outsourcing debt collection cannot be explained by gains from specialization or economies of scale alone. Rather, hiring a third-party debt collection agency serves as a coordination device for creditors, who can use it to impose harsh collection practices on defaulting consumers. I will develop a simple fully rational model to formalize this argument and will show that its implications are consistent with observable behavior of creditors and debt collectors.

Creditors' willingness to outsource debt collection to third-party agencies implies that such outsourcing generates economic value for them. However, as FTC discovered over a series of public hearings, account transfers from original creditors to third-party debt collectors often result in significant loss of information.[16] In particular, the FTC states that "debt collectors often have inadequate information when they contact consumers,"[17] A leading association of debt buyers, DBA International, acknowledged that it is common for a debt buyer to receive only a computerized summary of the creditors business records when it purchases a portfolio.[18] It is industry

[15]For a general treatment of coordination mechanisms in principal-agent games, see Myerson (1982). Goltsman, Hörner, Pavlov, and Squintani (2009) compare three common dispute resolution processes – negotiation, mediation, and arbitration – in the framework of strategic information transmission.

[16]See "Collecting Consumer Debts: The Challenges of Change," A Report by the Federal Trade Commission, February 2009 and "Repairing A Broken System: Protecting Consumers in Debt Collection Litigation and Arbitration," A Report by the Federal Trade Commission, July 2010.

[17]"Collecting Consumer Debts: The Challenges of Change," A Report by the Federal Trade Commission, February 2009, p. 20

[18]DBA Comment (June 2, 2007) before the FTC in connection with the workshop "Collecting Consumer Debts: The Challenges of Change."

practice that debt buyers can the receive relevant information from the creditor upon request, but only if such information is needed to address consumer disputes or to support a lawsuit. Even though the ability to share information exists, the technology necessary to enable a creditor to transfer all documentation at the time of sale of the portfolio may be prohibitively expensive for many creditors.[19] Often times debt collectors find it difficult to obtain even Social Security numbers of the debtors.[20]

How can third-party agencies be more effective in collecting consumer debts than original creditor when the former often lack basic information about the identity of the borrower and account history? It may be the case that economies of scale can compensate for the informational disadvantage that third-party agencies have. If this were true, the debt collection industry would be very concentrated. However, this is not the case. There are about 6,500 debt collection agencies in the United States,[21] most of which are rather small: 2,500 out of approximately 3,500 members of ACA International, the largest industry group that represents debt collectors, employ less than twenty people.[22] According to the 2007 Economic Census, four largest firms in the debt collection industry generated 14.9% of total industry revenues, with eight and twenty largest firms bringing in 23.1% and 34.3% of industry revenues, respectively. In contrast, four largest credit reporting agencies, for example, generated 78.2% of industry revenues in 2007. By this measure, the debt collection industry is closer to repossession agencies, which have to be local because the nature of their business is tied to the location of the collateral they repossess.[23] In terms of geographic disper-

[19]Bev Evancic, Transcript I at 296 of the public hearing in connection with the workshop "Collecting Consumer Debts: The Challenges of Change."

[20]A collection industry consultant, Bev Evancic, noted that, "although debt collectors find SSNs [Social Security numbers] particularly helpful in identifying the correct individual debtor, creditors remain cautious about releasing this sensitive information."

[21]Source: http://www.acainternational.org/publications-collections-information-5431.aspx

[22]Source: ACA Comment (June 6, 2007) in connection with the workshop "Collecting Consumer Debts: The Challenges of Change."

[23]According to the Census Bureau, four largest repossession agencies generated 10.0% of industry

sion, the industry appears to be even less concentrated.[24] In any given year between 1988 and 2007, there has been no state without a debt collection establishment. The highest concentration of debt collection establishments was in California in 1989, with 11% of all U.S. debt collection establishments (it was 10.5% in 2007). In 2007, thirty-two states had the number of debt collectors of at least 1% of the total number of debt collectors in the U.S.[25]

Without a doubt, technological advances such as auto-dialing have transformed the debt collection industry. However, it is unclear why original creditors could not use the same technologies to collect debts on their own. Moreover, it is common, especially for large creditors, to allocate debt collection to several third-party agencies. All of this suggests that economies of scale do not suffice to explain the existence of a third-party debt collection process.

What is it then that distinguishes third-party agencies from the original creditors? It is the harshness of debt collection practices they impose upon consumers. Imagine a world in which creditors have to collect all consumer debts on their own. Since credit provision is a repeated game, consumers will allocate their demand in subsequent period taking into account, among other factors, the way that they are treated by the creditor in the event of default. Consumers will shy away from banks that use harsh practices and will try to obtain credit from more lenient bank because consumers can establish a direct link between the bank and the debt collection practices it uses. As such, this setting becomes a Prisoner's Dilemma, in which every individual bank is better off using the most lenient practices available, regardless of what other banks do. As a result, the banking industry as a whole will use lenient practices

revenues in 2007, with eight and twenty largest firms bringing in 15.9% and 27.6% of industry revenues, respectively.

[24] Firm-level measures may overstate the degree of industry concentration because one firm can have establishments in several states.

[25] Source: Census Bureau, County Business Patterns Surveys, various years.

and the aggregate recovery rates, conditional on default, will be low. If banks could credibly coordinate by using harsh practices, all of them would be better off as long as consumers could not single out individual institutions.

Now imagine a world where banks can choose whether to delegate debt collection to a third party. It is important that this third party retains independence from the original creditor and can decide on its own how harsh it can be in collecting debts. If each bank hires a separate collection agency, consumers can still establish the link between debt collection practices and the individual creditor. In that case, this setting is no different from when the banks collect on their own. However, if one debt collection agency is used by several creditors, then whatever practices this agency is using will be associated with all those creditors and consumers will not be able to punish any individual one of them by switching to their competitor (because the competitor is effectively using the same practices).

What happens if a bank decides to avoid the agency and collect on its own while the other banks are still hiring the agency? To keep things simple, assume that there are two banks and one debt collection agency (I will relax this assumption later). Since the agency is independent of original creditors, it can switch to being lenient as soon as it sees that one of them deviates by collecting on their own. Because the agency is concerned with its own survival and this survival depends on the survival of the creditors who transfer their accounts to the agency, the agency will start using lenient practices if one of the banks deviates. As a result, this setting is different from the Prisoner's Dilemma because once a bank decides to use lenient practices it can no longer assume that the other bank will use harsh practices: the debt collection agency provides the feedback necessary to prevent this. Hence, if the gains from using harsh practices exceed the fees that need to be paid to the agency, both banks can be better off by outsourcing debt collection. In effect, they can now use the debt collection agency to coordinate on the more efficient equilibrium. This is the idea that is made precise in the model below.

2.3.1 Model setup: banks collecting debts on their own

Consider the case of credit provision in an infinitely repeated game by two competing banks, A and B. Each period, consumers demand a fixed amount of credit, Q, which they allocate between the two banks. Assume that ex ante consumers are indistinguishable in terms of their preferences and risk profiles. Each period, a fixed fraction of consumers, γ, defaults on their obligations, with $\gamma < 1$. After default, the bank decides what debt collection practices to use. Those practices can either be harsh or lenient. Let h denote the recovery rate from harsh practices and l denote the recovery rate from lenient practices. Assume that harsh practices are more effective, $h > l$. The common discount rate is $\beta, \beta < 1$.

The nature of the game is as follows. Initially, consumers allocate their demand across the two banks. Since consumers are homogenous, at the margin each consumer is indifferent between bank A and bank B. Assume for simplicity that if consumers are indifferent between the bank they are equally likely to allocate their demand to either of them. Some consumers default and the banks collect the debts owed to them by using either harsh or lenient practices. In the next period, consumers allocate demand by taking into account the harshness of the debt collection practices used before. If both banks used harsh practices, consumers are still indifferent between them and are equally likely to allocate demand to either of them. If both banks used lenient practices, consumers are again indifferent between them and are equally likely to allocate demand to either of them. If one bank uses harsh practices while the other uses lenient practices, however, the entire consumer demand in the next period is allocated to the bank that used lenient practices. Hence, if one of the banks uses harsh practices while the other bank uses lenient practices, the former disappears from the market: all future payoffs for that bank are zero. The game is then infinitely repeated. Bank maximize the net present value of all future cash flows.

Each bank's single-period profit (in period t) is given by

$$\pi_i^t(\lambda_i^t; \lambda_{-i}^t, r, \gamma,) = (r(1-\gamma) - \gamma)Q_i^t(\lambda_i^{t-1}; \lambda_{-i}^{t-1}) + \lambda_i^t \gamma Q_i^t(\lambda_i^{t-1}; \lambda_{-i}^{t-1}), \tag{2.1}$$

where:

- λ_i^t is the harshness of collection methods chosen by bank i at time t, $i \in \{A, B\}$, $\lambda_i^t \in \{l, h\}$;
- Q_i^t is the demand for credit that consumers allocate to bank i at time t, $i \in \{A, B\}$, and it is determined by the debt collection practices used in the previous period as follows:
  $Q_i^t(\lambda_i^{t-1}; \lambda_{-i}^{t-1}) = \frac{1}{2}Q$ if $\lambda_i^{t-1} = \lambda_{-i}^{t-1}$,
  $Q_i^t(\lambda_i^{t-1}; \lambda_{-i}^{t-1}) = Q$ if $\lambda_i^{t-1} = l$ and $\lambda_{-i}^{t-1} = h$,
  $Q_i^t(\lambda_i^{t-1}; \lambda_{-i}^{t-1}) = 0$ if $\lambda_i^{t-1} = h$ and $\lambda_{-i}^{t-1} = l$;
- r is the interest rate (identical across banks);
- γ is the proportion of consumers who default on their debts.

The first term in (2.1) represents profits from consumers who do not default (the proportion of such consumers is given by $1 - \gamma$). First, the bank lends Q_i^t, it then receives back the principal and interest from $1 - \gamma$ consumers who do not default. The second term in (2.1) represents recoveries from consumers who default: in this case, the bank has to use some method of debt collection, and the effectiveness of that method is denoted by λ. Notice that the amount of demand allocated to each bank depends on the collection practices used in the previous period: if at some point a bank chooses harsh practices while its competitor uses lenient practices, the former simply exits the market.

Proposition 2.1. *Let $l > h(1-\beta)$. Also assume that $r > \frac{\gamma}{1-\gamma}$ (that is, the default rate is not too high compared to the interest rate charged; otherwise, no credit will be provided). Then, there is no equilibrium in which both banks co-ordinate by using harsh collection practices. There exists a Nash equilibrium in which both banks always use lenient debt collection practices.*

Proposition 2.1 shows that when the discount rate is not very high, cooperative equilibrium in this infinitely repeated game is unsustainable. Only the non-cooperative equilibrium survives in this game, in which both banks always use lenient debt collection practices. As I show below, however, even under the above restrictions, outsourcing debt collection to a third party can solve this coordination problem and the cooperative outcome can be sustained in equilibrium.

2.3.2 Collecting debt via an intermediary

Assume that banks can delegate debt collection to a third-party agency, which in return receives a fee proportional to the amount collected for the bank that hires it (both banks can hire the agency). That proportion is f. Banks have a choice of whether to collect on their own or to hire the agency. If they hire the agency, it is now the agency who determines the harshness of debt collection practices to be used. If a bank chooses not to delegate debt collection to the agency, that bank (possibly both banks) has to determine the harshness of debt collection practices it will use. Each bank maximizes the net present value of all future cash flows. The agency maximizes the present value of all future fees it collects.

If only one bank hires the debt collection agency, consumers will be able to associate that bank with the practices of the agency. If both banks hire the agency, consumers cannot punish any individual bank since in effect both banks use the same collection practices. If only one bank chooses the agency and the agency engages in harsh collections while the other bank collects leniently, the bank that hired the agency will lose its customers in all subsequent periods. All future profits for that bank and the collection agency will be zero. However, as I show below, under certain conditions this game has an equilibrium in which both banks delegate debt collection to the agency and the agency uses harsh collection practices, even when a cooperative equilibrium is unsustainable in the absence of the agency.

Proposition 2.2. *Assume $h(1-f) > l > h(1-\beta)$. In addition, assume that $r > \frac{\gamma}{1-\gamma}$ and $h(f-\beta) < \frac{(r(1-\gamma)-\gamma)\beta}{\gamma}$.*[26] *Then, the game above has a Nash equilibrium in which both banks decide to delegate debt collection to the agency each period and agency uses harsh collection practices.*

Notice that Proposition 2.2 assumes that $l > h(1-\beta)$ and $r > \frac{\gamma}{1-\gamma}$. Hence, by Proposition 2.1, in a game without the agency a cooperative equilibrium is unsustainable because each bank has an incentive to deviate and use lenient practices. The difference in the game with the agency is the fact that the deviation is no longer costless. The agency adopts the following equilibrium strategy: it uses harsh practices if and only of both banks delegate debt collection to the agency and uses lenient practices otherwise. Hence, if one bank decides to deviate in the hope of uses lenient practices and thus attracting borrowers, the agency also starts to use lenient practices and that way it ensures that its client doesn't lose future customers, so that the

[26]The following set of parameters, for example, will satisfy all of these conditions: $f = 0.3, \beta = 0.7, h = 0.5, l = 0.2, r = 0.15, \gamma = 0.1$.

agency can continue to collect in the future. For such this equilibrium to exist, the fees charged by the agency should not be too high, $h(1-f) > l$ and $h(f-\beta) < \frac{(r(1-\gamma)-\gamma)\beta}{\gamma}$, so that the benefits of being able to use harsh practices exceed the costs of hiring the agency.

Proposition 2.2 can explain why delegating debt collection to third-party agencies may be economically feasible even in the absence of benefits from specialization and economies of scale. In effect, the debt collection industry serves as a coordination device and erects a shield that protects reputations of individual creditors. When consumers are contacted by debt collectors, they are told who the original creditor is. Hence, in principle, consumers who rationally understand that it is, for example, the Bank of America who ultimately benefits from debt collection, can avoid that bank in the future. However, this strategy will not work if, for example, Chase and other banks are using the same agency. As long as the entire banking industry can coordinate and engage in harsh collection practices, the consumers are powerless to take action against individual banks.

2.3.3 More than one agency

In the previous section we considered a game where the banks could delegate debt collection to only one agency. In reality, however, the number of debt collection agencies is large. In this section I show that the basic intuition developed above carries to the case of multiple debt collection agencies. The theory developed here is concerned with the relationship between creditors and debt collectors, and it therefore has nothing to say about why there may be multiple agencies. There are, however, implications for how creditors will hire those agencies if they have a choice.

Consider the game as before but assume that instead of two banks and one debt collection agency, there are two banks, A and B, and two agencies, a and b. Also assume that both agencies charge the same fee for their services.

Proposition 2.3. *Assume $h(1-f) > l > h(1-\beta)$. In addition, assume that $r > \frac{\gamma}{1-\gamma}$ and $h(f-\beta) < \frac{(r(1-\gamma)-\gamma)\beta}{\gamma}$. Also, assume that there are two banks, A and B, and two debt collection agencies, a and b. Then, if banks choose to delegate debt collection, either both banks delegate debt collection to one and the same agency or both banks delegate to both agencies. There is no equilibrium in which one bank delegates to agency a while the other bank delegates to agency b. There is no equilibrium in which one bank delegates to both agencies, a and b, while the other bank delegates to only one agency, a or b.*

The intuition behind Proposition 2.3 is as follows. If each bank hires its own debt collection agency, then consumers can establish a link between those agencies and their respective banks. As a result, the agencies will use lenient practices. Hence, there is no benefit in hiring third-party agencies and banks can save the costs by collecting on their own. If both banks hire both agencies, on the other hand, then consumers can no longer associate debt collection practices with individual banks. Since both agencies receive business from both banks, they can use harsh collection practices. This intuition generalizes to the case of more than two agencies.

There may be exogenous reasons for why the debt collection industry may be decentralized, and I discuss them later. However, the empirical implication of the previous result is that in a world with several debt collection agencies creditors should allocate their business to several of them while each agency should collect on behalf of several creditors. If there are exclusive relationships between creditors and debt collectors, those creditors and debt collectors should use lenient practices. Debt collection is a sensitive topic for lenders, which is why it is difficult to obtain data on their behavior in this area. From private conversations with industry professionals, however, it seems that allocating debt collection business among several debt collection agencies is the usual industry practice. The U.S. Department of Education, for example, is currently using twenty-two debt collection agencies and has a scoring system to assign accounts for collection.[27] Some hospitals, on the other hand, form exclusive relationships and in those cases the creditors often exist that consumers have to be treated well. Debt collection agencies usually collect from several creditors,[28] which is also consistent with the model's predictions.

2.3.4 Competing on interest rates and other terms

One of the simplifying assumptions I made above was that there are only two banks who are competing only on the harshness of debt collection practices they are using. In this section, I show that my results will go through as long as the harshness of debt collection practices is not contractible. In particular, if the pool of borrowers is homogenous, then conditional on default, the bank is always better off using the harshest practices it can use. Hence, regardless of the other terms of the contract, the banks will try to find arrangements that enable them to use harsh practices, independent of the initial contract terms.

The credit provision game each period can be split into two rounds: in the first round, consumers allocate credit based on contract characteristics other than debt collection practices, in the second round, they adjust the allocation of the first round based on the debt collection practices used in the previous period. Assume that several banks compete both on interest rates and on other terms of the contract. In the first

[27] Source: Fiscal Year 2010 Report to the Congress, "U.S. Government Receivables and Debt Collection Activities of Federal Agencies," Department of the Treasury, Financial Management Service, March 2011.

[28] See "Top Collection Markets Survey" by ACA International, various years.

round of each period, borrowers allocate their credit demand among competing banks based contract terms. Assume that bank i charges interest rate r_i and that it receives a share q_i of the total demand. In the second round, borrowers reallocate demand based on the debt collection practices used by banks in the previous period. If all banks used harsh practices, consumers allocate their demand based on contractible terms only (debt collection is irrelevant). If all banks used lenient practices, consumers again allocate their demand based on contractible terms only (debt collection is irrelevant). If one bank uses harsh practices while the other banks use lenient practices, however, the entire consumer demand in the next period is allocated to the bank that used lenient practices.

Proposition 2.4. *Assume $h(1-f) > l > h(1-\beta)$. In addition, assume that $r > \frac{\gamma}{1-\gamma}$ and $h(f-\beta) < \frac{(r(1-\gamma)-\gamma)\beta}{\gamma}$.*[29] *Then, the game above has a Nash equilibrium in which both banks decide to delegate debt collection to the agency each period and agency uses harsh collection practices.*

2.4 Debt collection and credit rationing

Effective debt collection implies that credit contracts can be effectively enforced ex-post, after the borrower defaults. In this section I examine whether this ex-post enforcement affects ex-ante incentives to screen borrowers and ration credit. Assume that there are two types of consumers, safe and risky. Let ω denote the share of safe consumers in the population, with $1-\omega$ being the share of risky consumers. Let γ_r denote the default rate by risky consumers and γ_s denote the default rate by safe consumers, $\gamma_s < \gamma_r$. In addition, assume that banks have access to a a screening

[29] The following set of parameters, for example, will satisfy all of these conditions: $f = 0.3, \beta = 0.7, h = 0.5, l = 0.2, r = 0.15, \gamma = 0.1$.

technology, which enables them to tell apart risky consumers from safe consumers with certainty. I normalize the cost of the screening technology to 0.[30]

Once the banks know the type of borrower, they can charge differential interest rates to different types of borrowers and may also choose to ration credit. Assume that safe borrowers are willing to accept interest rate r_s while risky borrowers are willing to accept r_r, $r_s < r_r$.[31] Also assume that credit is not rationed to safe borrowers (or, equivalently, $r_s > \frac{\gamma_s}{1-\gamma_s}$). As the next proposition shows, banks will lend to risky borrowers if and only if the effectiveness of debt collection is high enough while the fee charged by the debt collection agency is low enough. Also, the observed average interest rate charged in such equilibrium will be higher than in the equilibrium when risky borrowers are rationed out of the credit market.

Proposition 2.5. *Assume that* $h(1-f) > l > h(1-\beta)$, $r_s > \frac{\gamma_s}{1-\gamma_s}$, *and* $h(f-\beta) < \frac{(r_s(1-\gamma_s)-\gamma_s)\beta}{\gamma_s}$ *and* $h(f-\beta) < \frac{(r_r(1-\gamma_r)-\gamma_r)\beta}{\gamma_r}$. *Then, if banks invest in the screening technology, the following is true:*

(*i*) *Banks provide credit to both risky and safe borrowers if and only if* $h(1-f)\gamma_r > r_r(1-\gamma_r) - \gamma_r$. *Otherwise, only safe borrowers are able to obtain credit.*

(*ii*) *The average interest rate charged in equilibrium is higher when* $h(1-f)\gamma_r > r_r(1-\gamma_r) - \gamma_r$ *than when* $h(1-f)\gamma_r < r_r(1-\gamma_r) - \gamma_r$.

Proposition 2.5 is intuitive. Credit will not be rationed as long as the higher risk associated with risky borrowers is compensated by the effectiveness of ex-post debt collection. In particular, $h(1-f)\gamma_r$ is high when h is high or f is low: either harsh debt collection practices are very effective or the fee charged by third-party

[30]All results will go unchanged if I assume that the technology entails a fixed per period maintenance cost or a fixed lump-sum outlay at the beginning of the game.

[31]This assumption is in line with Stiglitz and Weiss (1981), where risky borrowers may be willing to accept a higher interest rate because their repayment probability is lower.

debt collectors is low, or both. Moreover, when credit is not rationed, banks are able to attract riskier borrowers and also charge those borrowers a higher interest rate on their loans. Hence, the average equilibrium interest rate is higher. It happens because when credit is rationed, all borrowers are safe and are charged r_s but when credit is not rationed, the average equilibrium interest rate is given by $r_s\omega + r_r(1-\omega)$. It is easy to see that $r_s\omega + r_r(1-\omega) > r_s$ because $r_r > r_s$ and $\omega < 1$.

