# The Impact of Financial Incentives on Firm Behavior

by

David Abraham Matsa

B.S., Economics
Massachusetts Institute of Technology, 2000

B.S., Mathematics
Massachusetts Institute of Technology, 2000

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2006



Author ..........................................................
Department of Economics
May 15, 2006

Certified by ..........................................................
Nancy L. Rose
Professor
Thesis Supervisor

Certified by ..........................................................
Joshua Angrist
Professor
Thesis Supervisor

Accepted by ..........................................................
Peter Temin
Elisha Gray II Professor of Economics
Chairman, Department Committee on Graduate Students

## The Impact of Financial Incentives on Firm Behavior

by

David Abraham Matsa

Submitted to the Department of Economics
on May 15, 2006, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Economics


### Abstract

This dissertation analyzes the impact of various financial incentives on firm behavior. The first two chapters examine product-market and input-market effects of a firm's capital structure and the incentives they create. The third chapter analyzes how incentives from the tort system affect physician location decisions.

Chapter 1 examines the impact of union bargaining on capital structure determination. If a firm maintains a high level of liquidity, workers may be encouraged to raise wage demands. In the presence of external finance constraints, a firm has an incentive to use the cash flow demands of debt service payments to improve its bargaining position. Using both cross-sectional estimates of firm-level collective bargaining coverage and state changes in labor law to identify changes in union bargaining power, I show that firms indeed appear to use financial leverage strategically to influence collective bargaining negotiations. These estimates suggest that strategic incentives from union bargaining have a substantial impact on financing decisions.

A firm's financial structure can also impact investments in marketing and operations management. Chapter 2 examines how capital structure affects a firm's provision of product availability – an important dimension of product quality in the retail sector. Using U.S. consumer price index microdata to measure the prevalence of out-of-stocks, I find that supermarket leveraged buyouts, which reduce liquidity, increase out-of-stocks by 10 percent. These findings suggest it is important for firms to consider these sorts of real effects on their operations when setting financial policy.

Chapter 3 examines financial incentives created by medical malpractice liability. If patients bear the full incidence of cost changes and market demand is inelastic, then marginal changes in malpractice liability will not affect physicians' net income or location decisions. Using county-level, specialty-specific data on physician location from 1970 to 2000, I find that damage caps do not affect physician supply for the average resident of states adopting reforms. On the other hand, caps appear to increase the supply of specialist physicians in the most rural areas by 10 to 12 percent. This is likely because rural doctors face greater uninsured litigation costs and a more elastic demand for medical services.


Thesis Supervisor: Nancy L. Rose
Title: Professor

Thesis Supervisor: Joshua Angrist
Title: Professor

*for mom and dad*

## Acknowledgments

I am deeply grateful to my advisors, Nancy Rose and Josh Angrist, for introducing me to economics research – first as an undergraduate and then as a doctoral student. I appreciate their instruction, steadfast encouragement, and guidance throughout. They have been most generous with their time and have been true mentors. While not officially on my thesis committee, Dirk Jenter has been a tremendous resource. His guidance has improved this research and helped shape my life's path. I also thank Glenn Ellison and Jean Tirole for valuable advice and suggestions.

Conversations with my peers, including Michael Anderson, Todd Gormley, Dominique Olie Lauga, Dan Hungerman, and Amir Sufi, bolstered various sections of this research. I also appreciate comments from Daniel Bergstresser, Henry Farber, Amy Finkelstein, Richard Freeman, Chris Hansen, Jonathan Klick, Steven Levitt, Michelle Mello, Amalia Miller, Joseph Newhouse, Michael Piore, Joshua Rauh, David Scharfstein, Antoinette Schoar, and members of the MIT industrial organization, finance, labor economics, and organization economics lunch groups, and various seminar participants around the globe.

I gratefully acknowledge financial support in the form of fellowships from the National Science Foundation and the MIT World Economy Lab. I would also like to thank Richard Freeman and Barry Hirsch for providing me with the firm-level estimates of collective bargaining coverage that I use in Chapter 1, and Robert Oshel for providing the custom extract from the National Practitioner Data Bank that I use in Chapter 3.

The research presented in Chapter 2 was conducted in coordination with the U.S. Bureau of Labor Statistics under an Agency Agreement. I am grateful to Bill Cook, Craig Brown, Mark Bowman, Dan Ginsburg, and others at the BLS who were extremely helpful throughout the project, and to Mark Bils for sharing portions of his computer code and insight into the data. I would also like to thank Garry Van Siclen and Trade Dimensions for providing the supermarket establishment data, and Judy Chevalier for sharing the information she collected on supermarket LBOs. This research would not have been possible without the support of the George and Obie Shultz Fund and the National Science Foundation (Grant No. SES-0551097). Any opinions, findings, or conclusions expressed herein are my own and do not necessarily reflect the views of the Bureau of Labor Statistics.