2.5 Empirical and policy implications

The model developed in this paper has several empirical implications, which are listed below.

(i) Debt collection industry and creditor behavior: The model in this paper provides a rationale for the existence of third-party debt collectors. Even when there are no benefits from specialization and economies of scale, creditors may use third-party debt collectors as a coordination device. If a third-party agency collects on behalf of several creditors, all of those creditors will be associated with the practices that the agency uses. Hence, consumers will be unable to single out individual creditors and punish them for using harsh practices. As a result, the third-party agency may use harsher debt collection practices than individual creditors collecting on their own. One implication that follows from this set-up is that in a world of several creditors and several debt collection agencies each creditor should allocate debt collection to several agencies and each agency should collect on behalf of several creditors. This is generally consistent with industry practice. It may be interesting, however, to test this implication directly by looking at the debt collection behavior of various banks.

(ii) Effective debt collection alleviates credit rationing and expands credit supply: When creditors can assess the riskiness of borrowers, they may ration credit. I show

that banks are willing to extend credit to risky borrowers only when debt collection practices are effective and expected recoveries after default compensate for the increase in the default probability. This result is consistent with evidence in Fedaseyeu (2011), who shows that effective third-party debt collection expands the supply of consumer credit.

(iii) Effective debt collection leads to higher interest rates and a riskier pool of borrowers: Effective debt collection enables banks to expand the pool of borrowers by providing credit to both safe and relatively risky applicants. As a result, the pool of borrowers becomes riskier. Also, since risky borrowers are charged a higher interest rate, the average observed equilibrium interest rate goes up. Fedaseyeu (2011) provides some evidence consistent with this implication but his evidence is not a direct test and further research is needed to confirm or falsify this result.

(iv) FTC enforcement actions should take into account the number of complaints per debt collector: Currently, the FTC tends to bring enforcement action against debt collection agencies that generate a large absolute number of consumer complaints. Keeping a debt collection agency small, therefore, enables it to engage in regulatory avoidance while the suffering it inflicts on consumers may be larger, on a per collector basis, than that of a bigger agency. It may very well be the case that several small debt collection agencies, collectively, inflict more damage on consumers than a larger agency which employs more debt collectors than the small agencies, combined. Bringing the enforcement mechanism in line with potential damage caused by individual collectors will alleviate this problem.

2.6 Conclusion

This paper is the first attempt in the literature to study the economics of debt collection in consumer credit markets. The existence of third-party debt collection agencies cannot be explained by the benefits of specialization and economies of scale alone. Rather, the debt collection industry can serve as a coordination mechanism between creditors who can effectively impose harsh debt collection practices on consumers. If a debt collection agency collects on behalf of several creditors, the practices it uses will be associated will all creditors that hired it. Hence, consumers will be unable to single out individual creditors and punish them for using harsh practices. As a result, the third-party agency may use harsher debt collection practices than individual creditors collecting on their own. As long as the costs of hiring third-party debt collectors are below the benefits from coordinated action by creditors, the debt collection industry will create economic value for creditors.

I show that effective debt collection expands the supply of credit because banks are willing to lend to relatively riskier borrowers who would otherwise be unable to obtain credit. As a result, the pool of borrowers becomes riskier and interest rates are higher when debt collection is effective. There is empirical evidence consistent with these implications (see Fedaseyeu (2011)). However, further research is needed to develop direct tests that will confirm or falsify these results.

2.7 Appendix: Proofs of Propositions

Proposition 2.1. *Let $l > h(1-\beta)$. Also assume that $r > \frac{\gamma}{1-\gamma}$ (that is, the default rate is not too high compared to the interest rate charged: this assumption insures that credit is not rationed). Then, there is no equilibrium in which both banks co-ordinate by using harsh collection practices. There exists a Nash equilibrium in which both banks always use lenient debt collection practices.*

Proof. I prove the first part by contradiction. Assume that there exists an equilibrium in which both banks co-ordinate and use harsh debt collection practices. The payoff to bank i in such equilibrium is given by

$$\pi_i = \sum_{t=0}^{\infty} \frac{1}{2}\beta^t (r(1-\gamma)-\gamma)Q + \sum_{t=0}^{\infty} \frac{1}{2}\beta^t h\gamma Q. \tag{2.7.1}$$

Assume that bank i decides to deviate by using lenient debt collection practices in period τ. Then, in period τ the recovery rate for this bank will equal l. However, it will be allocated the entire credit demand in all subsequent periods and will be able to use harsh collection practices in all of those periods (because there will be no competition from the other bank). Hence, the payoff to bank i from deviating in period τ is given by

$$\begin{aligned}\pi_i = &\sum_{t=0}^{\tau} \frac{1}{2}\beta^t (r(1-\gamma)-\gamma)Q + \sum_{t=0}^{\tau-1} \frac{1}{2}\beta^t h\gamma Q + \frac{1}{2}\beta^\tau l\gamma Q + \\ &\sum_{t=\tau+1}^{\infty} \beta^t (r(1-\gamma)-\gamma)Q + \sum_{t=\tau+1}^{\infty} \beta^t h\gamma Q.\end{aligned} \tag{2.7.2}$$

Subtract (2.7.1) from (2.7.2) to obtain

$$\frac{1}{2}\beta^\tau (l-h)\gamma Q + \sum_{t=\tau+1}^{\infty} \frac{1}{2}\beta^t (r(1-\gamma)-\gamma)Q + \sum_{t=\tau+1}^{\infty} \frac{1}{2}\beta^t h\gamma Q. \tag{2.7.3}$$

Rearranging (2.7.3) yields

$$\frac{1}{2}\beta^\tau Q\Big\{l\gamma - h\gamma\big(1-\sum_{t=1}^{\infty}\beta^t\big) + \sum_{t=1}^{\infty}\beta^t (r(1-\gamma)-\gamma)\Big\}. \tag{2.7.4}$$

Notice that since $\sum_{t=1}^{\infty}\beta^t(r(1-\gamma)-\gamma) > 0$ by assumption, (2.7.4) is greater than zero whenever $l > h\big(1-\sum_{t=1}^{\infty}\beta^t\big)$. We assumed that $l > h(1-\beta) > h\big(1-\sum_{t=1}^{\infty}\beta^t\big)$. Hence, each bank has an incentive to deviate and the co-ordination equilibrium does not exist. Contradiction.

The proof of existence of a non-cooperative equilibrium is similar. First, consider the payoff to bank i in a non-cooperative equilibrium:

$$\pi_i = \sum_{t=0}^{\infty} \frac{1}{2}\beta^t (r(1-\gamma)-\gamma)Q + \sum_{t=0}^{\infty} \frac{1}{2}\beta^t l\gamma Q. \tag{2.7.5}$$

Assume that bank i decides to deviate by using harsh debt collection practices in period τ. Then, in period τ the recovery rate for this bank will equal h. However, it will lose customers in all subsequent periods. Hence, the payoff to bank i from deviating in period τ is given by

$$\pi_i = \sum_{t=0}^{\tau} \frac{1}{2}\beta^t (r(1-\gamma)-\gamma)Q + \sum_{t=0}^{\tau-1} \frac{1}{2}\beta^t l\gamma Q + \frac{1}{2}\beta^\tau h\gamma Q. \tag{2.7.6}$$

Subtract (2.7.5) from (2.7.6) to obtain

$$\frac{1}{2}\beta^\tau (h-l)\gamma Q - \sum_{t=\tau+1}^{\infty} \frac{1}{2}\beta^t (r(1-\gamma)-\gamma)Q - \sum_{t=\tau+1}^{\infty} \frac{1}{2}\beta^t l\gamma Q. \tag{2.7.7}$$

Rearranging,

$$\frac{1}{2}\beta^\tau Q\Big\{h\gamma - l\gamma\big(1+\sum_{t=1}^{\infty}\beta^t\big) - \sum_{t=1}^{\infty}\beta^t(r(1-\gamma)-\gamma)\Big\}. \tag{2.7.8}$$

Notice that since $\sum_{t=1}^{\infty}\beta^t(r(1-\gamma)-\gamma) > 0$, (2.7.8) is less than zero whenever $h < l\big(1+\sum_{t=1}^{\infty}\beta^t\big)$, or, equivalently, $l > h(1-\beta)$. Hence, no bank has an incentive to deviate and non-cooperation is a Nash equilibrium. □

Proposition 2.2. *Assume $h(1-f) > l > h(1-\beta)$. In addition, assume that $r > \frac{\gamma}{1-\gamma}$ and $h(f-\beta) < \frac{(r(1-\gamma)-\gamma)\beta}{\gamma}$. Then, the game above has a Nash equilibrium in which both banks decide to delegate debt collection to the agency each period and agency uses harsh collection practices.*

Proof. Consider the following strategies. Each period, each bank delegates debt collection to the third-party agency. The agency uses harsh practices if both banks delegate debt collection to it and lenient practices if only one of the banks delegates debt collection to it. I show that these strategies constitute a Nash equilibrium.

First, consider the equilibrium payoff to bank i:

$$\pi_i = \sum_{t=0}^{\infty} \frac{1}{2}\beta^t (r(1-\gamma)-\gamma)Q + \sum_{t=0}^{\infty} \frac{1}{2}\beta^t h(1-f)\gamma Q. \tag{2.7.9}$$

The last term in (2.7.9) contains the $(1-f)$ because the agency remits to the bank only $(1-f)$ share of the amount it recovers from borrowers.

Consider bank i's payoff if it decides to deviate in period τ by collecting on its own and using lenient practices. Once the bank deviates, the agency collecting on behalf of the other bank immediately starts using lenient practices and so there is no reallocation of credit demand:

$$\pi_i = \sum_{t=0}^{\infty} \frac{1}{2}\beta^t (r(1-\gamma)-\gamma)Q + \sum_{t=0}^{\tau-1} \frac{1}{2}\beta^t h(1-f)\gamma Q + \sum_{t=\tau}^{\infty} \frac{1}{2}\beta^t l\gamma Q. \tag{2.7.10}$$

Subtract (2.7.9) from (2.7.10) to obtain

$$\sum_{t=\tau}^{\infty} \frac{1}{2}\beta^t \gamma Q(l - h(1-f)) < 0. \tag{2.7.11}$$

Hence, there is no incentive to deviate by using lenient practices.

Now consider bank i's payoff if it decides to deviate in period τ by collecting on its own and using harsh practices. Once the bank deviates, the agency collecting on behalf of the other bank immediately starts using lenient practices and in subsequent periods all consumer demand will be allocated to the other bank:

$$\pi_i = \sum_{t=0}^{\tau} \frac{1}{2}\beta^t (r(1-\gamma) - \gamma)Q + \sum_{t=0}^{\tau-1} \frac{1}{2}\beta^t h(1-f)\gamma Q + \frac{1}{2}\beta^\tau h\gamma Q. \tag{2.7.12}$$

Subtract (2.7.9) from (2.7.12) to obtain

$$-\sum_{t=\tau+1}^{\infty} \frac{1}{2}\beta^t (r(1-\gamma) - \gamma)Q - \sum_{t=\tau+1}^{\infty} \frac{1}{2}\beta^t h(1-f)\gamma Q + \frac{1}{2}\beta^\tau (h - h(1-f))\gamma Q. \tag{2.7.13}$$

(2.7.13) is less than zero if and only if

$$\begin{aligned}
&\frac{1}{2}\beta^\tau Q\gamma f h < \frac{1}{2}\beta^\tau Q\Big\{(r(1-\gamma) - \gamma)\sum_{t=1}^{\infty}\beta^t + \gamma h(1-f)\sum_{t=1}^{\infty}\beta^t\Big\} \Leftrightarrow \\
&\gamma f h < (r(1-\gamma) - \gamma)\frac{\beta}{1-\beta} + \gamma h(1-f)\frac{\beta}{1-\beta} \Leftrightarrow \\
&\gamma f h(1-\beta) < (r(1-\gamma) - \gamma)\beta + \gamma h(1-f)\beta \Leftrightarrow \\
&\gamma h(f-\beta) < (r(1-\gamma) - \gamma)\beta.
\end{aligned} \tag{2.7.14}$$

We assumed that $h(f-\beta) < \frac{(r(1-\gamma)-\gamma)\beta}{\gamma}$. Hence, there is no incentive to deviate by using harsh practices. Now, consider the agency's equilibrium strategy. Its equilibrium payoff is given by

$$\sum_{t=1}^{\infty} \beta^t h f \gamma Q. \tag{2.7.15}$$

Consider the agency's payoff if it deviates by using lenient practices during a single period, t_0:

$$\sum_{t=1}^{t_0-1} \beta^t l h \gamma Q + \beta^{t_0} l f \gamma Q + \sum_{t=t_0+1}^{\infty} \beta^t l h \gamma Q. \tag{2.7.16}$$

Since $l < h$, (2.7.16)<(2.7.15) and there is no incentive to deviate by always using lenient practices.

It is also easy to see that the agency has no incentive to deviate by using lenient practices whenever both banks delegate their debt collection to it and harsh practices when only one bank delegates debt collection to it. Given the banks' equilibrium strategies, the agency's payoff in this case is given by (2.7.16). Lastly, if the agency deviates by using harsh practices regardless of whether one or both banks delegate their debt collection to it, its payoff is given by (2.7.15) and it has no incentive to deviate. □

Proposition 2.3. *Assume $h(1-f) > l > h(1-\beta)$. In addition, assume that $r > \frac{\gamma}{1-\gamma}$ and $h(f-\beta) < \frac{(r(1-\gamma)-\gamma)\beta}{\gamma}$. Also assume that there are two banks, A and B, and two debt collection agencies, a and b. Then, if banks choose to delegate debt collection, either both banks delegate debt collection to one and the same agency or both banks delegate to both agencies. There is no equilibrium in which one bank delegates to agency a while the other bank delegates to agency b. There is no equilibrium in which one bank delegates to both agencies, a and b, while the other bank delegates to only one agency, a or b.*

Proof. As we showed in Proposition 2.2, there exists an equilibrium in which both banks delegate to one agency. If both banks delegate to both agencies, a and b, then each agency collects on behalf of both banks and consumers are unable to establish the link between the creditor and the collection practices that creditor uses. Hence, the payoff structure for the banks does not change and the payoff of each agency is proportional to the share of debt collection business that is delegated to that agency. Again, the equilibrium described above exists.

It is easy to see that there is no equilibrium in which one bank chooses agency a while the other bank chooses agency b. Assume without loss of generality that bank A chooses agency a and bank B chooses agency b. In that case, consumers can establish a direct link between the practices used by agency a and bank A and between practices used by agency b and bank B. The game is therefore equivalent to the game when both banks collect on their own but only receive $(1-f)$ share of the proceeds. Hence, by Proposition 2.1, both banks use lenient practices. Since they would have used lenient practices without hiring the agencies, both banks are better off collecting on their own and saving the fee paid to the agency. Hence, an equilibrium in which one bank chooses agency a while the other bank chooses agency b does not exist.

All that remains is to show that there is no equilibrium in which one bank delegates to both agencies, a and b, while the other bank delegates to only one agency, a or b. Assume without loss of generality that bank A delegates to both agencies and bank B uses agency b only. Consider agency a. Since it only receives assignment from one bank, A, it uses lenient practices to collect the debts. Agency b receives assignments from both banks and therefore uses harsh practices.

In this arrangement, consumers can establish a link between the practices used by agency a and bank A. Hence, they will switch to bank A to the extent that it assigns debt to agency a. However, bank A needs to use lenient practices to collect from those consumers. If lifetime benefits of attracting more consumers exceed the costs of using less productive collection practices, nothing prevents bank B from also using agency a. If the lifetime benefits of attracting more consumers are lower than the costs of using less productive collection practices, bank A would not have used agency a exclusively in the first place. Either way, the equilibrium in which one bank uses both agencies while the other uses only one of them does not exist. □

Proposition 2.4. *Assume $h(1-f) > l > h(1-\beta)$. In addition, assume that $r > \frac{\gamma}{1-\gamma}$ and $h(f-\beta) < \frac{(r(1-\gamma)-\gamma)\beta}{\gamma}$.*[32] *Then, the game above has a Nash equilibrium in which both banks decide to delegate debt collection to the agency each period and agency uses harsh collection practices.*

Proof. Consider the following strategies. Each period, each bank delegates debt collection to the third-party agency. The agency uses harsh practices if all banks delegate debt collection to it and lenient practices if only some of the banks delegate debt collection to it. I show that these strategies constitute a Nash equilibrium.

First, consider the equilibrium payoff to bank i:

$$\pi_i = \sum_{t=0}^{\infty} q_i \beta^t (r_i(1-\gamma)-\gamma)Q + \sum_{t=0}^{\infty} q_i \beta^t h(1-f)\gamma Q. \tag{2.7.17}$$

The last term in (2.7.17) contains the $(1-f)$ because the agency remits to the bank only $(1-f)$ share of the amount it recovers from borrowers.

Consider bank i's payoff if it decides to deviate in period τ by collecting on its own and using lenient practices. Once the bank deviates, the agency collecting on behalf of the other bank immediately starts using lenient practices and so there is no reallocation of credit demand:

$$\pi_i = \sum_{t=0}^{\infty} q_i \beta^t (r_i(1-\gamma)-\gamma)Q + \sum_{t=0}^{\tau-1} q_i \beta^t h(1-f)\gamma Q + \sum_{t=\tau}^{\infty} q_i \beta^t l \gamma Q. \tag{2.7.18}$$

Subtract (2.7.17) from (2.7.18) to obtain

$$\sum_{t=\tau}^{\infty} q_i \beta^t \gamma Q(l-h(1-f)) < 0. \tag{2.7.19}$$

Hence, there is no incentive to deviate by using lenient practices.

Now consider bank i's payoff if it decides to deviate in period τ by collecting on its own and using harsh practices. Once the bank deviates, the agency collecting on behalf of the other bank immediately starts using lenient practices and in subsequent periods all consumer demand will be allocated to the other bank:

$$\pi_i = \sum_{t=0}^{\tau} q_i \beta^t (r_i(1-\gamma)-\gamma)Q + \sum_{t=0}^{\tau-1} q_i \beta^t h(1-f)\gamma Q + q_i \beta^\tau h \gamma Q. \tag{2.7.20}$$

Subtract (2.7.17) from (2.7.20) to obtain

$$-\sum_{t=\tau+1}^{\infty} q_i \beta^t (r_i(1-\gamma)-\gamma)Q - \sum_{t=\tau+1}^{\infty} q_i \beta^t h(1-f)\gamma Q + q_i \beta^\tau (h-h(1-f))\gamma Q. \tag{2.7.21}$$

[32] The following set of parameters, for example, will satisfy all of these conditions: $f = 0.3, \beta = 0.7, h = 0.5, l = 0.2, r = 0.15, \gamma = 0.1$.

(2.7.21) is less than zero if and only if

$$
\begin{aligned}
&q_i\beta^\tau Q\gamma fh < q_i\beta^\tau Q\Big\{(r_i(1-\gamma)-\gamma)\sum_{t=1}^{\infty}\beta^t + \gamma h(1-f)\sum_{t=1}^{\infty}\beta^t\Big\} \Leftrightarrow \\
&\gamma fh < (r_i(1-\gamma)-\gamma)\frac{\beta}{1-\beta} + \gamma h(1-f)\frac{\beta}{1-\beta} \Leftrightarrow \\
&\gamma fh(1-\beta) < (r_i(1-\gamma)-\gamma)\beta + \gamma h(1-f)\beta \Leftrightarrow \\
&\gamma h(f-\beta) < (r_i(1-\gamma)-\gamma)\beta.
\end{aligned}
\tag{2.7.22}
$$

We assumed that $h(f-\beta) < \frac{(r_i(1-\gamma)-\gamma)\beta}{\gamma}$. Hence, there is no incentive to deviate by using harsh practices. Now, consider the agency's equilibrium strategy. Its equilibrium payoff is given by

$$
\sum_{t=1}^{\infty}\beta^t hf\gamma Q. \tag{2.7.23}
$$

Consider the agency's payoff if it deviates by using lenient practices during a single period, t_0:

$$
\sum_{t=1}^{t_0-1}\beta^t lh\gamma Q + \beta^{t_0} lf\gamma Q + \sum_{t=t_0+1}^{\infty}\beta^t lh\gamma Q. \tag{2.7.24}
$$

Since $l < h$, (2.7.24)<(2.7.23) and there is no incentive to deviate by always using lenient practices.

It is also easy to see that the agency has no incentive to deviate by using lenient practices whenever both banks delegate their debt collection to it and harsh practices when only one bank delegates debt collection to it. Given the banks' equilibrium strategies, the agency's payoff in this case is given by (2.7.24). Lastly, if the agency deviates by using harsh practices regardless of whether one or both banks delegate their debt collection to it, its payoff is given by (2.7.23) and it has no incentive to deviate. □

Proposition 2.5. *Assume that* $h(1-f) > l > h(1-\beta)$, $r_s > \frac{\gamma_s}{1-\gamma_s}$, *and* $h(f-\beta) < \frac{(r_s(1-\gamma_s)-\gamma_s)\beta}{\gamma_s}$ *and* $h(f-\beta) < \frac{(r_r(1-\gamma_r)-\gamma_r)\beta}{\gamma_r}$. *Then, if banks invest in the screening technology, the following is true:*

(*i*) *Banks provide credit to both risky and safe borrowers if and only if* $h(1-f)\gamma_r > \gamma_r - r_r(1-\gamma_r)$. *Otherwise, only safe borrowers are able to obtain credit.*

(*ii*) *The average interest rate charged in equilibrium is higher when* $h(1-f)\gamma_r > \gamma_r - r_r(1-\gamma_r)$ *than when* $h(1-f)\gamma_r < \gamma_r - r_r(1-\gamma_r)$.

Proof. We assumed that $h(1-f) > l > h(1-\beta)$, $r_s > \frac{\gamma_s}{1-\gamma_s}$, and $h(f-\beta) < \frac{(r_s(1-\gamma_s)-\gamma_s)\beta}{\gamma_s}$ and $h(f-\beta) < \frac{(r_r(1-\gamma_r)-\gamma_r)\beta}{\gamma_r}$. Hence, by Proposition 2.2, there exists a symmetric equilibrium in which all banks use the debt collection agency to collect both from safe and from risky borrowers. We need $h(f-\beta) < \frac{(r_s(1-\gamma_s)-\gamma_s)\beta}{\gamma_s}$ and $h(f-\beta) < \frac{(r_r(1-\gamma_r)-\gamma_r)\beta}{\gamma_r}$ to ensure that debt collection from both types of borrowers is delegated to the debt collection agency.

We will now differentiate between two types of the debt collection equilibrium: the equilibrium in which credit is rationed and the equilibrium in which credit is not rationed. If credit is rationed, the payoff to each bank is given by

$$\begin{aligned}&\sum_{t=0}^{\infty}\beta^t\Big[r_s\omega(1-\gamma_s)q_iQ-\omega\gamma_sq_iQ\Big]+\sum_{t=0}^{\infty}\beta^th(1-f)\omega\gamma_sq_iQ+\\&\sum_{t=0}^{\infty}\beta^t\Big[r_r(1-\omega)(1-\gamma_r)q_iQ-(1-\omega)\gamma_rq_iQ\Big]+\sum_{t=0}^{\infty}\beta^th(1-f)(1-\omega)\gamma_rq_iQ.\end{aligned} \tag{2.7.25}$$

In the equilibrium in which banks ration credit to risky borrowers, each bank's payoff will be given by

$$\sum_{t=0}^{\infty}\beta^t\Big[r_s\omega(1-\gamma_s)q_iQ-\omega\gamma_sq_iQ\Big]+\sum_{t=0}^{\infty}\beta^th(1-f)\omega\gamma_sq_iQ. \tag{2.7.26}$$

Subtract (2.7.26) from (2.7.25) to obtain

$$\sum_{t=0}^{\infty}\beta^t\Big[r_r(1-\omega)(1-\gamma_r)q_iQ-(1-\omega)\gamma_rq_iQ\Big]+\sum_{t=0}^{\infty}\beta^th(1-f)(1-\omega)\gamma_rq_iQ. \tag{2.7.27}$$

There is no incentive to deviate from the equilibrium in which credit is not rationed if and only if (2.7.27) is greater than zero. After a little bit of algebra this is equivalent to

$$h(1-f)\gamma_r>\gamma_r-r_r(1-\gamma_r). \tag{2.7.28}$$

This proves part (i).

To prove part (ii), notice that in the equilibrium when credit is rationed, all borrowers are charged interest rate r_s. In the equilibrium with no credit rationing, the average interest rate is given by $r_s\omega+r_r(1-\omega)>r_s$, because $r_r>r_s$ and $\omega<1$. □

Chapter 3

A Theory of Corporate Boards and Forced CEO Turnover

This essay has been written jointly with Thomas Chemmanur.

Abstract

We develop a theory of corporate boards and their role in forcing CEO turnover. We consider a firm with an incumbent CEO of uncertain management ability and a board consisting of a number of directors whose role is to evaluate the CEO and fire her if a better replacement can be found. Each board member receives an independent private signal about the CEO's ability, after which board members vote on firing the CEO (or not). If the CEO is fired, the board hires a new CEO from the pool of candidates available. The true ability of the firm's CEO is revealed in the long run; the firm's long-run share price is determined by this ability. Each board member owns some equity in the firm, and thus prefers to fire a CEO of poor ability. However, if a board member votes to fire the incumbent CEO but the number of other board members also voting to fire her is not enough to successfully oust her, the CEO can impose significant costs of dissent on him. In this setting, we show that the board faces a coordination problem, leading it to retain an incompetent CEO even when a majority of board members receive private signals indicating that she is of poor quality. We solve for the optimal board size, and show that it depends on various board and firm characteristics: one size does not fit all firms. We develop extensions to our basic model to analyze the optimal composition of the board between firm insiders and outsiders and the effect of board members observing imprecise public signals in addition to their private signals on board decision-making. Finally, we develop a dynamic extension to our basic model to analyze why many boards do not fire CEOs even when they preside over a significant, publicly observable, reduction in shareholder wealth over a long period of time.

3.1 Introduction

If you shoot at a king, you must kill him.

Ralph Waldo Emerson

The seeming inability of corporate boards to monitor the performance of firm management proactively, and thus prevent corporate crises, has become the topic of debate among both practitioners and academics over the last few years in the wake of various corporate scandals. Citing several case studies where boards allowed CEOs to destroy shareholder value over a number of years before pressuring the CEOs to resign, Monks and Minow (1995) ask: "Why does it take boards so long to respond to deep-seated competitive problems? And, if one of the leading responsibilities of directors is to evaluate the performance of the CEO, why do boards wait too long for proof of managerial incompetence before making a move?" A number of related questions also arise in the above context. First, are there situations where boards do not fire CEOs, even when a majority of board members are individually convinced that the CEO is of poor quality? In particular, what are the characteristics of boards that fire poor quality CEOs proactively versus waiting and firing such CEOs only after considerable destruction in shareholder value has occurred? Second, what is the relationship between board size and the effectiveness of board decision-making in terms of monitoring the CEO? Third, what is the relationship between board composition (the proportion of outside directors) and the effectiveness of board decision-making and forced CEO turnover? Finally, how will the various policy proposals for corporate governance and board reform that have been suggested by academics, practitioners, and policy-makers affect the ability of the board to appropriately monitor the CEO and fire him or her as necessary?

While there has been little research on the dynamics of board's CEO firing decisions, there has been a large empirical literature in finance partially addressing some of the other questions we have raised above.[1] However, the above evidence has been mixed. For example, while some papers argue that smaller boards of directors are more effective and therefore maximize shareholder value (e.g., Yermack (1996)), others challenge this notion, documenting that, depending on various firm characteristics, either large or small board sizes may be appropriate from the point of view of enhancing shareholder value (see, e.g., Coles, Daniel, and Naveen (2008)). Similarly, the evidence is also inconclusive on the question of whether a larger proportion of inside directors enhances shareholder value: see, e.g., Linck, Netter, and Yang (2008), Boone, Field, Karpoff, and Raheja (2007), Coles, Daniel, and Naveen (2008). While theoretical models of corporate boards (see, e.g., Hermalin and Weisbach (1998), Harris and Raviv (2008), Adams and Ferreira (2007)) have helped to resolve some of the above ambiguities by providing guidance for empirical research, a number of interesting questions remain unanswered. The objective of this paper is to provide answers to such questions by developing a theoretical analysis of corporate boards and their role in forcing CEO turnover based on trade-offs that have not been studied before in the literature.

We consider a firm with an incumbent CEO, whose quality is uncertain, and a board consisting of a number of directors whose role is to evaluate the CEO and decide whether to fire her or to retain her. Each board member receives an independent private signal about the CEO's quality, after which board members vote regarding whether or not to fire the CEO. While we consider other voting rules, the bulk of our analysis is carried out under the assumption that a majority of votes is required to fire the CEO. If the CEO is fired, the board hires a new CEO from the pool of

[1] Ertugrul and Krishnan (2011) empirically study the characteristics of boards that fire CEOs proactively, before the firm sustains a significant loss in shareholder value.

candidates available. The quality of the firm's CEO is publicly revealed in the long run, and the firm's long-run share price is determined by this revealed quality of its CEO. We assume that each board member owns some equity in the firm, so that they have an incentive to fire a bad CEO and hire a good one (or retain a good CEO), based on their assessment of the CEO's quality. However, if a board member votes to fire the incumbent CEO but fails to oust him, the CEO can impose significant costs of dissent on him (by not re-nominating that director to the board or otherwise imposing costs on him even if he continues on the board).[2]

We first study the benchmark equilibrium where the board members' dissent-costs are zero. We show that, in the absence of dissent-costs, each board member votes informatively (i.e., they vote to fire the CEO if they get a bad signal and vote to retain her if they get a good signal). Further, in the absence of dissent-costs and if the board's voting rule is optimal, we show that the board's firing decision efficiently aggregates all board members' private signals, in the sense that the board fires the CEO only if, in the opinion of a social planner who has access to all board members' signals, the incumbent CEO is worse (in terms of expected quality) than a potential replacement; the board retains the CEO if the reverse is true. We also show that, under reasonable conditions, a simple majority voting rule is optimal in the above sense.

We then analyze our basic model with dissent-costs. We show that, once such dissent-costs are significant, coordination problems arise between board members, so that the board may become suboptimally passive. In particular, the board may choose to retain the CEO even when a majority of board members are privately

[2] Mace (1971) discusses anecdotal evidence of CEOs exercising authority in selecting candidates for the board, in effect hand-picking nominees. Similarly, Lorsch and MacIver (1994) report survey evidence indicating that CEOs wield major influence in selecting new board members. Tejada (1997) presents a news account of an outside director of a prominent company being denied nomination for reelection after criticizing management.

convinced that the CEO is of poor quality. This is because, even when each board member has a bad signal about the CEO's quality, his probability assessment that enough other board members will vote against the CEO (and thus will be able to successfully fire the CEO and avoid incurring dissent-costs) is not large enough, so that the board member is better off voting to retain the CEO. In this case, we show that whether each board member votes informatively or not depends on how well they are incentivized: if each board member's equity holdings in the firm is large enough, so that his dissent-cost to equity loss ratio is low, then he votes informatively; he votes to retain the CEO even if he receives a bad signal about the CEO's ability if this ratio is high.

We then analyze the effect of board size on the quality of the board's decision-making. On the one hand, since each board member's signal about the CEO's quality is independent, a larger number of board members increases the amount of information that is potentially available (collectively) to the board. On the other hand, a larger board worsens the coordination problems across board members, since in a larger board, a larger number of votes is required to fire the CEO. We show that the optimal board size emerges from the above trade-off. Further, this optimal board size depends on the precision of each board member's signal and his dissent-cost to equity loss ratio. Our results contradict the hypothesis of Lipton and Lorsch (1992) and Jensen (1993) that larger boards are always less effective, and sheds light on the findings of the recent empirical literature, which documents that while smaller board sizes are value maximizing for simpler firms, larger board sizes seem to be optimal for more complex firms (with greater information requirements for evaluating the CEO's performance): see, e.g., Coles, Daniel, and Naveen (2008).

In section 5, we extend our basic model to analyze issues of optimal board composition. Here we distinguish between insider board members and independent outsiders:

while outsiders are similar to the board members in our basic model (who receive independent private signals about the CEO's quality, and who cannot credibly communicate their signals to each other prior to voting), insiders' signals are correlated with each other, and they can credibly communicate their private signals to one another.[3] The advantage of insider dominated boards is that they do not suffer from coordination problems; the disadvantage is that, since insiders' signals are correlated, the board is effectively making a decision based on much less information compared to outsider dominated boards. By the same token, the advantage of outsider dominated boards is that, since each outsider has an independent signal, such boards have a much greater amount of information (collectively) available to them; the disadvantage of outsider dominated boards is that, since outsider directors (similar to directors in our basic model) suffer from coordination problems, they may not be able to use this information efficiently in many situations.

The optimal composition of boards emerges from the above trade-off between having a larger number of outsiders versus insiders: while in some situations outsider dominated boards are optimal, in other situations a board consisting of a combination of insiders and outsiders is optimal; insider dominated boards are optimal in yet other situations. We show that the optimal composition of the board depends on the dissent-cost to equity loss ratio of board members, and the informational requirements of evaluating the CEO's performance in managing the firm. Our results thus contradict the conventional wisdom that outsider dominated boards are always value-maximizing. Further, they explain the recent empirical evidence documenting

[3]Given that insiders, by definition, are connected to the firm and the incumbent CEO, it seems to be natural to assume that insiders' signals about the CEO's quality will be correlated. For the same reason, it is likely that it is easier for insiders (given their common background and potential familiarity with each other) to credibly communicate their signals to each other compared to the case for outsiders (who may come from diverse backgrounds, and may be relatively unfamiliar with each other; outsiders may have competing interests as well, making credible communication difficult).

that, while a larger number of outsiders is value-maximizing for complex firms with greater informational requirements in evaluating and monitoring the CEO (see Coles, Daniel, and Naveen (2008)), this is not necessarily the case for simpler firms and those where firm-specific knowledge (likely to be available to insiders) is more important (see Linck, Netter, and Yang (2008) and Coles, Daniel, and Naveen (2008)). It also explains the evidence in Boone, Field, Karpoff, and Raheja (2007), who document that the proportion of independent outsiders on corporate boards is negatively related to measures of the CEO's influence over the board (which increase dissent-costs) but positively related to the equity ownership of outside directors (which increase their equity loss).

In section 6, we extend the basic model in a different direction: we study the effect of a public signal on the coordination problem among directors in the basic model. One might at first think that publicly observable information regarding the CEO's quality (for example, the accounting performance of the firm under the CEO's management) is a natural coordination mechanism. In order to examine this question, we assume in this section that, in addition to their own private signals, all directors observe a noisy public signal of the CEO's quality, before voting.[4] In this setting, we show that the effect of the signal on board decision-making is asymmetric. On the one hand, it is indeed the case that a negative public signal improves board decision-making, in the sense that board members vote informatively for ranges of the dissent-cost to equity loss ratio where, in the absence of such a signal, they would have voted to retain the CEO (ignoring their private signals). On the other hand (and more interestingly), a positive public signal worsens board decision-making, in

[4] Our assumption here is that the public signal is imprecise, so that it does not make the board members' private signals redundant. Of course, if the public signal is entirely precise, all board members will ignore their private signals and will vote to fire the CEO after they observe a negative public signal (for example).

the sense that board members vote to retain the CEO in the presence of such a signal for values of the dissent-cost to equity loss ratio for which they would have chosen to vote informatively (i.e., according to their private signals) otherwise. The intuition here is that, when a board member receives a negative private signal but a positive public signal, he may choose not to vote against the CEO, since, given the positive public signal, he assesses a greater probability that many other board members vote to retain the CEO, thus allowing the CEO to stay in power and impose dissent costs on any director who voted against her. In summary, we show that having noisy public measures of CEO performance available to the board will not always improve board decision-making, since such signals can sometimes make the board members' coordination problem worse rather than better.

Finally, in section 7 we develop a dynamic version of our basic model to address the question we raised at the beginning of this paper, namely, why many corporate boards wait too long in the face of shareholder value destruction before firing the CEOs involved, and the difference between the characteristics of such boards and those that fire CEOs proactively (before significant destruction in shareholder value takes place). In this dynamic model, we introduce a second round of voting on the incumbent CEO (if she was not fired as a result of the first round of voting), with a highly informative public signal (such as whether significant destruction in shareholder value has taken place or not) between the two rounds of voting. We also assume that the equity loss of the board is significantly higher if a low quality CEO is fired only after board members receive the public signal (consistent with the higher loss in the value of their equity holdings that board members would sustain compared to the situation where they replace the low quality CEO earlier).

In the above dynamic setting, we show that the coordination problem faced by board members may explain the board's reluctance to act to remove CEOs until considerable value-destruction has taken place. The dynamic trade-off faced by board members is the following. On the one hand, as time goes by, the firm's share price goes further and further down, thus increasing the board members' equity loss. On the other hand, as more and more evidence of the CEO's incompetence arrives (publicly) over time, each board members' assessment of a sufficient number of other members also voting against the CEO to oust her increases, thus reducing his own probability of incurring a dissent-cost if he votes against her. The timing of the CEO's dismissal emerges from this dynamic trade-off between expected equity loss and expected dissent-cost. We show that boards that are better incentivized (in terms of dissent-cost to equity loss ratio) will fire CEOs earlier, before significant shareholder value-destruction has taken place. The predictions of our dynamic model explain the empirical findings of Ertugrul and Krishnan (2011), who document that boards that dismiss CEOs early (i.e., in the absence of significant negative prior stock returns) are those that are characterized by higher equity ownership by board members, higher board equity compensation as a fraction of total compensation, higher institutional ownership, and better firm corporate governance characteristics.

Even though our model is set in the context of a board deciding whether to force a CEO turnover, the intuition behind our model generalizes to other binary corporate decisions made by a board as well. For example, consider the case of a CEO proposing a takeover of another firm, and the board voting on the CEO's proposal, with each board member having a private signal about the desirability of the takeover from the point of view of shareholder value maximization. While, unlike in the turnover setting, rejecting the CEO's preferred policy will not rid the firm of her, the basic trade-offs of our model will go through even in a non-turnover setting as long as we make an additional assumption that the costs imposed by the CEO on a dissenting board member will be lower if the majority of directors went against the

CEO's proposal. Here again the relationship between dissent-costs and equity-loss will determine the equilibrium. Larger boards are less likely to sustain informative voting if directors' signals are relatively imprecise and there emerges an optimal board size even in this case. Also, insiders (who we assume have correlated signals) will be more likely to avoid the co-ordination problem but will have to make decisions based on less information, so that there will again be an optimal combination of insiders and outsiders on the board in equilibrium. The role of publicly available information will also be similar to that in a turnover setting in the sense that having a public signal may not always improve the quality of decision-making.

The rest of the paper is organized as follows. Section 2 relates our paper to the existing literature. Section 3 describes the set-up of our basic model. Section 4 characterizes its equilibrium and analyzes issues related to optimal board size. Section 5 extends our basic model to analyze optimal board composition. Section 6 extends our basic model to study the effect of an imprecise public signal on the effectiveness of board decision-making. Section 7 presents a dynamic extension of our basic model to study the timing of forced CEO turnover. Section 8 discusses the empirical and policy implications of our model. Section 9 concludes. The proofs of all propositions are confined to the appendix.

3.2 Relation to existing literature

There have been several theoretical models of corporate boards in the existing literature, driven by trade-offs different from the ones we analyze here. The first theoretical analysis in this literature was by Hermalin and Weisbach (1998). In their model, board structure is the outcome of negotiation between the CEO and outside directors. CEOs who generate surplus for their firms (for whom good substitutes are unavailable) wield considerable influence over their outside directors and use this influence to capture some of the surplus they generate by placing insiders in open board positions. Adams and Ferreira (2007) develop a model in which the CEO's preferred projects, which yield him control benefits, differ from those of shareholders. The CEO faces a trade-off in disclosing information to the board: if he reveals his information, he receives better advice from it, but may lose control benefits, since an informed board will also monitor him more intensively. They thus show that management-friendly boards, which can pre-commit not to use the information provided by the CEO to monitor him, may be value-maximizing for the firm.[5]

Harris and Raviv (2008) develop a model of a corporate board consisting of both insiders and outsiders and which can profitably use the information held by both insiders and outsiders to make optimal project choices. Insiders have private information relevant to this choice, but have private benefits that lead their incentives to be misaligned with those of shareholders; outsiders, whose interests are perfectly aligned with those of shareholders, are initially uninformed but can engage in producing information relevant to project choice at a cost. Control of the board entitles the controlling party (insiders or outsiders) either to make decisions themselves or to delegate decisions to the other party. In this setting, they characterize the conditions under which there will be insider versus outsider control of the board, when each party will delegate decision-making to the other party, the extent of communication between the two parties, and the number of outside directors.

Raheja (2005) also models a board consisting of both insiders and outsiders in charge of project selection, where insiders have decision-relevant private information, have private benefits that distort their incentives, and where outsiders can engage in

[5] See also Ferreira, Ferreira, and Raposo (2011), who develop a model demonstrating that stock price informativeness affects the structure of corporate boards, and test its implications.

costly production of information relevant to project choice. Outsiders' information costs are reduced if insiders reveal their private information. She argues that, since insiders are the source of future CEOs in her model, outsiders can use their CEO succession votes to motivate insiders to reveal their private information. She solves for the combination of insiders and outsiders on the board that leads to optimal project selection. Gillette, Noe, and Rebello (2003) show both theoretically and experimentally that when agency problems are especially severe, having uninformed outsiders in control of a board can prevent inefficient outcomes. Hirshleifer and Thakor (1994) model the interaction between internal (corporate boards) and external (acquisitions) governance mechanisms in the maintenance of managerial quality.

While some of the above models also address two of the central issues regarding corporate boards that we study here, namely, board size and optimal board composition, the above models address these questions based on trade-offs quite different from that in our model.[6] Further, none of these models study issues related to the effect of public signals on board decision-making, and the dynamics of CEO firing decision that we study here.[7] Our paper is also related to the large empirical literature on corporate boards and the relation between boards and firm characteristics, which we discuss in section 8 in the context of the empirical and policy implications of our model: see Adams, Hermalin, and Weisbach (2010) for an excellent review of

[6]In a subsequent paper, Malenko (2011) analyzes a setting very similar to ours to study issues of communication among board members. While some of the implications of her model are similar to ours, she introduces the possibility of board members being able to credibly communicate with each other by incurring an additional cost and develops the implications of this assumption.

[7]Warther (1998) develops a simple model of a two-member board in which the CEO can eject hostile board members, which can be viewed as the imposition of a dissent-cost on board members who disagree with the CEO. He shows that this results in board members being reluctant to oppose management and votes being often unanimous in favor of management. However, given that he exogenously assumes a two-member board, he is unable to study any of the issues of coordination among board members that we analyze here; neither does he analyze optimal board size, board composition, the effect of public signals, or the dynamics of CEO turnover. Almazan and Suarez (2003) show that passive (weak) boards may be optimal in a setting where severance pay and weak boards are substitutes for costly incentive compensation paid to the CEO.

the literature on corporate boards.[8]

3.3 The model

Our basic model consists of four dates. It starts at time 0 with an incumbent CEO in place. The CEO can be of either good or bad quality (denoted G for good and B for bad). The board consists of n directors and does not include the CEO (we will use the terms board member and director interchangeably throughout the paper).[9] At time 1 each of n board members obtains a signal, s_i, $i = 1, \ldots, n$, about the CEO's quality. At time 2, based on the signals they received at time 1, board members vote whether to keep the current CEO or to replace her. If the board fires the CEO, it must hire a replacement from the pool of available candidates. The quality of this pool is indexed by γ: γ is the ex-ante probability of hiring a high-quality replacement. At time 3 the true quality of the then-incumbent CEO is revealed (either of the CEO existing at time 0 if she was not fired or the new CEO hired at time 2). The game ends and all payoffs are distributed. The timeline of the model is depicted in figure 1. All agents are risk-neutral, the risk-free rate of return is zero.