I would also like to thank the support staff in the economics department, especially Gary

King, whose good nature, resourcefulness, and quick thinking have been indispensable.

Graduate school would have been far more difficult if not for the support of close friends. I am grateful to classmates, officemates, and flatmates — including Patricia Cortes, Antara Dutta, Josh Fischman, Alan Grant, Guy Michaels, Analia Schlosser, and Pete Sperber — who provided comraderie, good humor, and many insights. I would especially like to thank Mark Histed, Mike Wagner, and Kevin Amonlirdvaman for all those lunches at the food trucks and an infinite willingness to listen. All of these friendships helped make my graduate experience as rewarding personally as it was professionally.

Most of all, I thank my partner and soon-to-be wife, Lesley. She has provided me with a constant source of joy and stability throughout the ups and downs of my thesis writing experience. Her love, devotion, and strength are truly inspiring.

Last but certainly not least, I want to thank my parents. I am honored to share this graduation day with my father's 50th MIT reunion and all of the pageantry that rightfully entails. I fondly recall, shortly after starting college, my mother boastingly sending me a "news" article reporting that a boy's brains can be attributed to his mother. In addition to the genes, I am indebted to both my parents for instilling in me insatiable intellectual curiosity and a sound moral compass. With love and appreciation, I dedicate this thesis to them.

*David A. Matsa*
*May 15, 2006*

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Capital Structure as a Strategic Variable: Evidence from Collective Bargaining

*Leverage means debt. More debt for Eastern meant greater pressure to cut costs.... We're going to have a fiasco in this country, because twenty percent of the [passenger aircraft] capacity in this country has been allowed to come under the control of a highly leveraged company which is embarked on a confrontation between labor and interest costs. It's not labor and management. It's labor and interest cost.*

Farrell Kupersmith
Pilots' Union Representative[1]

## 1.1 Introduction

The product-market and input-market effects of capital structure link the financial and real activities of a firm. A vast theoretical literature relates financial structure to market conduct and postulates that firms use leverage strategically.[2] However, empirical work in this area is far less developed. There is considerable evidence that changes in firm capital structure affect product-market behavior, including entry, exit, and pricing (Chevalier 1995a,b; Chevalier and Scharfstein 1996; Zingales 1998).[3] But how do firms respond to

[1] *Frontline*, "The Battle for Eastern Airlines," January 31, 1989.
[2] For surveys, see Harris and Raviv (1991) and Franck and Huyghebaert (2004).
[3] Other related empirical papers include Opler and Titman (1994), Phillips (1995), Kovenock and Phillips (1997), Campello (2003), Campello and Fluck (2004), and Brown, Fee, and Thomas (2005). The Franck and

such incentives? This chapter fills an important gap by showing that strategic incentives from input markets have a substantial impact on financing decisions.

A firm may use its financial structure to alter the behavior of competitors, customers, and suppliers. For example, the presence of debt, and particularly debt due in the near-term ("current debt"), may improve a firm's bargaining position with suppliers or customers who possess market power. This chapter focuses specifically on the effects of union power, which is widely associated with raising wages and imposing other costs on employers (Lewis 1986).[4] Although U.S. firms are required by law to bargain with employee collectives in good faith, firms can also act to reduce the impact of bargaining on profits. Just as some firms attempt to prevent unions from organizing in the first place (Freeman 1986), they may also undertake costly actions that improve their negotiating position. One possible strategy is increasing leverage to reduce liquidity. By increasing the cash flow demands of debt service, firms may credibly commit to be tougher in labor negotiations.[5]

Delta Airlines' recent experience exemplifies how excessive liquidity can hurt a firm's bargaining position with workers. Delta was generally considered to be one of the strongest carriers leading up the September 11, 2001, shock to the U.S. airline industry. Following its history of fiscally conservative management, Delta weathered this downturn by building up cash and liquidity. But increased liquidity also reduced the need to cut costs, hurting Delta's bargaining position with workers (Perez 2004).[6] By 2004, Delta found itself far behind the other big carriers in restructuring, and in severe financial distress.

To illustrate the impact of collective bargaining on a firm's optimal debt policy, I embed efficient Nash bargaining over the wage bill in an agency-cost model of corporate financing. I show that market power in the hands of a supplier (such as organized labor) has two effects on the firm's capital structure: a balance sheet effect and a strategic effect. First, collective

Huyghebaert (2004) survey article also covers this literature.

[4] A recent paper by Dinardo and Lee (2004) is an exception.