The signal that each board member receives can be either good, denoted g, or bad, denoted b. The signals are assumed to be independent across directors and informative. In particular, $p(s_i = g|G) = p(s_i = b|B) = \alpha > \frac{1}{2}$. Thus, α gives the precision of each board member's signal about the CEO's quality. We assume that all board members have signals of identical precision. The prior probability assessment

[8]Our paper is also related to the literature on voting in committees: see, e.g, Crawford and Sobel (1982), Coughlan (2000), Doraszelski, Gerardi, and Squintani (2003) and Austen-Smith and Feddersen (2005), who allow for pre-vote communication and study a cheap-talk setting. Our setting differs from the above papers in the sense that we do not allow pre-vote communication and do not focus on a cheap-talk setting.

[9]This assumption is introduced for ease of exposition and is innocuous if one assumes that the CEO always votes to retain herself. Adding the CEO to the board in this case will simply decrease the effective board size to $n - 1$.

of a board member that the incumbent CEO is good is μ.

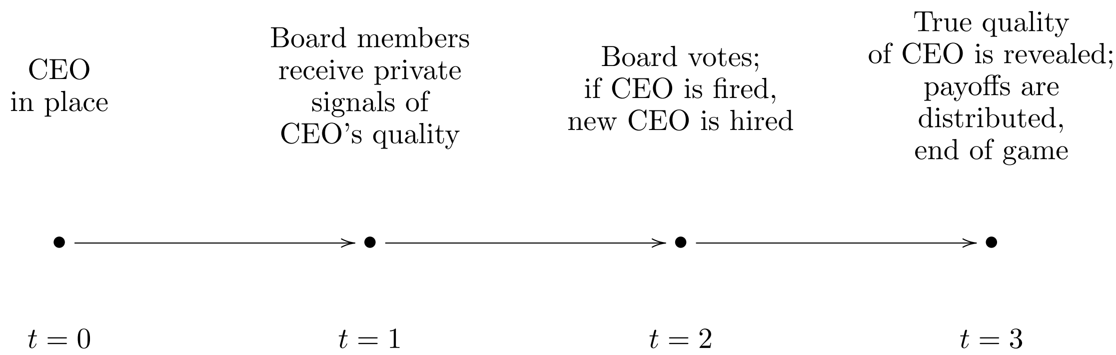

Figure 3.1: Timeline of the basic model.

3.3.1 Board members' objective

The objective of each board member in making his voting decision is to retain the CEO if she is of good quality and fire her if she is of bad quality (and hire a good quality replacement if the incumbent CEO is fired by the board) but only to the extent that this affects his own costs and benefits, as described below.[10] We assume without loss of generality that the maximum payoff a director can receive is 0.[11] If it is revealed at time 3 that the CEO chosen by the board is of bad quality, each board member suffers a cost, $q, q > 0$. q is the loss in value to directors from making a bad decision that arises from the reduction in the value of the firm's equity held by individual board members (due to the mismanagement of the firm by a low quality CEO). We will refer to q from now on the the directors' "equity loss." Under this interpretation, one can think of the equity loss as being proportional to the difference in the long-run value

[10] If good quality CEOs can generate higher long-run cash flows for the firm than bad quality CEOs, this translates into the assumption that the board wishes to maximize the firm's long-run value net of personal costs incurred.

[11] This is a normalization. If we assumed a fixed amount here (e.g., representing the board member's compensation that is independent of firm value), our results would go through qualitatively unchanged.

of the firms's equity under a good quality CEO versus a bad quality CEO. Thus, one way to increase a board member's equity loss would be to increase the equity fraction in his compensation. q can also be thought of as reputational damage or a direct cost imposed by shareholders on the members of a board which retains a bad CEO.

In addition to the above cost of choosing a low quality CEO, each board member may suffer a cost of dissent. In particular, if a director votes to fire the CEO but fails to receive enough support from fellow board members, he suffers a cost, $c, c > 0$. This "dissent-cost" can be thought of as potential punishment imposed by the CEO (such as the CEO's reluctance to nominate the director in the next election cycle) or as the reputational damage from being perceived as a "troublemaker" by other directors on this board or by other boards who refuse to grant this director a seat. What matters to individual directors is the magnitude of c relative to q. We will refer to $\frac{c}{q}$ as the directors' dissent-cost to equity loss ratio.

Director i's vote is denoted by v_i. We say that $v_i = 0$ if director i votes to keep the CEO and $v_i = 1$ if director i votes to fire the CEO. Given the above, director i's objective when making his voting decision, denoted by π_i, is given by:[12]

$$\begin{aligned}\pi_i(v_i; s_i, v_{-i}, k, n, c, q, \alpha, \gamma, \mu) = &- \Big\{ I\big(\sum_{j=1}^{n} v_j \geq k\big) q(1-\gamma) + I\big(\sum_{j=1}^{n} v_j < k\big) q p(B|s_i) \Big\} \\ &- I\big(\sum_{j=1}^{n} v_j < k\big) c v_i, \end{aligned} \tag{3.1}$$

where I denotes an indicator function in each term (it takes the value of 1 if the relation in parenthesis is satisfied and 0 otherwise), v_{-i} is the voting configuration of all directors except i, and $p(B|s_i)$ denotes the probability that the current CEO is of

[12]Notice that there is no discounting since the risk-free rate is assumed to be zero and all agents are risk-neutral.

low quality given director i's signal. The terms in curly brackets in (3.1) represent the cost of choosing a bad CEO. The first term is the expected cost of choosing a bad CEO when the current CEO is fired: the corresponding indicator function, $I\big(\sum_{j=1}^{n} v_j \geq k\big)$, takes on the value of 1 when there are enough votes to oust the CEO ($\sum_{j=1}^{n} v_j \geq k$). In that case, the probability that the replacement is of low quality is given by $1-\gamma$, and hence the expected equity loss is $-q(1-\gamma)$. The second term in curly brackets is the expected equity loss from retaining the incumbent: the corresponding indicator function takes on the value of 1 when there are not enough votes to oust the CEO ($\sum_{j=1}^{n} v_j < k$). In that case, the probability that the board member will suffer equity loss is given by $p(B|s_i)$ so that the expected equity loss is $-qp(B|s_i)$. The last term in (3.1), outside the curly brackets, represents the expected dissent-cost. A board member incurs it when two conditions are satisfied: (i) he votes to fire the CEO ($v_i = 1$), and (ii) there are not enough votes to oust the incumbent ($\sum_{j=1}^{n} v_j < k$, so that the corresponding indicator function takes on the value of 1). For example, if board member i votes to fire the CEO but fails to attain enough support from other board members, his expected payoff will be $-qp(B|s_i) - c$.

3.3.2 Board members' voting rule

Each director decides how to vote based on his updated assessment (based on his prior probability μ and his private signal) of the quality of the CEO, his conjecture about the voting behavior of other directors (since this affects his chance of being on the winning side if he chooses to vote to fire the CEO, for example), and his equity loss and dissent-cost, which together determine the value of his objective (3.1). Using Bayes' rule, each board member's updated probability assessment of the CEO being of bad quality conditional on his private signal, denoted $p(B|s_i)$, is given by:

$$p(B|g) \equiv p(B|s_i = g) = \frac{(1-\mu)(1-\alpha)}{(1-\mu)(1-\alpha)+\mu\alpha}, \tag{3.2}$$

$$p(B|b) \equiv p(B|s_i = b) = \frac{(1-\mu)\alpha}{(1-\mu)\alpha+\mu(1-\alpha)}. \tag{3.3}$$

The assumption of informative signals guarantees that

$$p(B|b) > 1-\mu > p(B|g). \tag{3.4}$$

Definition 3.1. *Voting rule, k, is the minimum number of votes required to fire the CEO.*

Definition 3.2. *Voting is sincere if $v_i = 1$ iff $p(B|s_i) > 1-\gamma$.*

Definition 3.3. *Voting is informative if $v_i = 1$ iff $s_i = b$.*

Sincere voting implies that each board member will vote to fire the CEO if after observing the private signal, this director's assessment of the CEO's quality is lower than that of the potential replacement. Informative voting implies that the director reveals his private information: i.e., the board member will vote in the direction of his private signal about the CEO's quality. We will assume throughout that $p(B|b) > 1-\gamma > p(B|g)$; otherwise, sincere voting can never be informative. Consider, for instance, the case when $p(B|b) > p(B|g) > 1-\gamma$. Then, even after observing private signal g, a director's private assessment of the incumbent CEO's quality is lower than of a potential replacement. With this assumption, on the other hand, informative voting is always sincere.

The board makes decisions according to a voting rule defined above. Assume that there exists a benevolent social planner who can observe all board members' private signals. Assuming that the number of directors who observe a bad private signal is r, the social planner's updated probability of the incumbent's quality is given by

$$p(B|r) = \frac{(1-\mu)\alpha^r(1-\alpha)^{n-r}}{(1-\mu)\alpha^r(1-\alpha)^{n-r} + \mu(1-\alpha)^r\alpha^{n-r}}. \tag{3.5}$$

The social planner would therefore fire the incumbent if and only if

$$\frac{(1-\mu)\alpha^r(1-\alpha)^{n-r}}{(1-\mu)\alpha^r(1-\alpha)^{n-r} + \mu(1-\alpha)^r\alpha^{n-r}} > 1-\gamma. \tag{3.6}$$

Let r^* be the smallest integer that makes (3.6) hold. We then define the optimal voting rule as the one that sets $k = r^*$.[13] In other words, under the optimal voting rule, the CEO is fired only when she is considered worse than outside replacement, given all directors' cumulative private information. Notice that r^* generally depends on the parameters of the situation. However, in an important special case that may be applicable to real life boards, simple majority is the optimal voting rule.

Proposition 3.1. *(Optimal voting rule). Assume that $\mu = \gamma$ and that n is odd. Then, the optimal voting rule is given by $k = \frac{n+1}{2}$ (simple majority).*

It seems feasible that in real-life boards, the quality of potential replacement (as captured by γ) is often equal to board members' prior (μ) about the underlying quality of the current CEO, so that a simple majority voting rule is the optimal rule.

3.4 Equilibrium of the basic model

In this section we characterize the equilibrium of our basic model. The equilibrium concept we use is symmetric pure strategy Efficient Bayesian Nash equilibrium. An

[13] Our definition of an optimal voting rule parallels the existing voting literature: see, e.g., Austen-Smith and Banks (1996).

equilibrium consists of: (*i*) a vote by each director (either to fire or to keep the current CEO); (*ii*) each director's conjecture about the signals of the other directors. Each director's vote maximizes his expected payoff given his beliefs about the distribution of signals among the other board members and their voting strategies.

3.4.1 The benchmark equilibrium with no dissent-costs

Before we solve for the equilibrium of our basic model and its enriched versions (in later sections), we consider the benchmark case in which each director's dissent-cost, c, is zero, i.e., the case when the only cost imposed on board members is the equity loss. The following proposition describes the equilibrium in this case.

Proposition 3.2. *(Benchmark Equilibrium).*

(*i*) *Assume that the board adopts the optimal voting rule. Then, the benchmark case described above has a unique equilibrium, in which each board member votes informatively. In particular, board member i votes to keep the CEO iff $s_i = g$ and to fire the CEO iff $s_i = b$.*

(*ii*) *A board that adopts an optimal voting rule efficiently aggregates all board members' information in the absence of dissent-costs, in the sense that if the board decides to retain the CEO, her expected quality is better than that of a potential replacement; if the board decides to fire the CEO, her expected quality is worse than that of a potential replacement.*

This equilibrium achieves the efficient outcome preferred by a benevolent social planner. Since all directors vote informatively and the voting rule is chosen optimally, the outcome of this vote is equivalent to aggregating all directors' private information in making the final decision of the board regarding whether to fire the incumbent CEO or to retain her.

Another important observation concerns the optimal board size. If the board effectively aggregates private information, it is clearly optimal to have as many independent signals is possible. This logic is made precise in the following proposition.

Proposition 3.3. *(Optimal board size without dissent-costs). Assume that n is odd and majority is the optimal voting rule. Then, the quality of the social planner's decision is increasing in the number of board members, n.*

Proposition 3.3 says that when board members are willing to reveal their private valuation it is better to have as many of them as possible, since this increases the amount of information collectively available to the board (or the social planner, in the benchmark case).

3.4.2 Equilibrium of the basic model with dissent-costs

We now characterize the equilibrium of our model in the presence of dissent-costs, which leads to imperfect coordination among board members.

Proposition 3.4. *(Equilibrium of the basic model). Index board members by i and let $e_i = 1$ if the i^{th} board member observes a bad signal ($s_i = b$); $e_i = 0$ if the signal is good ($s_i = g$). Then the game described above has a unique equilibrium, so that:*

(i) If the dissent-cost to equity loss ratio $\frac{c}{q} \leq \Upsilon(k, n, \alpha, \gamma, \mu|b)$, each director votes informatively, in particular, director i votes to retain the CEO iff $s_i = g$ and to fire the CEO iff $s_i = b$;

(ii) If $\frac{c}{q} > \Upsilon(k, n, \alpha, \gamma, \mu|b)$, each director votes to retain the CEO regardless of his private signal, where $\Upsilon(k, n, \alpha, \gamma, \mu|b) = \frac{p\left(\sum_{j=1}^{n} e_j \geq k \middle| b\right)\left(p(B|b) - (1-\gamma)\right)}{p\left(\sum_{j=1}^{n} e_j < k \middle| b\right)}$ and all conditional probabilities represent beliefs of director i given his signal.

The intuition for Proposition 3.4 is as follows. The benefit from keeping the current CEO for a director with a good signal is the chance to avoid the equity loss q. Since directors who vote to keep the current CEO do not suffer the cost of dissent if the CEO is fired, they have nothing to lose by voting truthfully. It follows that board members who receive bad signals of the current CEO's quality determine the outcome of the vote. If they vote to fire the CEO, they reduce the probability of suffering the equity loss q. However, if they vote truthfully and fail to attain enough support they will suffer c, the cost of dissent, regardless of whether the CEO is good or bad. Hence, if a director observes a signal b, he faces the trade-off between equity loss and the dissent-cost. When c is high relative to q, board members will choose to disregard their signals and vote to keep the current CEO. Hence, it is the ratio $\frac{c}{q}$ that will determine whether board members vote informatively or not. The threshold value of $\frac{c}{q}$ above which truth telling is impossible is given by

$$\Upsilon(k, n, \alpha, \gamma, \mu|b) = \frac{p\left(\sum_{j=1}^{n} e_j \geq k \middle| b\right)\left(p(B|b) - (1-\gamma)\right)}{p\left(\sum_{j=1}^{n} e_j < k \middle| b\right)}. \tag{3.1}$$

Notice that (3.1) makes intuitive sense. $\Upsilon(k, n, \alpha, \gamma, \mu|b)$ is positively related to the probability that enough directors observe bad signals, given that director i observes b. It is also positively related to the probability that the current CEO is of low quality given that director i observes b: the higher this probability, the more strongly directors with a bad signal believe that the incumbent's quality is low. It is also positively related to the quality of outside replacement, γ (negatively related to $1-\gamma$): a high expected quality of replacement makes it more worthwhile to fire the incumbent. On the other hand, $\Upsilon(k, n, \alpha, \gamma, \mu|b)$ is inversely related to the conjectured number of positive signals observed by board members (given director i's own signal is b): if director i believes that only a small number of other board members obtained a bad signal, he understands that it is unlikely they will have enough votes to oust the current CEO, so that he will vote to retain the CEO even after observing a bad signal in order to avoid the dissent-cost c.

It is natural to consider the ratio $\frac{c}{q}$ to measure the board's incentives.

Definition 3.4. *Consider two boards, A and B, and denote the corresponding dissent-costs and equity loss by c_A, q_A, c_B, and q_B, respectively. We say that board A is better incentivized than board B if $\frac{c_A}{q_A} < \frac{c_B}{q_B}$.*

It is immediate from Proposition 3.4 that better incentivized boards are more likely to sustain informative voting (in the sense defined above). Our analysis emphasizes two sources of influence that a director faces: the CEO (and potentially other directors) via dissent-costs c, and the shareholders and directors' own shareholdings through the equity loss q. Providing better incentives involves changes to either or both of those parameters.

Proposition 3.5. *(Suboptimally passive boards). In the presence of dissent-costs, boards are likely to keep the incumbent CEO more often than is socially optimal.*

In particular, if dissent-costs are sufficiently high relative to the equity loss the board will not fire the CEO even if the majority of directors observe a bad signal of CEO's quality. This may explain the criticism of current boards in the financial press: in general, they are less likely to fire bad CEOs than is desirable. As long as some boards suffer from coordination problems, the average turnover decision in the economy will be suboptimal.[14]

[14]Some readers may conjecture that one way to solve the problem of suboptimally passive boards is to reduce k, the number of board members' votes required to fire the CEO. This is, however, not the case, since informative voting leads to optimal decision-making by the board only if k is chosen following the optimal voting rule (i.e., to satisfy (3.6)). For example, if k is reduced below the optimal level, the board will fire good CEOs more often than is optimal, even when all board members vote informatively. Instead of being suboptimally passive, the board becomes suboptimally active.

Proposition 3.6. *(Large vs. small boards). When simple majority is the optimal voting rule and n is odd:*

(*i*) *if $\mu < \alpha z(\alpha, n)$, then adding directors to the board increases the threshold value of $\frac{c}{q}$ above which informative voting is unsustainable,*

(*ii*) *if $\mu > \alpha z(\alpha, n)$, then adding directors to the board decreases the threshold value of $\frac{c}{q}$ above which informative voting is unsustainable,*

(*iii*) *if $\mu = \alpha z(\alpha, n)$, then adding directors to the board doesn't change the threshold value of $\frac{c}{q}$ above which informative voting is unsustainable, where* $z(\alpha, n) = \frac{2n\alpha-(n+1)}{2n\alpha-(n+1)+2(1-\alpha)} < 1$.

Proposition 3.6 implies, in particular, that when $\alpha \leq \mu$, then larger board are less likely to sustain informative equilibria. The trade-off driving the relationship between board size and informative voting is as follows. On the one hand, a larger board brings more information into the decision making process. On the other hand, bigger boards increase the difficulty of coordinating across board members. Whether a larger board fosters or inhibits informative voting depends on whether the first effect dominates the second, or vice versa. If simple majority is the optimal voting rule two cases need to be considered. When the signal available to an individual director is very precise ($\mu < \alpha z(\alpha, n)$), then the first effects dominates and as the board's size grows, so does the likelihood of informative voting. On the other hand, if the signal is not very precise ($\mu > \alpha z(\alpha, n)$), then the second effect dominates the first, and informative voting becomes less likely when the board becomes larger.

Proposition 3.7. *(Optimal board size). If $\alpha \leq \mu$ and majority is the optimal voting rule, then there exists an optimal board size, n^*.*

The trade-off driving the preceding proposition is the same as in Proposition 3.6. A larger board collectively has access to more information but size increases the difficulty of coordination among directors. Compare Proposition 3.3 to Proposition 3.6. The former states that larger boards obtain a more precise cumulative signal, conditional on having all directors reveal their private valuation. The latter, however, provides conditions under which large boards inhibit informative voting. In general, boards must be large enough to obtain a reasonably precise signal but small enough to enable informative voting. More directors bring additional information into decision making. However, this will occur only if voting is informative. If they are unwilling to reveal their information, additional board members only exacerbate the coordination problem among board members.

Proposition 3.8. *(Comparative statics).*

(*i*) *The threshold value of* $\frac{c}{q}$ *above which informative voting is unsustainable is decreasing in* μ.

(*ii*) *If* $\mu = \frac{1}{2}$, *then the threshold value of* $\frac{c}{q}$ *above which informative voting is unsustainable is increasing in* α.

(*iii*) *The threshold value of* $\frac{c}{q}$ *above which informative voting is unsustainable is increasing in* γ.

Proposition 3.8 is quite intuitive. If the assessment of the incumbent's quality is high, it is unlikely that there will be enough negative votes to oust her, even conditional on observing a bad private signal. When signals are more precise, directors are more confident abut the quality of their information and are more willing to vote informatively. A high expected quality of a potential replacement makes it more worthwhile to fire the incumbent CEO since conditional on firing, the board is more likely to obtain a good CEO.

3.5 Insiders versus outsiders

One of the recurring topics in the literature on corporate boards is the relationship between insiders and outsiders on the board. In this section we explore the role that these types of directors play in the CEO turnover decision. As before, we assume that there are no public signals and that each director observes a private signal of the incumbent CEO's quality.

We introduce two types of board members: insiders and outsiders. We assume that there are h insiders and $n-h$ outsiders on the board. We define insiders as board members employed by the firm. Since insiders work in the same environment they are likely to have the ability to predict what signals other insiders observe. To keep the model tractable, we assume that this ability is perfect: i.e., insiders can see each other's signals. At the same time, the similarity of information received by insiders means that their signals are correlated. For simplicity of modeling, we make the extreme assumption that the correlation across insiders' signals is perfect: i.e., all insiders receive either a good signal or a bad signal.[15] We continue to assume, however, the same precision of each insider's signal as in the basic model. Thus, $p(s_i = g|G) = p(s_i = b|B) = \alpha > \frac{1}{2}$. Outsiders have the same characteristics as board members in our basic model.

In summary, we introduce two differences between insiders and outsiders in this

[15] Our results would be unchanged if, instead of assuming perfect correlation across insiders' signals, we were to assume that a fixed proportion of insiders, $\rho, 0 < \rho < 1$, receive the same signal: If the CEO is of good quality, then with probability α, ρh insiders observe a good signal, with $(1-\rho)h$ receiving a bad signal. Thus, ρ would represent the degree to which insiders as a group get similar signals. Similarly, if the CEO is bad, then with probability α, ρh insiders observe a bad signal, with $(1-\rho)h$ receiving a good signal.

section. First, on the positive side, insiders do not suffer from coordination problems, unlike outsiders (who suffer from the inability to coordinate their voting behavior that we discussed for all directors throughout previous sections). On the negative side, however, since insiders' signals are correlated, each insider does not bring as much information to the board's decision as an outsider, who receives a signal that is independent (of other outsiders as well as insiders). All other characteristics (in terms of dissent-cost, equity loss and so on) are the same for insiders and outsiders, and various other assumptions in this section remain the same as in the basic model.[16]

Let f denote insiders' aggregate signal. In particular, we let $f = b$ if all insiders observe a bad signal and $f = g$ if insiders receive a good signal. Notice that the inference that insiders make is equivalent to having one director observe a signal based on their cumulative information. In particular,

$$p(B|f = g) = \frac{(1-\mu)(1-\alpha)}{(1-\mu)(1-\alpha) + \mu\alpha} = p(B|g), \tag{3.1}$$

$$p(B|f = b) = \frac{(1-\mu)\alpha}{(1-\mu)\alpha + \mu(1-\alpha)} = p(B|b). \tag{3.2}$$

3.5.1 Equilibria

Before we discuss equilibria we note that due to their perfect coordination insiders will vote as a block. Since they observe each other's signal, they aggregate this information to update their private valuation. In fact, after observing every other insider's signal, any inside director has the same information as any other inside director. Hence, they cannot vote differently in a symmetric equilibrium.

[16]One may conjecture that inside directors may also have larger dissent costs compared to outside directors since their careers may be tied closely to that of the CEO: see, e.g., Weisbach (1988). As we argue later (see footnote 18), incorporating this feature will not change the qualitative nature of our results.

Equilibrium when the board is insider dominated

First we consider the case when insiders dominate the board ($h \geq k$). Since insiders always vote as a block and their block is large enough to determine the outcome, the way outsiders vote is irrelevant. Hence, it is optimal for outsiders to always vote to keep the incumbent CEO. This way outsiders always avoid the cost of dissent. To see this clearly, consider the case when an outsider observes a bad signal of CEO's quality. If he votes to fire the incumbent while insiders don't, this outsider will suffer the cost of dissent. If, on the other hand, insiders vote to fire the CEO, his vote makes no difference: the CEO will leave even if this outsider votes to keep her. Thus, the payoff from voting to keep the CEO is always at least as high as the payoff from voting to fire the CEO even for outsiders with a bad signal. The following proposition characterizes the equilibrium in this case.

Proposition 3.9. *(Equilibrium when insiders dominate the board). If insiders dominate the board ($h \geq k$), they vote to fire the CEO if and only if $f = b$. Outsiders always vote to retain the CEO, regardless of their private signals.*

Proposition 3.9 implies that when insiders dominate the board, they completely ignore outsiders' information and either retain the CEO or fire her in a "palace coup." The fact that insiders dominate the board and vote as a block means that they can always avoid the cost of dissent. If they vote to fire the CEO, she is let go and insiders do not suffer any dissent-cost. However, a board dominated by insiders may act on very little information, since insiders' signals are correlated with each other, and the board ignores any information held by outsiders.

Equilibrium when the board is outsider dominated

We now consider boards where insiders do not have enough power to determine the outcome of the game ($h < k$). Insiders still have access to every insider's information and therefore they still vote as a block. However, they need to assess the probability of outsiders observing a particular signal and voting according to that signal. Outsiders, in their turn, have to assess the probability that insiders observe a bad signal and side with outsiders in ousting the CEO. Otherwise, they need to assess the probability that they can have enough votes to fire the CEO without insiders' support. Recall that, since outsiders receive independent signals (and cannot see each other's signals), they suffer from the coordination problems that we discussed under our basic model.

Proposition 3.10. *(Equilibrium when outsiders dominate the board). Index outside board members by i and let $e_i = 1$ if the i^{th} board member observes a bad signal ($s_i = b$); $e_i = 0$ if the signal is good ($s_i = g$). Then the game described above has a unique equilibrium, so that:*

(*i*) *If the dissent-cost to equity loss ratio $\frac{c}{q} \leq \Theta(k, n, q, \alpha, \gamma, \mu, h)$, each outside director votes informatively, in particular, director i votes to retain the CEO iff $s_i = g$ and to fire the CEO iff $s_i = b$, and all insiders vote to fire the CEO iff $f = b$ and to retain her if $f = g$;*

(*ii*) *If $\frac{c}{q} > \Theta(k, n, q, \alpha, \gamma, \mu, h)$, each director votes to retain the CEO regardless of his private signal, where*

$$\Theta(\cdot, h) = \frac{p(f{=}b|b)p\left(\sum_{j=1}^{n-h} e_j \geq k-h \middle| b\right) + p(f{=}g|b)p\left(\sum_{j=1}^{n-h} e_j \geq k \middle| b\right)}{1 - p(f{=}b|b)p\left(\sum_{j=1}^{n-h} e_j \geq k-h \middle| b\right) - p(f{=}g|b)p\left(\sum_{j=1}^{n-h} e_j \geq k \middle| b\right)} \big(p(B|b) - (1-\gamma)\big).$$

All probabilities above represent beliefs of director i given his signal (in the case of insiders it's their cumulative signal based on f).

Proposition 3.10 is in line with the basic intuition of our model: outsiders vote informatively only if they expect that there will be a sufficient number of directors who will join them. Notice that in this case it is outsiders' assessment that matters. If outsiders are unwilling to vote informatively, insiders will never vote to fire the CEO since they do not have a large enough block.

3.5.2 Optimal board composition

The more interesting and perhaps relevant question is whether there is an optimal composition of the board. We characterize optimal board composition in the following proposition.

Proposition 3.11. *(Optimal board composition).*

(*i*) *If the dissent-cost to equity loss ratio* $\frac{c}{q} \leq \Upsilon(k, n, \alpha, \gamma, \mu|b)$, *then the board should include at most one insider.*

(*ii*) *If* $\Upsilon(k, n, \alpha, \gamma, \mu|b) < \frac{c}{q} \leq \Theta(k, n, q, \alpha, \gamma, \mu, h)$, *then the board should include both insiders and outsiders.*

(*iii*) *If* $\frac{c}{q} > \Theta(k, n, q, \alpha, \gamma, \mu, h)$, *then only insiders should be on the board, where*
$\Upsilon(k, n, \alpha, \gamma, \mu|b) = \frac{p\left(\sum_{j=1}^{n} e_j \geq k \middle| b\right)\left(p(B|b)-(1-\gamma)\right)}{p\left(\sum_{j=1}^{n} e_j < k \middle| b\right)}$ *and*
$\Theta(\cdot, h) = \frac{p(f=b|b)p\left(\sum_{j=1}^{n-h} e_j \geq k-h \middle| b\right)+p(f=g|b)p\left(\sum_{j=1}^{n-h} e_j \geq k \middle| b\right)}{1-p(f=b|b)p\left(\sum_{j=1}^{n-h} e_j \geq k-h \middle| b\right)-p(f=g|b)p\left(\sum_{j=1}^{n-h} e_j \geq k \middle| b\right)}\left(p(B|b) - (1-\gamma)\right).$
All probabilities above represent beliefs of director i given his signal (in the case of insiders it's their cumulative signal based on f).

The intuition behind Proposition 3.11 is as follows. Since insiders' signals are correlated, having them on the board provides little informational advantage. In fact, their cumulative inference is the same as the inference made by a single director with an independent signal. Hence, to improve decision-making it is optimal to have as many outsiders on the board as possible as long as they are willing to reveal their private information. When the dissent-cost to equity loss ratio is low, as in (i) above, outsiders would be willing to vote informatively even if the board included only outsiders. In this case it is optimal to have at most one insider on the board (since when there's only one insider, he is equivalent to an outsider), so that the informational benefits of a greater number of independent signals provided by outsiders are realized.

Insiders, on the other hand, are able alleviate the coordination problem that outside director face because inside directors observe each other's signal. When the cost of dissent is moderate but high enough that a board consisting exclusively of outsiders will not vote informatively (as in (ii)), having insiders can improve decision-making, so that the board should consist of both insiders and outsiders. In this case, having insiders reduces the effective board size and thus mitigates the coordination problem. Once the effective board size is smaller due to the presence of insiders, outsiders will vote informatively. Notice that we explicitly emphasize the dependence of Θ on h, the number of insiders. It is only for those values of h for which the dissent-cost to equity loss ratio does not go beyond this threshold that informative voting by outsiders is feasible.[17]

Finally, if the dissent-cost to equity loss ratio is very high, as in (iii), so that outside directors are never willing to vote informatively, only insiders should be on the board. This ensures that the board will avoid the cost of dissent and will be able to fire a bad CEO with a positive probability. Outsiders will always vote to keep the CEO, so there is no informational advantage in adding them to the board in this situation. However, the disadvantage of such an insider dominated board is

[17] In general, there can be several values of h that would satisfy this requirement. On the other hand, if $\Theta(k, n, q, \alpha, \gamma, \mu, h) < \frac{p\left(\sum_{j=1}^{n} e_j \geq k \middle| b\right)\left(p(B|b) - (1-\gamma)\right)}{p\left(\sum_{j=1}^{n} e_j < k \middle| b\right)}$ for all $h = 1, \ldots, n$, then $\Theta(k, n, q, \alpha, \gamma, \mu, h)$ is irrelevant.

that, while having only insiders avoids the coordination problem, the board will have access to significantly less information in this case.[18]

3.6 Public signals

So far we have been assuming that directors were only able to obtain private information about the CEO's quality. We have shown that in this case lack of coordination may lead to inefficient decision-making. In this section we investigate what implications adding a public signal brings to our model. We argue that publicly observable information is a natural coordination mechanism. Simple intuition would suggest that having such a mechanism should improve decision-making. Somewhat surprisingly, however, we show that it is not always the case.

We assume that, in addition to their own private signals, all directors observe a public signal of CEO's quality. The public signal can be either H (high) or L (low). As in our basic model, we assume that directors are homogeneous (there's no distinction between insiders and outsiders). The informativeness of the public signal is determined by parameter ϕ:[19]

$$
\begin{aligned}
& p(H|G) = p(L|B) = \phi; \\
& p(L|G) = p(H|B) = 1 - \phi; \\
& \phi > \frac{1}{2}.
\end{aligned}
\tag{3.1}
$$

[18]Assuming that insiders have larger dissent costs compared to outsiders will not change the qualitative nature of our results. Clearly, given our assumption that insiders can credibly communicate their signals to each other, the greater dissent costs are irrelevant for communication among insiders. However, the larger dissent costs of insiders will affect outsiders' conjectures about their voting behavior. This reduces the value of insiders in terms of inducing informative voting by the board, which in turn, shrinks the range of parameter values for which insiders will be part of the optimal composition of the board.

[19]For simplicity of modeling, we use only one parameter to describe the public signal's precision.

We also assume that public and private signals are conditionally independent. That is, $p(s_{private}, S_{public}|Q) = p(s_{private}|Q)p(S_{public}|Q)$ for $s_{private} \in \{g, b\}, S_{public} \in \{H, L\}$, and $Q \in \{G, B\}$.

3.6.1 Equilibria

Due to the efficiency requirement, if the public signal is very precise compared to private signals, directors should disregard their private information and fire the CEO upon observing public signal L but retain her upon observing public signal H. This happens when $p(B|b, H) < 1 - \gamma < p(B|g, L)$. In order to keep matters interesting, we assume that the precision of the public signal does not allow directors to disregard their private signals. This is a reasonable assumption, since otherwise the board of directors brings no value into the turnover decision and it would be optimal to simply dismiss the CEO after observing negative public information (such as a fall in the stock price or in operating performance).

To capture the above ideas, in this section only we impose the following restriction:[20]

$$p(B|g, H) < p(B|g, L) < 1 - \gamma < p(B|b, H) < p(B|b, L). \tag{3.2}$$

Proposition 3.12. *(Equilibrium with a public signal). Index board members by i and let $e_i = 1$ if the i^{th} board member observes a bad private signal ($s_i = b$); $e_i = 0$ if the private signal is good ($s_i = g$). Then, in the presence of a public signal, the game described above has a unique equilibrium, so that:*

(i) *After observing public signal H:*

(a) If the dissent-cost to equity loss ratio $\frac{c}{q} \leq \Upsilon(k, n, q, \alpha, \gamma, \mu|b, H)$, each

[20] This assumption ensures that the private signal continues to be informative to board members even in the presence of a public signal (i.e., the public signal is not too precise).

director votes informatively, in particular, director i votes to retain the CEO iff $s_i = g$ and to fire the CEO iff $s_i = b$;

(b) *If $\frac{c}{q} > \Upsilon(k, n, q, \alpha, \gamma, \mu|b, H)$, each director votes to retain the CEO regardless of his private signal, where*
$\Upsilon(k, n, q, \alpha, \gamma, \mu|b, H) = \frac{p\left(\sum_{j=1}^{n} e_j \geq k \middle| b,H\right)\left(p(B|b,H)-(1-\gamma)\right)}{p\left(\sum_{j=1}^{n} e_j < k \middle| b,H\right)}$.

(ii) *After observing public signal L:*

(a) *If the dissent-cost to equity loss ratio $\frac{c}{q} \leq \Upsilon(k, n, q, \alpha, \gamma, \mu|b, L)$, each director votes informatively, in particular, director i votes to retain the CEO iff $s_i = g$ and to fire the CEO iff $s_i = b$;*

(b) *If $\frac{c}{q} > \Upsilon(k, n, q, \alpha, \gamma, \mu|b, L)$, each director votes to retain the CEO regardless of his private signal, where*
$\Upsilon(k, n, q, \alpha, \gamma, \mu|b, L) = \frac{p\left(\sum_{j=1}^{n} e_j \geq k \middle| b,L\right)\left(p(B|b,L)-(1-\gamma)\right)}{p\left(\sum_{j=1}^{n} e_j < k \middle| b,L\right)}$.

Notice that unlike in the basic model there are two thresholds in this case, depending on the type of the public signal that board members observe. This is because, after observing a public signal each director reassesses the validity of his own private information, conditional on the public information. The quality of decision-making by the board depends on the relation between $\frac{c}{q}$ and these conditional thresholds.

Proposition 3.13. *(Quality of decision making with a public signal).*

(i) *If $\frac{c}{q} \leq \Upsilon(k, n, q, \alpha, \gamma, \mu|b, H)$, directors vote informatively regardless of the public signal, i.e., the public signal is irrelevant.*

(ii) *If $\Upsilon(k, n, q, \alpha, \gamma, \mu|b, H) < \frac{c}{q} \leq \Upsilon(k, n, \alpha, \gamma, \mu|b)$, directors vote informatively only if there is no public signal or if the public signal is L, i.e., the public signal worsens the board's decision-making on average.*

(*iii*) *If* $\Upsilon(k,n,\alpha,\gamma,\mu|b) < \frac{c}{q} \leq \Upsilon(k,n,q,\alpha,\gamma,\mu|b,L)$*, directors vote informatively only when there is a public signal and it is* L*, i.e., the public signal improves the board's decision-making on average.*

(*iv*) *If* $\Upsilon(k,n,q,\alpha,\gamma,\mu|b,L) < \frac{c}{q}$*, directors never reveal their private signals and always vote to keep the CEO, i.e., the public signal is irrelevant again, where* $\Upsilon(k,n,q,\alpha,\gamma,\mu|x) \equiv \frac{p\left(\sum_{j=1}^{n} e_j \geq k-1|x\right)\left(p(B|x)-(1-\gamma)\right)}{p\left(\sum_{j=1}^{n} e_j < k-1|x\right)}$.

Part (*i*) of Proposition 3.13 states that when the dissent-cost to equity loss ratio is low, the public signal is irrelevant. Since $\frac{c}{q}$ is low, the directors would have revealed their private information even in our basic model. On the other hand, when this ratio goes up but remains below $\Upsilon(k,n,\alpha,\gamma,\mu|b)$ of our basic model, as in part (*ii*) of the above proposition, the presence of a public signal may worsen decision-making. This happens because conditional on observing a low public signal or no public signal at all, directors vote informatively. Observing a high public signal, however, may prevent board members with bad private signals from voicing their opinion, since they may assess that many other directors may vote for retaining the CEO, given the high public signal. A favorable public signal may thus make it less likely that there will be enough directors to overturn the incumbent CEO.

Part (*iii*) of Proposition 3.13 implies that when the dissent-cost to equity loss ratio is higher than $\Upsilon(k,n,\alpha,\gamma,\mu|b)$ of our basic model but is below $\Upsilon(k,n,q,\alpha,\gamma,\mu|b,H)$, the presence of a public signal is beneficial. This happens because in this case the only way for directors to vote informatively is to observe a low public signal. Directors with bad private signals would have sided with the CEO in our basic model. Upon learning negative public information about her quality, however, there is a higher probability that enough other directors will also want to oust her. This higher chance of success facilitates informative voting. When $\frac{c}{q}$ is very high, as in part (*iv*) of

Proposition 3.13, even a low public signal will not make board members reveal their private information. In this case, the public signal is irrelevant.

3.7 The dynamic model: Proactive boards and early firing decisions

We now turn to the dynamic aspect of the board's decision-making. In particular, we ask the following question: which boards are likely to fire their CEO sooner rather than later? The financial press has been long asking the question why some boards seem to keep their CEOs until the damage is so obvious that the stock price is in a free fall. This subject, however, has seen little attention in the academic literature, with Ertugrul and Krishnan (2011) being the sole empirical paper on the topic.

In this section we modify our initial game to allow two rounds of voting: one after directors observe their private signals and another after some public information about the current CEO is revealed. We also assume the board is homogeneous (there's no difference between insiders and outsiders). To make the notion of early versus late firing decision operational, we say that the CEO is fired early if the board made this decision prior to the revelation of a negative public signal. The CEO is said to be fired late if the decision came only after a negative public signal. The negative public signal we have in mind is a drop in share price (loss of shareholder value) under the current CEO.

The timeline of the extended model is as follows. At time 1 each director observes a private signal about the CEO's ability. At time 2 the board votes on whether or not to fire the CEO. If the CEO is retained, at date 3 the public signal of her quality is revealed. If a new CEO is hired, at time 3 directors observe private signals about her

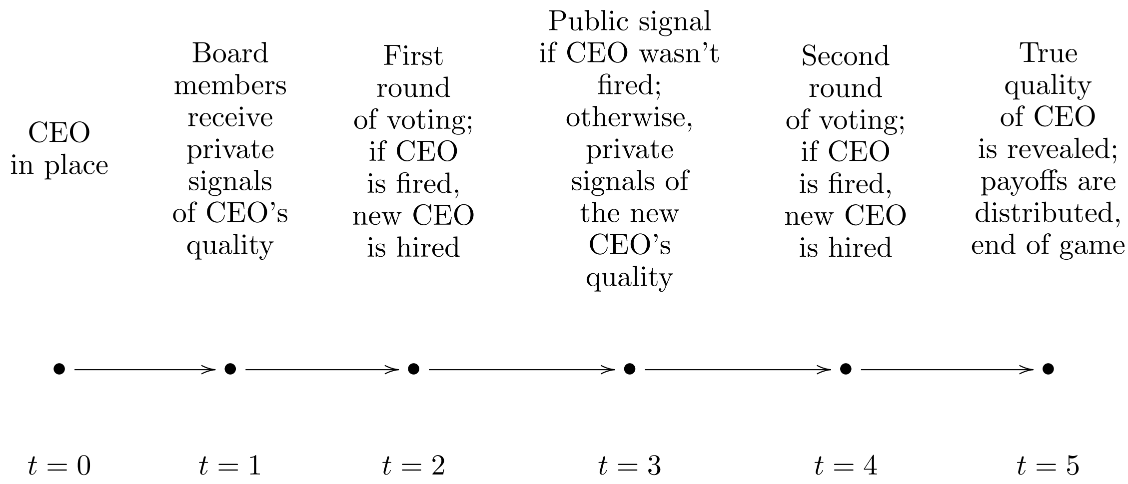

Figure 3.2: Timeline of the dynamic model.

ability, and there will be no public signal at time 3 if a new CEO is hired at time 2. At time 4 the second round of voting takes place and at time 5 the true quality of the then-incumbent CEO is revealed. Unlike in section 3.6, we assume that the public signal is highly informative, that is, $p(B|g, H) < p(B|b, H) < 1 - \gamma < p(B|g, L) < p(B|b, L)$. Given this, directors should disregard their private signals upon observing the public signal. Apart from modeling simplicity, we introduce this assumption because it represents the trade-off we intend to study in this section. One may view this public signal as the result of a series of revelations, which are compressed into a single publicly observed sufficient statistic (H for high or L for low). As in the previous section, the precision of the public signal is determined by the parameter ϕ.

In addition, we impose a penalty for keeping a bad CEO for too long. We increase the amount of potential equity loss to reflect the idea that some damage is being done to the firm during a bad CEO's tenure.[21] Specifically, if a low public signal is

[21] Without such an increase it would always be optimal to wait until the public signal is revealed, since this would allow board members to always avoid the dissent-cost without suffering any additional equity loss.

observed at time 3, each board member suffers an equity loss q: this loss is in addition to any equity loss they will suffer if it is revealed that the quality of the CEO at time 5 is bad (e.g., if the board fired the incumbent CEO but the replacement CEO was of bad quality, the total equity loss by each board member will be $2q$). Also, if a bad CEO was retained from time 0 until time 5, each board member's dissent cost rises to $2q$. The potential equity loss can be as high as $3q$ in the following scenario: a bad incumbent is retained after both rounds of voting even though the public signal revealed at time 3 was L. This never happens in equilibrium, however, since board members will disregard their private signals and vote in the direction of the more precise public signal. If a new CEO was hired at time 2 (i.e., f the incumbent CEO was fired early, before the realization of the public signal), there is no increase in the equity loss.[22] For simplicity, we also assume that $\mu = \gamma$, i.e., the board members' probability assessment of the incumbent CEO being of type G is the same as the probability assessment of a CEO from the outside pool being of type G.