[5] Various other actions may improve a firm's bargaining position as well, such as underfunding pension plans (Ippolito 1985) and strategically fashioning managerial compensation (Wilson 2004). For some firms, employment policy may also serve as an anti-takeover device (Pagano and Volpin 2005). To encourage specific capital investment, labor may also be interested in making pre-commitments to reduce the holdup problem, such as organizing in multiple, separate unions (Ulph 1989) and choosing a certain type of union leader (Skatun 1997).

[6] Delta's then-CEO, Leo Mullin, said in a recent interview, "Because we managed ourselves in a financially responsible way,... we ended up in the tougher bargaining position." Then-board member, James Broadhead, describes, "The pilots wouldn't give anything until the last possible minute." Many of the other legacy carriers, who face powerful unions throughout their workforce, are highly leveraged, and labor has claimed that the use of leverage is strategic (for example, see the quote on page 15).

bargaining effectively weakens the balance sheet of the firm by redistributing project returns from investors to workers, forcing the firm to hold more current debt. Second, the firm uses additional current debt strategically to reduce excess liquidity and improve its bargaining position.[7] Distinguishing these effects is important; while the balance sheet effect is a response to the outcome of labor negotiations, the strategic effect is an attempt to influence them.

Existing evidence on the link between collective bargaining and capital structure determination relies on cross-sectional comparisons that may be affected by omitted variables bias. For example, Bronars and Deere (1991) show that unionization rates are correlated with financial leverage at the industry level. This correlation may be interpreted as evidence that non-unionized firms issue debt when faced with the threat of unionization, but this relation is also consistent with a range of noncausal scenarios.[8] This sort of simple correlation also fails to distinguish strategic increases in debt from the more "mechanical" balance sheet effect.[9]

I identify the strategic effect empirically using variation in profit variability across firms, which reflects differences in the specific product markets in which firms compete. When labor and management bargain, a union can claim a portion of a firm's excess liquidity — its operating cash flow net of any required debt payments. Collective bargaining thus imposes a greater threat to a firm when the firm maintains higher levels of excess liquidity. With limited liability and positive debt balances, greater underlying profit variability is one factor that increases expected excess liquidity and a firm's exposure to union rent seeking. Greater variability in the profitability of potential projects implies that the firm must, on average,

[7]The firm adopts this strategy in order to attract investors in the capital market. When a credit-constrained firm is unionized, it would have difficulty obtaining financing if not for the greater levels of current debt.

[8]For example, unions are likely attracted to established, profitable industries, which may have a greater capacity for debt. In other cross-sectional analyses, Hirsch (1991) finds that the ratio of debt to equity is higher in union companies. Sarig (1998) finds that an estimate of labor's share of the firm's quasi-rent is positively correlated with financial leverage. Gorton and Schmid (2004) use data on German firms to show that firms subject to codeterminiation laws (requiring partial employee corporate control) have greater leverage than other firms. Hanka (1998) finds that debt is negatively correlated with employment, wages, and pension funding, and positively correlated with the use of part-time and seasonal employees. Kale and Shahrur (2004) show that a firm's leverage is negatively related to the R&D intensity in its supplier and customer industries. Brown, Fee, and Thomas (2005) find that leveraged buyouts are associated with a reduction in firms' costs of goods sold.

[9]While I do not know of any papers that explicitly aim to disentangle the balance sheet and strategic effects, empirical correlations presented in Cavanaugh and Garen (1997) can be interpreted as evidence of the strategic effect. They show that the correlation between collective bargaining and debt increases with rough proxies for the specificity of a firm's assets.

maintain greater excess liquidity in order to fund the same marginal project.[10] Firms with greater profit variability are thus more vulnerable to union rent seeking and have a greater incentive to use debt to shield liquidity from workers in bargaining. Consequently, evidence of the strategic effect can be found by analyzing the interaction between union bargaining power and profit variability.

I explore these implications using two very different estimation strategies. Both approaches regress measures of financial debt on proxies for union bargaining power and exploit the interaction between union bargaining power and profit variability to capture the strategic effect. The first approach measures cross-sectional correlations using data on collective bargaining coverage (a direct measure of union power) for samples of manufacturing firms from the 1970s, 1980s, and 1990s. The second approach is less direct but avoids the potential biases of cross-sectional estimation by using state adoption of right-to-work laws in the 1950s and state repeal of unemployment insurance work stoppage provisions in the 1960s and early 1970s as sources of exogenous variation in union power.

The results suggest that union bargaining power leads firms to increase financial leverage, with larger increases in current debt at firms with greater profit variability. Estimates suggest this effect is sizeable: for firms with profit variability one standard deviation above the mean, the ratio of current debt to total firm value is 5 to 10 percent greater when an additional 10 percent of employees bargain collectively. In contrast, for firms with little profit variability, differences in union coverage rates seem to have little to no effect. Because profit variability increases exposure to union rent seeking, the effect of its interaction with union power is evidence of the strategic effect.