Proposition 3.14. *(Firing in the dynamic game). In the two-round voting game described above, there exists a threshold,* $\Psi(\gamma, \alpha, \phi, k, n)$*, such that if* $\frac{c}{q} > \Psi(\gamma, \alpha, \phi, k, n)$*, the board never fires the incumbent CEO after the first round of voting, but waits for the second round of voting. If* $\frac{c}{q} \leq \Psi(\gamma, \alpha, \phi, k, n)$*, the board will fire the incumbent CEO after the first round of voting with a positive probability.*

The trade-off faced by board members is as follows. As in the basic model, they face an equity loss if it is revealed that they chose a bad CEO and the dissent-cost if they attempt to fire the incumbent but fail. However, in this dynamic setting, they have the additional choice of whether to fire the CEO in the first round by voting according to their private signals, or to wait until some public information is observed.

[22]If a new CEO was hired after the first round of voting, the game after time 2 is equivalent to the original one-period game.

Since we assume that the public signal is more precise than directors' private signals, waiting is beneficial. There is some residual uncertainty, however, so that even after observing a positive public signal the CEO's quality may be revealed as bad at time 5. In that case, the board members' equity loss will be significantly greater than it would have been had the board fired the CEO in the first round of voting (at time 2). The equity loss will also be higher if a low public signal is revealed, even if the board fires the CEO upon observing this public signal. If the dissent-cost to equity loss ratio is high, board members would rather risk a greater equity loss than face the higher probability of suffering dissent-costs. When this ratio is low, directors will vote according to their private information in the first round since the potential of an increased equity loss outweighs the benefits of avoiding the cost of dissent.

To summarize, as time goes by and more evidence of poor firm performance by the CEO becomes available to board members, their equity loss becomes greater; however, the coordination problem faced by board members is also reduced over time, since, once each board member becomes aware that others are also receiving more and more such evidence of the CEO's incompetence (over time), their probability assessment of a sufficient number of board members voting against the CEO goes up (and correspondingly, their probability assessment of having to incur the dissent-cost goes down). Proposition 3.14 implies that better incentivized boards (smaller $\frac{c}{q}$) are more likely to fire the CEO earlier, before significant damage is done to shareholder wealth.

3.8 Empirical and policy implications

Our model has several testable and policy implications. We highlight some of these below.

(i) Passive boards, suboptimal firing decisions, and the effects of shareholder activism: Our model develops a new rationale for why many boards are suboptimally passive (i.e., they choose to retain the existing CEO for too long). We showed that when the dissent-cost to equity loss ratio is large (for example, when the CEO has the ability to dictate the composition of the board without any uncertainty), coordination problems between board members prevent the board from firing the CEO even when collectively, the board has enough information to know that it is optimal for shareholders to fire the existing CEO and hire a new one from outside. Further, our model implies that if the board's assessment of the quality of the existing CEO relative to the pool of outside CEOs is low, then the board is more likely to fire the existing CEO and hire a new CEO from the outside.

The above has several testable predictions. First, more entrenched CEOs with greater ability to influence the payoff of board members are less likely to be fired. Second, there will be a positive relationship between a CEO being fired and the pool of outside candidates for the CEO's job and a negative relationship between industry-adjusted firm performance and the CEO being fired. Third, firms in industries where board members are able to evaluate CEOs more precisely (i.e., where board members receive more precise signals of CEO's quality) are more likely to make better CEO firing decisions. Thus, boards of firms in more homogeneous industries are likely to make better firing decisions than those in more heterogeneous industries (where it is harder to benchmark the CEO's performance relative to other firms). Further, boards in which members have, on average, larger stock ownership (larger q in our model) and are institutional or otherwise activist shareholders (lower c) are more likely to make better decisions in terms of firing the CEO.

Evidence consistent with the first prediction above is provided by Allen (1981), who documents that CEOs having greater power over directors have longer tenures.

Strong evidence supporting the second and third predictions above is provided by Parrino (1997). First, he documents a strong negative relationship between industry-adjusted firm performance and the likelihood that an outsider is appointed CEO. Second, he documents that the likelihood of forced CEO turnover and outside succession are both greater in homogeneous industries (that consist of similar firms) than in heterogeneous industries.

(ii) Optimal board size: The optimal board size in our model is determined by the trade-off between the information requirements of the board in evaluating the CEO, which favors a larger board (the larger the board is, the greater is the amount of information available to the board as a whole), and the greater coordination problems that arise with larger boards. Thus, our model predicts that, in order to maximize firm value, firms will have the smallest board size consistent with the information requirements of the board in evaluating and monitoring the performance of the CEO. Thus, our model predicts that larger, multi-divisional, and more complex firms (in the sense that the board requires more information in evaluating the CEO's performance) will have larger boards than simple firms.[23] In the latter case, our model predicts that the board members will (optimally) be better incentivized (smaller $\frac{c}{q}$) either through larger holdings in the firm's equity (larger equity loss q) or through appointing board members (like activist shareholders) who suffer from smaller dissent-costs.

Several empirical papers provide evidence consistent with the first prediction above of our model. Yermack (1996) shows that, controlling for firm size and other firm characteristics, there is an inverse relationship between market value (Tobin's Q) and the size of the board of directors. He also shows that measures of operating

[23] As we demonstrated in Proposition 3.6(a), if board members' signals are relatively precise, a larger board makes better decisions. In large and complex firms, each board member can specialize in evaluating the CEO's performance in a particular division, so that each board member will have a signal of higher precision compared to the case where each board member has to evaluate the entire firm. This, in turn, implies that larger boards are optimal for larger and more complex firms.

efficiency and profitability are negatively related over time to board size within firms, and that smaller boards are more likely to dismiss CEOs following periods of poor performance. Related evidence is provided by Kini, Kracaw, and Mian (1995), who document that board size shrinks after successful tender offers of underperforming firms. Evidence supporting the second prediction of our model above is provided by Mikkelson, Megan Partch, and Shah (1997) and Boone, Field, Karpoff, and Raheja (2007), who document that firms start out with smaller boards at IPO, with board size increasing significantly after the firm has become seasoned. The latter paper also documents that board size is related to measures of the scope and complexity of the firm's operations such as firm size, firm age, and the number of business segments the firm operates in. Finally, Coles, Daniel, and Naveen (2008) document that complex firms, which have greater advising requirements than simple firms, have larger boards with more outside directors. They also document that the relation between a firm's market value (Tobin's Q) and board size is U-shaped: Tobin's Q increases (decreases) in board size for complex (simple) firms.

(iii) Optimal board composition: Conventional wisdom among practitioners seems to be that a greater level of board independence (a larger fraction of outside directors) unambiguously increases firm performance. For example, TIAA-CREF, one of the largest pension funds in the world, has stated that it will invest only in companies that have a majority of outside directors on its board; similarly, CALPERS, another large pension fund, recommends that the CEO should be the only inside director on a firm's board. The empirical evidence, however, has been mixed: For example, Coles, Daniel, and Naveen (2008) present evidence challenging the notion that restrictions on board size and insider representation necessarily enhance firm value (see also Bhagat and Black (2001)).

Our model contributes significantly to the above debate about the optimal proportion of insiders versus outsiders on corporate boards. Our analysis suggests that there may be an optimal mix of insiders versus outsiders on corporate boards: the advantage of having insiders arises from reducing coordination problems, while the advantage of having outsiders arises from the additional information they bring to the board (at the cost of greater coordination problems). Thus, our analysis suggests that the fraction of insiders to outsiders (and whether the board should be insider or outsider dominated) depends on the dissent-cost to equity loss ratio. Consequently, our model predicts that a greater proportion of independent directors (outsiders) is not necessarily value improving in all situations: rather, it suggests that firms where the CEO has greater influence, generating higher dissent-costs for board members (as measured by the CEO's share ownership and job tenure), will be associated with a smaller fraction of independent outsiders. Evidence supporting this prediction is provided by Boone, Field, Karpoff, and Raheja (2007), who document such a relationship. On the other hand, our model predicts that more complex firms (with greater informational requirements for the board in evaluating the firm's and the CEO's performance) should have a greater proportion of outsiders; further, in such boards, the outside board members should have greater equity holdings, so as to reduce the dissent-cost to equity loss ratio. Evidence supporting the former prediction is provided by Coles, Daniel, and Naveen (2008), who document that more complex firms, with greater advising requirements than simpler firms, have larger boards with more outside directors. Evidence supporting the latter prediction that the proportion of independent outside directors will be greater in firms with lower dissent-cost to equity loss ratio is provided by Boone, Field, Karpoff, and Raheja (2007), who document that the proportion of independent outsiders on a firm's board is positively related to the equity ownership of outside directors. Finally, our model predicts that a board with a greater proportion of (appropriately incentivized) outsiders is more likely to replace a poorly performing CEO. Evidence supporting this prediction is supported by Weisbach (1988) and Borokhovich, Parrino, and Trapani (1996).

(iv) The ambiguous effect of low-precision public signals on board performance: Our model demonstrates that, contrary to common intuition, low-precision public signals may in fact worsen the board's decision-making instead of improving it (relative to the case of no public signal). Thus, short-term improvements in accounting performance by a firm may induce a board member to vote to retain the current CEO, even though he is privately convinced that the firm is pursuing a wrong long-term strategy under the current CEO. This arises from the fact that such a public signal worsens coordination problems among board members by increasing each member's fear that other members may vote to retain the CEO, thus increasing his likelihood of having to incur dissent-costs arising from voting against the CEO but failing to oust her.

(v) The dynamics of forced CEO turnover and its relation to board characteristics: Both practitioners and academics have bemoaned the tendency of boards to allow CEOs to pursue wrong-headed policies for a number of years, and fire them only after millions (if not billions) of dollars of shareholder value has been destroyed. For example, Monks and Minow (1995) document a number of cases where CEOs were able to destroy shareholder value for a number of years, and raise the following questions: "Why does it take boards so long to respond to deep-seated competitive problems? And, if one of the leading responsibilities of directors is to evaluate the performance of the CEO, why do boards wait too long for proof of managerial incompetence before making a move?" Jensen (1993) makes a similar point and calls this a failure of corporate internal control systems, and comments: "They seldom respond in the absence of a crisis."[24]

[24]Monks and Minow (1995) give the case studies of the CEOs of General Motors, Westinghouse,

The dynamic version of our model provides some answers to the above questions raised by Monks and Minow (1995), regarding why, even when board members are individually aware of the wrong-headedness of the CEO's policies, they do not immediately vote to fire the CEO. Our dynamic analysis indicates that the timing of a board's firing of a CEO who is destroying shareholder value is driven by a trade-off between the equity loss incurred by board members versus the probability of incurring a dissent-cost. On the one hand, as time goes by and more evidence of poor firm performance becomes available to board members, their equity loss becomes greater; on the other hand, the coordination problem faced by board members is also reduced over time, since, once each board member becomes aware that others are also receiving more and more such evidence (over time), their probability assessment of of having to incur a dissent-cost also goes down. Thus, our dynamic analysis has two predictions. First, CEOs who are entrenched and have greater influence over the board (e.g., as measured by longer tenure) are likely to be allowed to pursue policies that reduce share price for a longer time period before being asked to resign by the board. Second, boards that are better incentivized (smaller dissent-cost to equity loss ratio) are more likely to fire CEOs proactively, before considerable destruction of shareholder value has taken place. Evidence consistent with the latter prediction of our model is provided by Ertugrul and Krishnan (2011), who document that boards that dismiss CEOs early (i.e., in the absence of significant negative prior stock returns) are those that are characterized by higher equity ownership by board members, higher

American Express, IBM, Eastman Kodak, Scott Paper, and Borden, who were pressured to resign in the face of their companies' long-term underperformance. They comment that, while the above moves by their companies' boards were heralded in the media as breakthroughs in boardroom activism, in all these instances the board took the necessary drastic action years too late. Jensen (1993) points out that "The [...] GM board revolt [...] which resulted in the firing of CEO Robert Stempel exemplifies the failure, not the success, of GM's governance system. General Motors, one of the world's high-cost producers in a market with substantial excess capacity, avoided making major changes in its strategy for over a decade. The revolt came too late: the board acted to remove the CEO only in 1992, after the company had reported losses of $6.5 billion in 1990 and 1991."

board equity compensation as a fraction of total compensation, higher institutional ownership, and better firm corporate governance characteristics.[25]

(vi) Policy implications for corporate governance and board reform: Our model suggests several ways in which corporate governance, especially corporate boards, can be reformed to improve firm performance. Our analysis suggests that reforms that reduce the dissent-costs of board members or increase their equity loss will significantly increase the likelihood of boards firing incompetent CEOs. One proposal that would reduce dissent-costs would be to grant security of tenure (i.e., guarantee their presence on the board for some length of time) for current board members, thus reducing their dissent-costs in the event they vote against the current CEO but end up being unable to oust him or her. Another reform proposal to reduce dissent-costs would make it easier for dissidents to acquire a seat on the board: due to the cost of proxy fights, few dissidents bother to make board challenges.[26] One such proposal currently being considered by the SEC would let owners of at least 1% of a company with an equity value of $700 million or more (3% for smaller companies) include information about their board nominees in corporate proxy materials. Another such proposal would reimburse successful dissident board candidates for all their campaign expenses (and partially reimburse candidates who obtained at least 40% of the votes cast).[27] Another such proposal that would reduce dissent-costs would be to reduce

[25] Ertugrul and Krishnan (2011) document that firms that dismiss CEOs early do not experience poor operating performance after the CEO dismissal relative to a control sample, suggesting that such dismissals are not cases of value-reducing mistakes by the board.

[26] It can cost hundreds of thousands of dollars or more for an outside contender to run against a director. For example, a campaign for a board seat may involve repeated mailings to every investor; directors endorsed by the firm management, on the other hand, can make use of the company coffers to finance their candidacies. Thus, RiskMetrics, a proxy-advisory firm, points out that there were just 75 shareholder contests in 2009 (as of October of that year); challengers won one seat each in 58 of those fights.

[27] Such a proposal was adopted by HealthSouth Corporation in October 2009. For details of this and other proposals being considered by the SEC, see Joann S. Lublin, *Reimbursements Aim for a Fairer Proxy Fight.* The Wall Street Journal, page A22, October 27, 2009.

the involvement of CEOs in the selection of board members and have boards choose directors through nominating committees composed only of independent members of the board (see, e.g., The Working Group on Corporate Governance (1991)).[28]

A second set of board reform proposals suggested by our analysis would increase the equity loss of board members. One such proposal would be to increase the proportion of board members' compensation that depends directly on the firm's long-run share value. Another proposal (perhaps harder to implement) would require board members to invest a significant amount of their own wealth in the firm's equity.

3.9 Conclusion

In this paper, we have developed a theory of corporate boards and their role in forcing CEO turnover. We considered a firm with an incumbent CEO of uncertain management ability and a board consisting of a number of directors whose role is to evaluate the CEO and fire her if a better replacement can be found. In our setting, each board member receives an independent private signal about the CEO's ability, after which board members vote on firing the CEO (or not). If the CEO is fired, the board hires a new CEO from the pool of candidates available. The true ability of the firm's CEO is revealed in the long run; the firm's long-run share price is determined by this ability. Each board member owns some equity in the firm, and thus prefers to fire a CEO of poor ability. However, if a board member votes to fire the incumbent CEO but the number of other board members also voting to fire her is not enough to successfully oust her, the CEO can impose significant costs of dissent

[28]Shivdasani and Yermack (1999) document that when the CEO is involved in the selection of a new director, firms appoint fewer independent outside directors and more "grey" outsiders. They define a CEO as "involved" in such selection (*i*) if the board has a separate nominating committee and the CEO serves as a member or (*ii*) if such a committee does not exist and directors are selected by the entire board including the CEO.

on him. In this setting, we show that the board faces a coordination problem, leading it to retain an incompetent CEO even when a majority of board members' private signals indicate that she is of poor quality. We solved for the optimal board size, and show that this depends on various board and firm characteristics: one size does not fit all firms. We developed extensions to our basic model to analyze the optimal composition of the board between insiders and outsiders and the effect of board members observing imprecise public signals (such as the firm's short-term operating performance) in addition to their private signals on the effectiveness of board decision-making. Finally, we developed a dynamic extension to our basic model to analyze why many boards do not fire CEOs even when they preside over a significant, publicly observable, reduction in shareholder wealth over a long period of time, and used this dynamic model to distinguish between the characteristics of such boards from those that fire bad CEOs proactively, before significant wealth reductions take place.

3.10 Appendix: Proofs of Propositions

Proof of Proposition 3.1

When $\gamma = \mu$, (3.6) can be rewritten as

$$\frac{(1-\mu)\alpha^r(1-\alpha)^{n-r}}{(1-\mu)\alpha^r(1-\alpha)^{n-r} + \mu(1-\alpha)^r\alpha^{n-r}} > 1-\mu. \tag{3.10.1}$$

With a little bit of algebra, (3.10.1) transforms into

$$\begin{aligned} &\alpha^r(1-\alpha)^{n-r} > (1-\alpha)^r\alpha^{n-r} \;\Leftrightarrow \\ &2r > n \text{ (since } \alpha > \tfrac{1}{2}) \;. \end{aligned} \tag{3.10.2}$$

We assumed that n is odd; r is an integer, hence, the lowest r that satisfies (3.10.1) is $r^* = \frac{n+1}{2}$.

Proof of Proposition 3.2

With no dissent-costs, the directors' objective is no longer given by (3.1), but rather by

$$\pi_i^{benchmark}(v_i; s_i, v_{-i}, k, n, q, \alpha, \gamma, \mu) = -I\big(\sum_{j=1}^{n} v_j \geq k\big)q(1-\gamma) - I\big(\sum_{j=1}^{n} v_j < k\big)qp(B|s_i). \tag{3.10.3}$$

We say that a director is pivotal if his vote determines the outcome of the vote. First consider the case when the director is not pivotal. Since there are no costs of dissent, his vote does not alter his payoff. He may as well vote informatively. Now consider the case when the director is pivotal. It only happens when exactly $k-1$ other directors vote to fire the CEO. Since k is the optimal voting rule, if the pivotal director's signal is g, then his inference is $p(B|g^{pivotal}) < 1-\gamma$. If, on the other hand, this director's signal is b, then his inference is $p(B|b^{pivotal}) > 1-\gamma$. Thus, $-qp(B|g^{pivotal}) > -q(1-\gamma)$ and $-qp(B|b^{pivotal}) < -q(1-\gamma)$, and it follows that

$$\pi_i^{benchmark}(v_i = 0; g^{pivotal}, \cdot) > \pi_i^{benchmark}(v_i = 1; g^{pivotal}, \cdot), \tag{3.10.4}$$

and

$$\pi_i^{benchmark}(v_i = 0; b^{pivotal}, \cdot) < \pi_i^{benchmark}(v_i = 1; b^{pivotal}, \cdot). \tag{3.10.5}$$

Hence, the pivotal board member is better off voting informatively. Thus, informative voting is an equilibrium. This equilibrium is efficient since its outcome coincides with what a social planner, with access to all directors' signals, would choose.

Proof of Proposition 3.3

Since n is odd, let $n = 2m+1$. Let $P(k;n)$ be the probability that there are at least k votes out of n for the correct outcome (to fire the CEO if she is of poor quality and to keep the CEO if she is of high quality). We will show that $P(m+1; 2m+1) > P(m; 2m-1)$. That is, it is more likely that the majority of votes will be in the right direction when the number of directors grows from $2m-1$ to $2m+1$ (we increase the number of directors by 2 to keep it odd).

There are three possible combinations of the signals observed by the added directors (after $2m-1$ initial signals):

1. their signal is split (one observes a bad signal and the other observes a good signal), which happens with probability $2\alpha(1-\alpha)$;
2. both of them observe the right signal, with probability α^2;
3. both of them observe the wrong signal, with probability $(1-\alpha)^2$.

Notice that

$$\begin{aligned} P(m+1, 2m+1) = 2\alpha(1-\alpha)P(m; 2m-1) + \alpha^2 P(m-1; 2m-1) \\ + (1-\alpha)^2 P(m+1; 2m-1). \end{aligned} \tag{3.10.6}$$

(3.10.6) follows by the following logic. The left-hand side is the probability that at least $m+1$ directors observe the correct signal. If the signal of the new board members is split, which happens with probability $2\alpha(1-\alpha)$, then to have at least $m+1$ correct signals it must be the case that there were at least m correct signals among the initial $2m-1$ directors, which happens with probability $P(m, 2m-1)$. If both of the added board members observe correct signals, $m-1$ correct signals among the initial $2m-1$ directors suffices to bring the total to $m+1$. If both of the added board members observe wrong signals, there must already be $m+1$ correct signals.

Observe that

$$\begin{aligned} P(m-1, 2m-1) &= P(m, 2m-1) + \binom{2m-1}{m-1}\alpha^{m-1}(1-\alpha)^m, \\ P(n, k+1) &= P(m, 2m-1) - \binom{2m-1}{m}\alpha^m(1-\alpha)^{m-1}. \end{aligned} \tag{3.10.7}$$

Hence,

$$\begin{aligned} &P(m+1; 2m+1) - P(m; 2m-1) = \\ &\alpha^2\binom{2m-1}{m-1}\alpha^{m-1}(1-\alpha)^m - (1-\alpha)^2\binom{2m-1}{m}\alpha^m(1-\alpha)^{m-1} = \\ &\binom{2m-1}{m-1}\alpha^n(1-\alpha)^n(2\alpha-1) > 0, \end{aligned} \tag{3.10.8}$$

since $\binom{2m-1}{m-1} = \binom{2m-1}{m}$ and $\alpha > \frac{1}{2}$.

Proof of Proposition 3.4

There are four possible symmetric pure voting strategies given director i's signal:

1. always vote to fire the CEO, $\{v_i = 1|s_i = g; v_i = 1|s_i = b\}$;

2. always vote to keep the CEO, $\{v_i = 0|s_i = g; v_i = 0|s_i = b\}$;

3. vote according to the signal, $\{v_i = 0|s_i = g; v_i = 1|s_i = b\}$;

4. vote in the opposite direction to the signal, $\{v_i = 1|s_i = g; v_i = 0|s_i = b\}$.

It will not suffice to consider only pivotal board members since even when a director isn't pivotal, his vote influences his payoff via the dissent-cost. In the following, we characterize equilibria in terms of board members' signals rather than their votes. We denote $e_i = 1$ if $s_i = b$ and $e_i = 0$ if $s_i = g$. We can now rewrite (3.1) to obtain directors' expected payoffs from each of the above strategies, conditional on their private signals and assuming that the other directors are playing the same symmetric equilibrium. The payoffs to a director with $s_i = g$ from strategies 1, 2, 3, and 4, respectively, are given by

$$-q(1-\gamma), \tag{3.10.9}$$

$$-qp(B|g), \tag{3.10.10}$$

$$-qp(B|g)p\Big(\sum_{j=1}^{n} e_j < k\big|g\Big) - q(1-\gamma)p\Big(\sum_{j=1}^{n} e_j \geq k\big|g\Big), \text{ and} \tag{3.10.11}$$

$$-\big(c + qp(B|g)\big)p\Big(\sum_{j=1}^{n}(1-e_j) < k\big|g\Big) - q(1-\gamma)p\Big(\sum_{j=1}^{n}(1-e_j) \geq k\big|g\Big). \tag{3.10.12}$$

The payoffs to a director with $s_i = b$ from strategies 1, 2, 3, and 4, respectively are given by

$$-q(1-\gamma), \tag{3.10.13}$$

$$-qp(B|b), \tag{3.10.14}$$

$$-\big(c + qp(B|b)\big)p\Big(\sum_{j=1}^{n} e_j < k\big|b\Big) - q(1-\gamma)p\Big(\sum_{j=1}^{n} e_j \geq k\big|b\Big), \text{ and} \tag{3.10.15}$$

$$-qp(B|b)p\Big(\sum_{j=1}^{n}(1-e_j) < k\big|b\Big) - q(1-\gamma)p\Big(\sum_{j=1}^{n}(1-e_j) \geq k\big|b\Big). \tag{3.10.16}$$

We assumed that $p(B|g) < 1-\gamma < p(B|b)$. Hence, $-qp(B|g) > -q(1-\gamma) > -qp(B|b)$ and (3.10.10) $>$ (3.10.9) but (3.10.13) $>$ (3.10.14). This implies that directors who observe signal g prefer to retain the CEO while directors who observe b want to replace her.

Strategy 4 cannot survive in equilibrium. We prove by contradiction. Assume that voting in the opposite direction to one's signal and believing that everyone else follows this strategy is an equilibrium. Consider directors who observe g. They believe that all directors with signal b vote to retain the CEO. Hence, if directors with good signals vote in accordance with strategy 4, their payoff is given by (3.10.12). If, however, they vote to retain the current CEO, their payoff is given by (3.10.10). Since (3.10.10) $>$ (3.10.12), directors who observe g always deviate, yielding a contradiction. Among the three remaining strategies, strategy 1 can also be eliminated, as follows. Consider board members with signal g. If directors with this signal deviate from strategy 1, the lowest payoff they can attain is (3.10.11), which is greater than (3.10.9). Hence, it is profitable to deviate.

Only two possible equilibrium strategies remain: strategy 2 and strategy 3. In both of these, directors who observe g vote in accordance with their signal. Thus, it is directors with a bad signal of the CEO's quality who determine the equilibrium. They face the trade-off between (3.10.14) and (3.10.15) and vote according to their signal (informatively) if and only if (3.10.15) $\geq$ (3.10.14). Hence, a truth-telling equilibrium exists if and only if

$$-qp(B|b) \leq -\big(c + qp(B|b)\big)p\Big(\sum_{j=1}^{n} e_j < k\big|b\Big) - q(1-\gamma)p\Big(\sum_{j=1}^{n} e_j \geq k\big|b\Big). \tag{3.10.17}$$

Otherwise, directors who observe b vote to retain the CEO. With some algebra, we can transform (3.10.17) into

$$\frac{c}{q} \leq \frac{p\big(\sum_{j=1}^{n} e_j \geq k\big|b\big)\big(p(B|b) - (1-\gamma)\big)}{p\big(\sum_{j=1}^{n} e_j < k\big|b\big)}. \tag{3.10.18}$$

Proof of Proposition 3.5

Compare equilibria in Proposition 3.2 and Proposition 3.4. It is immediately clear that they coincide except when $\frac{c}{q} > \frac{p\left(\sum_{j=1}^{n} e_j \geq k\middle|b\right)\left(p(B|b)-(1-\gamma)\right)}{p\left(\sum_{j=1}^{n} e_j < k\middle|b\right)}$. In that case, had there been no dissent-cost, the CEO would have been fired with a positive probability. In the presence of dissent-costs the board will never fire the CEO. Since the equilibrium in Proposition 3.2 produces the same outcome as a social planner with access to all directors' signals, any deviation from this equilibrium is suboptimal.

Proof of Proposition 3.6

Let $\omega(n,k) \equiv p\big(\sum_{j=1}^{n} e_j \geq k\big|b\big)$. It suffices to show that $\omega(n+2,k+1) < \omega(n,k)$ when $\mu > \alpha\frac{2n\alpha-(n+1)}{2n\alpha-(n+1)+2(1-\alpha)}$.

Let $\lambda \equiv p(e_j = 1|b) = \frac{\mu-2\mu\alpha+\alpha^2}{\mu-2\mu\alpha+\alpha}$. Then,

$$\omega(n,k) \equiv p\Big(\sum_{j=1}^{n} e_j \geq k\big|b\Big) = p\Big(\sum_{j=1}^{n-1} e_j \geq k-1\big|b\Big) = \sum_{m=k-1}^{n-1} \binom{n-1}{m} \lambda^m (1-\lambda)^{n-1-m}. \tag{3.10.19}$$

The second step in (3.10.19) follows from the fact that once a board member observes b, he only needs to consider consider the probability that $k-1$ other directors obtained a bad signal.

Notice that:

$$\omega(n+2,k+1) = 2\lambda(1-\lambda)\omega(n,k) + \lambda^2\omega(n,k-1) + \lambda(1-\lambda)\omega(n,k+1). \tag{3.10.20}$$

Further observe that using (3.10.19), the following probabilities can be rewritten:

$$\begin{aligned}\omega(n,k-1) &= \omega(n,k) + \binom{n-1}{k-1}\lambda^{k-1}(1-\lambda)^{n-k},\\ \omega(n,k+1) &= \omega(n,k) - \binom{n-1}{k}\lambda^{k}(1-\lambda)^{n-k-1}.\end{aligned} \tag{3.10.21}$$

Combining (3.10.20) with (3.10.21) yields

$$\begin{aligned}&\omega(n+2,k+1)-\omega(n,k) =\\ &\lambda^2\binom{n-1}{k-1}\lambda^{k-1}(1-\lambda)^{n-k} - (1-\lambda)^2\binom{n-1}{k}\lambda^{k}(1-\lambda)^{n-k-1} =\\ &\frac{(n-1)!}{(k-1)!(n-k-1)!}\lambda^k(1-\lambda)^{n-k}\Big(\frac{\lambda}{n-k}-\frac{1-\lambda}{k}\Big).\end{aligned} \tag{3.10.22}$$

The sign of (3.10.22) is determined by $\big(\frac{\lambda}{n-k}-\frac{1-\lambda}{k}\big)$. Noting that $k=\frac{n-1}{2}$, we have:

$$\frac{\lambda}{n-k}-\frac{1-\lambda}{k}=\frac{2\lambda}{n+1}-\frac{2-2\lambda}{n-1}. \tag{3.10.23}$$

Hence,

$$\begin{aligned}\omega(n+2,k+1) < \omega(n,k) &\Leftrightarrow\\ \frac{\lambda}{n+1}<\frac{1-\lambda}{n-1} &\Leftrightarrow\\ \mu > \alpha\frac{2n\alpha-(n+1)}{2n\alpha-(n+1)+2(1-\alpha)}&.\end{aligned} \tag{3.10.24}$$

In deriving (3.10.24) we used the fact that $2n\alpha-(n+1)+2(1-\alpha)>0$ given $\alpha>\frac{1}{2}>\frac{n-1}{2n-1}$.

Proof of Proposition 3.7

Let n^* be the maximum n so that (3.10.18) holds. We will show that n^* dominates any $n>n^*$ and any $n<n^*$. First, consider some $n'>n^*$. First, notice that if $\alpha\le\mu$, then $\mu>\alpha z(\alpha,n)$ for any n since $z(\alpha,n)<1$. Hence, it follows from Proposition 3.6 that a board consisting of n' members will always retain the incumbent, regardless of directors' signals, while at n^* a bad CEO is fired with a positive probability. Hence, n^* dominates n'. Now, consider some $n''<n^*$. The board votes informatively. Hence, the outcome of the vote coincides with the social planner's choice. By Proposition 3.3, however, n^* produces more efficient outcomes. Hence, n^* dominates n'' as well.

Proof of Proposition 3.8

Part (i). We want to show that $\Upsilon(k,n,\alpha,\gamma,\mu|b)$ is decreasing in μ. It is easy to see that:

$$\begin{aligned}\frac{\partial p(B|b)}{\partial \mu} &= \frac{-\alpha(\alpha+\mu-2\alpha\mu)-(\alpha-\alpha\mu)(1-2\alpha)}{(\mu-2\mu\alpha+\alpha)^2} \\ &= \frac{\alpha(\alpha-1)}{(\mu-2\mu\alpha+\alpha)^2} < 0, \text{ since } \alpha < 1.\end{aligned} \tag{3.10.25}$$

Hence, the difference $p(B|b)-(1-\gamma)$ is decreasing in μ. It remains to show that $p\big(\sum_{j=1}^{n} e_j \geq k \big| b\big)$ is also decreasing in μ.

As before, let $\lambda \equiv p(e_j=1|b) = \frac{\mu-2\mu\alpha+\alpha^2}{\mu-2\mu\alpha+\alpha}$. Then,

$$\begin{aligned}\frac{\partial \lambda}{\partial \mu} &= \frac{(1-2\alpha)(\mu-2\mu\alpha+\alpha)-(\mu-2\mu\alpha+\alpha^2)(1-2\alpha)}{(\mu-2\mu\alpha+\alpha)^2} \\ &= \frac{(1-2\alpha)\alpha(1-\alpha)}{(\mu-2\mu\alpha+\alpha)^2} < 0, \text{ since } \frac{1}{2} < \alpha < 1.\end{aligned} \tag{3.10.26}$$

To complete the proof, notice that $p\big(\sum_{j=1}^{n} e_j \geq k \big| b\big)$ is 1 minus the cumulative distribution function of a Binomial distribution with parameter λ. It is therefore increasing in λ.

Part (*ii*). We want to show that when $\mu = \frac{1}{2}$, $\Upsilon(k, n, \alpha, \gamma, \mu|b)$ is increasing in α. First,

$$\begin{aligned}\frac{\partial p(B|b)}{\partial \alpha} &= \frac{(1-\mu)\big((1-\mu)\alpha+\mu(1-\alpha)\big)-(1-\mu)\alpha(1-2\mu)}{\big((1-\mu)\alpha+\mu(1-\alpha)\big)^2} \\ &= \frac{(1-\mu)\mu}{\big((1-\mu)\alpha+\mu(1-\alpha)\big)^2} > 0, \text{ since } \mu < 1.\end{aligned} \tag{3.10.27}$$

Hence, the difference $p(B|b)-(1-\gamma)$ is increasing in α. It remains to show that $p\big(\sum_{j=1}^{n} e_j \geq k \big| b\big)$ is also increasing in α when $\mu = \frac{1}{2}$. Notice that when $\mu = \frac{1}{2}$, $\lambda = \frac{\mu-2\mu\alpha+\alpha^2}{\mu-2\mu\alpha+\alpha} = 1+2\alpha^2-2\alpha$. Thus,

$$\frac{\partial \lambda}{\partial \alpha} = 4\alpha - 2 > 0 \text{ since } \frac{1}{2} < \alpha < 1. \tag{3.10.28}$$

To complete the proof, notice that $p\big(\sum_{j=1}^{n} e_j \geq k \big| b\big)$ is 1 minus the cumulative distribution function of a Binomial distribution with parameter λ. It is therefore increasing in λ.

Part (*iii*). Observe directly that $\frac{p\left(\sum_{j=1}^{n} e_j \geq k \middle| b\right)\left(p(B|b)-(1-\gamma)\right)}{p\left(\sum_{j=1}^{n} e_j < k \middle| b\right)}$ is increasing in γ.

Proof of Proposition 3.9

Insiders' vote determines the outcome. Hence, outsiders are irrelevant to the board's decision-making and always vote to retain the CEO in order to avoid incurring the cost of dissent. Since insiders always avoid the cost of dissent by voting as a block, they solve the following problem:

$$\max_{v_i} \pi_i^{insider}(v_i; f) = -qp(B|f)(1-v_i) - q(1-\gamma)v_i. \tag{3.10.29}$$

Hence, they vote to fire the CEO ($v_i = 1$) if and only if $-qp(B|f) < -q(1-\gamma)$, or, equivalently,

$$p(B|f) > 1-\gamma. \tag{3.10.30}$$

The required result follows since $(B|b) > 1-\gamma > p(B|g)$ by assumption.

Proof of Proposition 3.10

First consider insiders. We only need to consider the case when insiders collectively observe a bad signal, since if the signal is good they never vote to fire the CEO. Insiders with a bad signal vote informatively if and only if

$$cp\Big(\sum_{j=1}^{n-h} e_j \le k-h \Big| f=b\Big) < qp\Big(\sum_{j=1}^{n-h} e_j \ge k-h \Big| f=b\Big)\big(p(B|f=b)-(1-\gamma)\big). \tag{3.10.31}$$

(3.10.31) reflect the fact that there must be at least $k-h$ votes against the CEO cast by outsiders (whose number is $n-h$). (3.10.31) can be rewritten as

$$\frac{c}{q} \le \frac{p\big(\sum_{j=1}^{n-h} e_j \ge k-h \big| f=b\big)\big(p(B|f=b)-(1-\gamma)\big)}{p\big(\sum_{j=1}^{n-h} e_j < k-h \big| f=b\big)}. \tag{3.10.32}$$

Consider outsiders. Again, it is outsiders with bad private signals who determine the equilibrium since after observing a good signal they always vote to retain the CEO. Outsiders with bad signals must consider two scenarios: when insiders vote to fire the CEO and when insiders vote to keep the CEO. If insiders vote to fire the CEO, there needs to be $k-h$ outsiders to oust her. If insiders vote to retain the CEO, the number of outsiders necessary to oust her is k. It follows that outsiders vote informatively when

$$\frac{c}{q} \le \frac{p(f=b|b)p\big(\sum_{j=1}^{n-h} e_j \ge k-h \big| b\big) + p(f=g|b)p\big(\sum_{j=1}^{n-h} e_j \ge k \big| b\big)}{1-p(f=b|b)p\big(\sum_{j=1}^{n-h} e_j \ge k-h \big| b\big) - p(f=g|b)p\big(\sum_{j=1}^{n-h} e_j \ge k \big| b\big)} \times \\ \times \big(p(B|b)-(1-\gamma)\big). \tag{3.10.33}$$

We will show that (3.10.33) is the binding constraint. Notice that since insiders' cumulative signal is equivalent to having a single independent signal, $p(B|f=b) = p(B|b)$, $p\big(\sum_{j=1}^{n-h} e_j \ge k-h \big| f=b\big) = p\big(\sum_{j=1}^{n-h} e_j \ge k-h \big| b\big)$, $p(f=b|b) = \lambda$, $p(f=g|b) = 1-\lambda$, where $\lambda \equiv p(e_j = 1|b) = \frac{\mu - 2\mu\alpha + \alpha^2}{\mu - 2\mu\alpha + \alpha}$ as before.

(3.10.32) can now be rewritten as

$$\frac{c}{q} \le \frac{p\big(\sum_{j=1}^{n-h} e_j \ge k-h \big| b\big)}{1-p\big(\sum_{j=1}^{n-h} e_j \ge k-h \big| b\big)}\big(p(B|b)-(1-\gamma)\big). \tag{3.10.34}$$

(3.10.33) can be rewritten as

$$\frac{c}{q} \leq \frac{\lambda p\big(\sum_{j=1}^{n-h} e_j \geq k-h\big|b\big) + (1-\lambda)p\big(\sum_{j=1}^{n-h} e_j \geq k\big|b\big)}{1-\lambda p\big(\sum_{j=1}^{n-h} e_j \geq k-h\big|b\big) - (1-\lambda)p\big(\sum_{j=1}^{n-h} e_j \geq k\big|b\big)} \times \\ \times \big(p(B|b) - (1-\gamma)\big). \qquad (3.10.35)$$

Now notice that

$$p\big(\sum_{j=1}^{n-h} e_j \geq k-h\big|b\big) > \lambda p\big(\sum_{j=1}^{n-h} e_j \geq k-h\big|b\big) + (1-\lambda)p\big(\sum_{j=1}^{n-h} e_j \geq k\big|b\big). \qquad (3.10.36)$$

(3.10.36) follows because $\lambda < 1$ and $p\big(\sum_{j=1}^{n-h} e_j \geq k-h\big|b\big) > p\big(\sum_{j=1}^{n-h} e_j \geq k\big|b\big)$. Hence, whenever $\frac{c}{q}$ satisfies (3.10.33) it also satisfies (3.10.32). If $\frac{c}{q}$ satisfies (3.10.32) but does not satisfy (3.10.33), then outsiders always vote to retain the CEO. Since insiders don't have enough votes to overturn that decision, they also vote to retain the CEO.

Proof of Proposition 3.11

Part (i). The optimal board is the one with the greatest likelihood of firing a bad CEO. By Proposition 3.3, the larger the number of independent signals, the higher this likelihood is. Notice that when $\frac{c}{q} \leq \Upsilon(k, n, \alpha, \gamma, \mu|b)$, a board consisting exclusively of outsiders or a board with at most one insider will vote informatively. Such a board reveals the maximum possible number of independent signals about the CEO's ability, n. Hence, this board composition is optimal when $\frac{c}{q} \leq \Upsilon(k, n, \alpha, \gamma, \mu|b)$.

Part (ii). When $\frac{c}{q} > \Upsilon(k, n, \alpha, \gamma, \mu|b)$, a board with at most one insider will always vote to retain the CEO. Such a board obtains n independent signals about the CEO's quality, but none of these signals are reflected in the voting by the board, so that a bad CEO is fired with zero probability. If $\frac{c}{q} \leq \Theta(k, n, q, \alpha, \gamma, \mu, h)$ for some $h = 2, \ldots, n-1$, then a board with more than one insider will vote informatively by Proposition 3.10. Hence, $n-h+1$ independent signals will be reflected in the voting by the board, so that it will fire a bad CEO with a positive probability.

Part (iii). When $\frac{c}{q} > \Theta(k, n, q, \alpha, \gamma, \mu, h) > \Upsilon(k, n, \alpha, \gamma, \mu|b)$, for all $h = 2, \ldots, n-1$, then by the proof of Proposition 3.10, outsiders always vote to retain the CEO. If insiders do not dominate the board, they also vote to retain the CEO (by Proposition 3.10). If they dominate the board, they vote in accordance with their cumulative inference and a bad CEO is fired with a positive probability. Further, since outside directors always vote to retain the CEO, there is no informational advantage of adding them to the board. Hence, it is optimal to have only inside board members.

Proof of Proposition 3.12

Conditional on observing public signal H, the game with public signals is equivalent to our basic game with dissent costs, with $p(B|g, H)$ replacing $p(B|g)$ and $p(B|b, H)$ replacing $p(B|b)$. Hence, by Proposition 3.4 the board either always keeps the CEO or each director votes according to his private information. The latter outcome is possible if and only if

$$\frac{c}{q} \leq \frac{p\big(\sum_{j=1}^{n} e_j \geq k-1|b,H\big)\big(p(B|b,H)-(1-\gamma)\big)}{p\big(\sum_{j=1}^{n} e_j < k-1|b,H\big)}. \tag{3.10.37}$$

Analogously, conditional on observing L, directors will vote truthfully and reveal their private valuations if and only if

$$\frac{c}{q} \leq \frac{p\big(\sum_{j=1}^{n} e_j \geq k-1|b,L\big)\big(p(B|b,L)-(1-\gamma)\big)}{p\big(\sum_{j=1}^{n} e_j < k-1|b,L\big)}. \tag{3.10.38}$$

Proof of Proposition 3.13

First we calculate updated probabilities conditional on observing a particular public signal.

$$\begin{aligned} p(B|g,H) &= \frac{p(B)p(g,H|B)}{p(B)p(g,H|B)+p(G)p(g,H|G)} \\ &= \frac{p(B)p(g|B)p(H|B)}{p(B)p(g|B)p(H|B)+p(G)p(g|G)p(H|G)} \\ &= \frac{(1-\mu)(1-\alpha)(1-\phi)}{(1-\mu)(1-\alpha)(1-\phi)+\mu\alpha\phi}, \end{aligned} \tag{3.10.39}$$

where the second line follows from the assumption of conditional independence between public and private signals. It is easy to see that $p(B|g,H) < p(B|g)$:

$$\begin{aligned} &\frac{(1-\mu)(1-\alpha)(1-\phi)}{(1-\mu)(1-\alpha)(1-\phi)+\mu\alpha\phi} < \frac{(1-\mu)(1-\alpha)}{(1-\mu)(1-\alpha)+\mu\alpha} \Leftrightarrow \\ &\frac{\mu\alpha}{(1-\mu)(1-\alpha)} < \frac{\mu\alpha\phi}{(1-\mu)(1-\alpha)(1-\phi)} \Leftrightarrow \\ &\phi > \frac{1}{2}. \end{aligned} \tag{3.10.40}$$

Similarly,

$$p(B|g,L) = \frac{(1-\mu)(1-\alpha)\phi}{(1-\mu)(1-\alpha)\phi+\mu\alpha(1-\phi)} > p(B|g), \tag{3.10.41}$$

$$p(B|b,H) = \frac{(1-\mu)\alpha(1-\phi)}{(1-\mu)\alpha(1-\phi)+\mu(1-\alpha)\phi} < p(B|b), \tag{3.10.42}$$

$$p(B|b,L) = \frac{(1-\mu)\alpha\phi}{(1-\mu)\alpha\phi+\mu(1-\alpha)(1-\phi)} > p(B|b). \tag{3.10.43}$$

The probability of director j observing private signal b, given director i's signal, can be rewritten as

$$p(s_j = b|s_i) = (1-\alpha) + (2\alpha-1)p(B|s_i). \tag{3.10.44}$$

(3.10.44) is increasing in $p(B|s_i)$. We showed above that $p(B|b,H) < p(B|b) < p(B|b,L)$. Hence, $p(s_j = b|b,H) < p(s_j = b|b) < p(s_j = b|b,L)$. It follows from the properties of the Binomial distribution that $p(\sum_{j=1}^{n} e_j \geq k-1|b,H) < p(\sum_{j=1}^{n} e_j \geq k-1|b) < p(\sum_{j=1}^{n} e_j \geq k-1|b,L)$ and $p(\sum_{j=1}^{n} e_j < k|b,H) > p(\sum_{j=1}^{n} e_j < k|b) > p(\sum_{j=1}^{n} e_j < k|b,L)$. Therefore,

$$\frac{p\big(\sum_{j=1}^{n} e_j \geq k-1\big|b,H\big)\big(p(B|b,H)-(1-\gamma)\big)}{p\big(\sum_{j=1}^{n} e_j < k-1\big|b,H\big)} < \frac{p\big(\sum_{j=1}^{n} e_j \geq k-1\big|b\big)\big(p(B|b)-(1-\gamma)\big)}{p\big(\sum_{j=1}^{n} e_j < k-1\big|b\big)} < \frac{p\big(\sum_{j=1}^{n} e_j \geq k-1\big|b,L\big)\big(p(B|b,L)-(1-\gamma)\big)}{p\big(\sum_{j=1}^{n} e_j < k-1\big|b,L\big)}. \tag{3.10.45}$$

(3.10.45) is equivalent to $\Upsilon(k,n,q,\alpha,\gamma,\mu|b,H) < \Upsilon(k,n,\alpha,\gamma,\mu|b) < \Upsilon(k,n,q,\alpha,\gamma,\mu|b,L)$. The rest of this proposition follows from the equilibrium described in Proposition 3.12 and the relation between $\Upsilon(k,n,q,\alpha,\gamma,\mu|b,H)$, $\Upsilon(k,n,\alpha,\gamma,\mu|b)$ and $\Upsilon(k,n,q,\alpha,\gamma,\mu|b,L)$ that we have just derived.

Proof of Proposition 3.14

We solve the problem within the rational expectations framework by starting with the outcome of the second round of voting. It is easy to see that if $\frac{c}{q} > \frac{p\big(\sum_{j=1}^{n} e_j \geq k-1\big|b\big)\big(p(B|b)-(1-\gamma)\big)}{p\big(\sum_{j=1}^{n} e_j < k-1\big|b\big)}$ then the CEO is never fired in the first round. We prove by contradiction. Assume the CEO was fired in the first round (at time 2). In this case, the board hired a replacement and in the second round the game is equivalent to the initial one-period game (with γ replacing μ). Given the above restriction on $\frac{c}{q}$, Proposition 3.4 implies that the board doesn't fire the CEO in the second round (at time 4): directors simply keep the replacement they hired after the first round. The true quality of this replacement will be revealed at time 5. Going back to time 2, the game is now equivalent to the initial game because each director knows that if a replacement is hired, she will stay in the second round and therefore they will learn her true quality at date 5. Given the restriction on $\frac{c}{q}$, however, they can never fire the CEO in the initial game, yielding a contradiction.

If the CEO was retained in the first round of voting, then the outcome of the second round of voting is completely determined by the public signal observed at time 3. This follows because $p(B|g,H) < p(B|b,H) < 1-\gamma < p(B|g,L) < p(B|b,L)$ and therefore the only efficient equilibrium is the one in which all board members vote to retain the CEO if the public signal is H and all board members vote to fire the CEO if the public signal is L.

Given that the outcome of the second round of voting is public knowledge, we can calculate the expected payoff from having the CEO retained in the first round of voting conditional on director i's signal:

$$\begin{aligned}\pi^{round1}(retained|s_i) =& - qp(G|s_i)p(L|G) - q(1-\gamma)p(G|s_i)p(L|G)\\ &- qp(B|s_i)p(L|B) - q(1-\gamma)p(B|s_i)p(L|B) - 2qp(B|s_i)p(H|B)\\ =& - q(2-\gamma)p(G|s_i)(1-\phi) - 2qp(B|s_i)(1-\phi) - q(2-\gamma)p(B|s_i)\phi\\ =& - q(2-\gamma)(1-\phi) - qp(B|s_i)(1+(2\phi-1)(1-\gamma))\\ \equiv& \; qA(\gamma,\phi,\alpha,n,k|s_i).\end{aligned} \tag{3.10.46}$$

$qA(\gamma,\phi,\alpha,n,k|g) > qA(\gamma,\phi,\alpha,n,k|b)$ since $1+(2\phi-1)(1-\gamma) > 0$ and $p(B|b) > p(B|g)$. Hence, the value of waiting is greatest for directors who already think that the CEO is good with a high probability.

The expected payoff from having the CEO fired in the first round is easy to calculate. Since $\frac{c}{q} < \frac{p\left(\sum_{j=1}^n e_j \geq k-1 \middle| b\right)\left(p(B|b)-(1-\gamma)\right)}{p\left(\sum_{j=1}^n e_j < k-1 \middle| b\right)}$, board members will vote informatively in the second round. Hence, the CEO will be fired with the probability that k or more directors observe a bad signal, and will be retained otherwise. Let V denote the number of directors who observe a bad signal in the second round. V is binomially distributed, with the distribution parameter $1-\alpha$ if the true quality of the replacement hired after the first round of voting is good, and α otherwise. The payoff from having the CEO fired in the first round (without accounting for the dissent cost, which we will do below) is given by:

$$\begin{aligned}\pi^{round1}(fired|s_i) =& - \gamma p(b|G)cp(V<k|G) - \gamma p(b|G)q(1-\gamma)p(V\geq k|G)\\ &- \gamma p(g|G)q(1-\gamma)p(V\geq k|G) - (1-\gamma)p(b|B)(c+q)p(V<k|B)\\ &- (1-\gamma)p(b|B)q(1-\gamma)p(V\geq k|B) - (1-\gamma)p(g|B)qp(V<k|B)\\ &- (1-\gamma)p(g|B)q(1-\gamma)p(V\geq k|B)\\ =& - c\gamma(1-\alpha)p(V<k|G) - c(1-\gamma)\alpha p(V<k|B)\\ &- q\gamma(1-\alpha)(1-\gamma)p(V\geq k|G) - q\gamma\alpha(1-\gamma)p(V\geq k|G)\\ &- q(1-\gamma)\alpha p(V<k|B) - q(1-\gamma)\alpha(1-\gamma)p(V\geq k|B)\\ &- q(1-\gamma)(1-\alpha)p(V<k|B) - q(1-\gamma)(1-\alpha)(1-\gamma)p(V\geq k|B)\\ \equiv& \; cB(\gamma,\alpha,n,k) + qC(\gamma,\alpha,n,k).\end{aligned} \tag{3.10.47}$$

If $qA(\gamma,\phi,\alpha,n,k|g) > qA(\gamma,\phi,\alpha,n,k|b) > cB(\gamma,\alpha,n,k) + qC(\gamma,\alpha,n,k)$, it is always optimal to retain the incumbent CEO, regardless of the signal. If, on the other hand, $cB(\gamma,\alpha,n,k)+qC(\gamma,\alpha,n,k) > qA(\gamma,\phi,\alpha,n,k|g) > qA(\gamma,\phi,\alpha,n,k|b)$, then it is always optimal to fire the incumbent. We therefore focus on the interesting case when $qA(\gamma,\phi,\alpha,n,k|g) > cB(\gamma,\alpha,n,k) + qC(\gamma,\alpha,n,k) > qA(\gamma,\phi,\alpha,n,k|b)$. Directors with signal g always vote to keep the current CEO, so that it is the directors with signal b who determine the equilibrium. Considering the cost of dissent, they choose to vote according to their private signals if and only if:

$$
\begin{aligned}
&qA(\gamma,\phi,\alpha,n,k|b) \leq -\big(c - qA(\gamma,\phi,\alpha,n,k|b)\big)p\big(\sum_{j=1}^{n} e_j < k-1\big|b\big) \\
&+ cB(\gamma,\alpha,n,k)p\big(\sum_{j=1}^{n} e_j \geq k-1\big|b\big) + qC(\gamma,\alpha,n,k)p\big(\sum_{j=1}^{n} e_j \geq k-1\big|b\big).
\end{aligned}
\tag{3.10.48}
$$

(3.10.48) can be rewritten as

$$
\frac{c}{q} \leq \frac{\Big(-A(\gamma,\phi,\alpha,n,k|b) + C(\gamma,\alpha,n,k)\Big)p\big(\sum_{j=1}^{n} e_j \geq k-1\big|b\big)}{p\big(\sum_{j=1}^{n} e_j < k-1\big|b\big) - B(\gamma,\alpha,n,k)p\big(\sum_{j=1}^{n} e_j \geq k-1\big|b\big)}. \tag{3.10.49}
$$

Define

$$
\begin{aligned}
\Psi(\gamma,\alpha,\phi,k,n) \equiv \min\Bigg\{&\frac{\Big(-A(\gamma,\phi,\alpha,n,k|b) + C(\gamma,\alpha,n,k)\Big)p\big(\sum_{j=1}^{n} e_j \geq k-1\big|b\big)}{p\big(\sum_{j=1}^{n} e_j < k-1\big|b\big) - B(\gamma,\alpha,n,k)p\big(\sum_{j=1}^{n} e_j \geq k-1\big|b\big)}, \\
&\frac{p\big(\sum_{j=1}^{n} e_j \geq k-1\big|b\big)\big(p(B|b) - (1-\gamma)\big)}{p\big(\sum_{j=1}^{n} e_j < k-1\big|b\big)}\Bigg\}.
\end{aligned}
\tag{3.10.50}
$$

Then, if $\frac{c}{q} > \Psi(\gamma,\alpha,\phi,k,n)$, the CEO is never fired in the first round and if $\frac{c}{q} \leq \Psi(\gamma,\alpha,\phi,k,n)$, the CEO is fired in the first round with a positive probability (equal to the probability that k directors observe a bad signal given the CEO's true quality).

Bibliography

Operational Efficiency in the ARM Industry. Kaulkin Ginsberg Whitepaper, October 2006.

Value of Third-Party Debt Collection to the U.S. Economy in 2007: Survey and Analysis. Technical report, PriceWaterhouseCoopers, June 2007.

Collecting Consumer Debts: The Challenges of Change. Report, Federal Trade Comission, Washington, DC, February 2009.

Economic impact of the payday lending industry. IHS Global Insight (USA) Inc., May 2009.

Annual Report 2009: Fair Debt Collection Practices Act. Federal Trade Commission, Washington, DC, February 2009.

Repairing A Broken System: Protecting Consumers in Debt Collection Litigation and Arbitration. Report, Federal Trade Comission, Washington, DC, July 2010.

Annual Report 2011: Fair Debt Collection Practices Act. Federal Trade Commission, Washington, DC, March 2011.

Flow of Funds Accounts of the United States, Flows and Outstandings. Report, Board of Governors of the Federal Reserve System, Washington, DC, various years.

Judicial Business of the United States Courts. Report, Administrative Office of the United States Courts, Washington, DC, various years.

R. Adams and D. Ferreira. A theory of friendly boards. *Journal of Finance*, 62(1): 217, 2007.

R. Adams, B. Hermalin, and M. Weisbach. The role of boards of directors in corporate governance: A conceptual framework and survey. *Journal of Economic Literature*, 48(1):58–107, 2010.

P. Aghion and P. Bolton. An incomplete contracts approach to financial contracting. *The Review of Economic Studies*, 59(3):473–494, 1992.

M. P. Allen. Managerial power and tenure in the large corporation. *Social Forces*, 60 (2):482–494, 1981.

A. Almazan and J. Suarez. Entrenchment and severance pay in optimal governance structures. *Journal of Finance*, 58(2):519–547, 2003.

D. Austen-Smith and J. S. Banks. Information Aggregation, Rationality, and the Condorcet Jury Theorem. *American Political Science Review*, 90(1):34–45, 1996.

D. Austen-Smith and T. Feddersen. Deliberation and voting rules. *Social Choice and Strategic Decisions*, pages 269–316, 2005.

S. Bhagat and B. Black. The Non-Correlation between Board Independence and Long-Term Firm Performance. *Journal of Corporation Law*, 27:231, 2001.

P. Bolton and D. Scharfstein. A theory of predation based on agency problems in financial contracting. *The American Economic Review*, 80(1):93–106, 1990.

A. L. Boone, L. C. Field, J. M. Karpoff, and C. G. Raheja. The determinants of corporate board size and composition: An empirical analysis. *Journal of Financial Economics*, 85(1):66–101, 2007.

K. Borokhovich, R. Parrino, and T. Trapani. Outside directors and CEO selection. *Journal of Financial and Quantitative Analysis*, 31(3):337–355, 1996.

J. Y. Campbell. Household finance. *Journal of Finance*, 61(4):1553–1604, 08 2006.

J. P. Caskey. Pawnbroking in America: The Economics of a Forgotten Credit Market. *Journal of Money, Credit and Banking*, 23(1):85–99, 1991.

J. P. Caskey. *Fringe Banking: Check-Cashing Outlets, Pawnshops, and the Poor*. Russell Sage Foundation, New York, 1994.

J. P. Caskey. Fringe banking and the rise of payday lending. In P. Bolton and H. Rosenthal, editors, *Credit Markets for the Poor*. Russell Sage Foundation, New York, 2005.

T. Catan. Spain's showy debt collectors wear a tux, collect the bucks — their goal: Publicly humiliate non-payers. *The Wall Street Journal*, page A1, October 11, 2008.

E. Cohen-Cole, B. Duygan-Bump, and J. Montoriol-Garriga. Forgive and forget: who gets credit after bankruptcy and why? *Quantitative Analysis Unit Working Paper*, 2009.

J. Coles, N. Daniel, and L. Naveen. Boards: Does one size fit all? *Journal of Financial Economics*, 87(2):329–356, 2008.

P. Coughlan. In defense of unanimous jury verdicts: Mistrials, communication, and strategic voting. *American Political Science Review*, 94(2):375–393, 2000.

V. Crawford and J. Sobel. Strategic information transmission. *Econometrica*, 50(6): 1431–1451, 1982.

A. A. Dick and A. Lehnert. Personal bankruptcy and credit market competition. *Journal of Finance*, 65(2):655–686, April 2010.

I. Domowitz and R. Sartain. Determinants of the consumer bankruptcy decision. *Journal of Finance*, 54(1):403–420, 1999.

U. Doraszelski, D. Gerardi, and F. Squintani. Communication and voting with double-sided information. *Contributions to Theoretical Economics*, 3(1):6, 2003.

M. Ertugrul and K. Krishnan. Can CEO dismissals be proactive? *Journal of Corporate Finance*, 17(1):134 – 151, 2011.

S. Fay, E. Hurst, and M. White. The household bankruptcy decision. *American Economic Review*, 92(3):706–718, 2002.

V. Fedaseyeu. Debt Collection Agencies and the Supply of Consumer Credit. *Working paper, Boston College*, March 2011.

D. Ferreira, M. A. Ferreira, and C. C. Raposo. Board structure and price informativeness. *Journal of Financial Economics*, 99(3):523 – 545, 2011.

C. Fershtman, K. Judd, and E. Kalai. Observable contracts: Strategic delegation and cooperation. *International Economic Review*, 32(3):551–559, 1991.

M. J. Flannery and K. Samolyk. Payday Lending: Do the Costs Justify the Price? *SSRN eLibrary*, 2005.

M. Furletti. Credit card pricing developments and their disclosure. Technical Report 03-02, Federal Reserve Bank of Philadelphia, 2003.

A. Gillette, T. Noe, and M. Rebello. Corporate board composition, protocols, and voting behavior: experimental evidence. *Journal of Finance*, pages 1997–2031, 2003.

M. Goltsman, J. Hörner, G. Pavlov, and F. Squintani. Mediation, arbitration and negotiation. *Journal of Economic Theory*, 144(4):1397 – 1420, 2009.

R. Gropp, J. Scholz, and M. White. Personal Bankruptcy and Credit Supply and Demand. *Quarterly Journal of Economics*, 112(1):217–251, 1997.

D. Gross and N. Souleles. An empirical analysis of personal bankruptcy and delinquency. *Review of Financial Studies*, 15(1):319, 2002.

M. Harris and A. Raviv. A theory of board control and size. *Review of Financial Studies*, 21(4):1797, 2008.

O. Hart and J. Moore. Default and Renegotiation: A Dynamic Model of Debt. *Quarterly Journal of Economics*, 113(1):1–41, 1998.

B. Hermalin and M. Weisbach. Endogenously chosen boards of directors and their monitoring of the CEO. *American Economic Review*, 88(1):96–118, 1998.

D. Hirshleifer and A. Thakor. Managerial performance, boards of directors and takeover bidding. *Journal of Corporate Finance*, 1(1):63–90, 1994.

R. Hunt. Collecting consumer debt in America. *Business Review*, (Q2):11–24, 2007.

R. Hynes. Broke But Not Bankrupt: Consumer Debt Collection in State Courts. *Fla. L. Rev.*, 60, 2008.

R. M. Hynes, A. E. Dawsey, and L. M. Ausubel. The Regulation of Non-Judicial Debt Collection and the Consumer's Choice Among Repayment, Bankruptcy and Informal Bankruptcy. *SSRN eLibrary*, 2009.

M. Jensen. The Modern Industrial Revolution, Exit, and the Failure of Internal Control Systems. *Journal of Finance*, 48(3), 1993.

D. Karlan and J. Zinman. Expanding credit access: Using randomized supply decisions to estimate the impacts. *Review of Financial Studies*, 2009.

O. Kini, W. Kracaw, and S. Mian. Corporate takeovers, firm performance, and board composition. *Journal of Corporate Finance*, 1(3):383–412, 1995.

W. Krumbein. Collection Agencies. *University Journal of Business*, pages 48–67, 1924.

R. La Porta, F. Lopez-de Silanes, A. Shleifer, and R. Vishny. Law and finance. *Journal of political Economy*, 106(6):1113–1155, 1998. ISSN 0022-3808.

J. Linck, J. Netter, and T. Yang. The determinants of board structure. *Journal of Financial Economics*, 87(2):308–328, 2008.

M. Lipton and J. Lorsch. A Modest Proposal for Improved Corporate Governance,. *Business Lawyer*, 48(1):59–77, 1992.

J. Lorsch and E. MacIver. *Pawns or Potentates? The Reality of America's Corporate Boards*. Harvard Business School Press, Boston, MA, 1994.

J. S. Lublin. Reimbursements aim for a fairer proxy fight. *The Wall Street Journal*, page A22, October 27, 2009.

M. Mace. *Directors: Myth and Reality.* Harvard Business School Press, Boston, MA, 1971.

N. Malenko. Communication and Decision-Making in Corporate Boards. *Working paper, Stanford University*, 2011.

B. Melzer. The real costs of credit access: Evidence from the payday lending market. Working paper, Kellogg School of Management, Northwestern University, January 2009.

W. Mikkelson, M. Megan Partch, and K. Shah. Ownership and operating performance of companies that go public. *Journal of Financial Economics*, 44(3):281–307, 1997.

R. A. G. Monks and N. Minow. *Corporate Governance.* Basil Blackwell, Cambridge, 1995.

D. Morgan and M. Strain. Payday holiday: How households fare after payday credit bans. *Federal Reserve Bank of New York Staff Reports*, 309, 2007.

A. Morse. Payday Lenders: Heroes or Villains? *SSRN eLibrary*, January 2009.

R. Myerson. Optimal coordination mechanisms in generalized principal-agent problems. *Journal of Mathematical Economics*, 10(1):67–81, 1982.

R. Parrino. CEO turnover and outside succession a cross-sectional analysis. *Journal of Financial Economics*, 46(2):165–197, 1997.

C. Raheja. Determinants of board size and composition: A theory of corporate boards. *Journal of Financial and Quantitative Analysis*, pages 283–306, 2005.

J. Scott and T. Smith. The effect of the bankruptcy reform act of 1978 on small business loan pricing. *Journal of Financial Economics*, 16(1):119–140, 1986.

A. Shivdasani and D. Yermack. CEO involvement in the selection of new board members: An empirical analysis. *Journal of Finance*, 54(5):1829–1853, 1999.

J. Stiglitz and A. Weiss. Credit rationing in markets with imperfect information. *The American Economic Review*, 71(3):393–410, 1981.

C. Tejada. Longtime Tandy director quits board, says he was punished for faulting CEO. *The Wall Street Journal*, January 16, 1997.

P. Tufano. Consumer finance. *Annual Review of Financial Economics*, 1(1):227–247, 2009.

V. A. Warther. Board effectiveness and board dissent: A model of the board's relationship to management and shareholders. *Journal of Corporate Finance*, 4(1): 53–70, 1998.

M. Weisbach. Outside directors and CEO turnover. *Journal of Financial Economics*, 20:431–460, 1988.

M. White. Bankruptcy reform and credit cards. *The Journal of Economic Perspectives*, pages 175–200, 2007.

D. Yermack. Higher market valuation for firms with a small board of directors. *Journal of Financial Economics*, 40(2):185–211, 1996.

J. Zinman. Restricting consumer credit access: Household survey evidence on effects around the Oregon rate cap. *Journal of Banking & Finance*, 2009.