Analysis of the effects of changes in labor laws also suggests that strategic considerations with respect to input markets substantially influence capital structure determination. After states adopt legislation that reduces union bargaining power, firms that face concentrated labor markets and have more variable profits decrease current debt relative to otherwise similar firms with less variable profits. For firms with profit variability one standard deviation above the mean, the ratio of current debt to total firm value decreases by up to two-thirds after a right-to-work law is passed, and by one-fifth after a work stoppage provision is repealed. In contrast, these changes in labor law have little effect on firms with

[10]A union's claim on excess liquidity can be thought of as a real option. Greater underlying variability increases the value of the option.

little profit variability. As a falsification test, I show that these changes in labor law also do not affect financial policy at firms in industries with low union presence. Various tests demonstrate the robustness of the profit variability interaction.

The remainder of the chapter is organized as follows. Section 1.2 develops a model of capital structure determination at a firm with collective bargaining and derives comparative statics which can be tested empirically. Cross-sectional estimates are presented in Section 1.3, and the labor law evidence is presented in Section 1.4. Section 1.5 verifies that, in contrast to their use of debt service, firms are unable to use dividend policy as a commitment device with respect to organized workers. Section 1.6 concludes.

## 1.2 Theoretical model

Since Modigliani and Miller (1958), economists have been studying frictions that affect a firm's choice of financial structure.[11] In one traditional explanation, financial leverage is thought to mediate agency problems within firms. Easterbrook (1984) and Jensen (1986) consider the free cash flow problem at cash-rich firms which generate cash inflows exceeding their efficient reinvestment needs. Such firms have excess liquidity that must be "pumped out" to prevent it from being spent unprofitably, such as on poor projects, unwarranted diversification, or wasteful perks. Just as the cash flow demand of debt service payments may serve this function, it may be used to influence negotiations with suppliers, including organized labor. At a firm with collective bargaining, debt may commit a firm to disgorge excess free cash flow which might otherwise bolster wage demands.

I use generalized Nash bargaining to model negotiations with unionized workers to show how supplier market power affects a firm's optimal debt policy. A key feature of debt policy is that it is generally set unilaterally, without labor's consent (Baldwin 1983). Management fixes a firm's capital structure subject to capital market constraints. When a union has market power but cannot commit to future negotiating positions (Grout 1984), management will likely consider labor market ramifications in choosing its debt policy.[12] Even with

[11]Other than product/input market interactions, theories of capital structure determination focus on agency costs, asymmetric information, corporate control considerations, and taxes (Harris and Raviv 1991).

[12]If the union could commit to future negotiating positions, and the firm aims to maximize the returns to shareholders, and labor is interested in maximizing total income, then a generalized Nash bargaining solution implies that the capital structure will not depend on bargaining power and that economic profit will be shared by labor and the firm. However, without the ability to credibly commit not to demand more

efficient bargaining over wages, employment levels, and work rules, firms will structure financing to maximize the returns to shareholders at the expense of efficiency.[13]

Notions of the capacity to use debt as a means to partially control wage demands date back to Baldwin (1983), if not before, and formal models were developed by Bronars and Deere (1991) and Dasgupta and Sengupta (1993).[14] Bronars and Deere focus specifically on the use of debt by a firm facing the threat of unionization, and Dasgupta and Sengupta model a firm where labor is already organized. But these papers conflate the distinct channels through which unionization affects a firm's capital structure. In addition to providing a strategic incentive for debt financing, union bargaining likely weakens the balance sheet of a firm by increasing wages, leading to additional increases in leverage. At a credit-constrained firm, interactions of the various functions of capital structure generate these effects. Current debt serves the dual functions of (1) specifying the contingencies for continued financing and (2) removing excess liquidity from the firm. The model presented here embeds collective bargaining within a broader model of debt maturity structure, incorporating the various functions of capital structure and generating a testable implication of the *strategic* use of debt at unionized firms.

To fully understand how a supplier's bargaining power influences a firm's optimal capital structure, it is important to consider how some sort of financing friction can drive a wedge between internal and external financing. I introduce managerial moral hazard into a model of multi-stage financing, building on the Holmström and Tirole (1996) model of liquidity management. Firms are considered "cash-rich" ongoing entities, generating cash beyond any investment cost overruns that must be paid to complete the project. To ensure the firm does not overinvest, investors hold current debt, thereby pumping the excess cash out of the firm.

---

later, and without deep enough pockets to subsidize the debt issuance *ex ante*, labor cannot ensure there is efficient continuation.

[13]Grout (1984) and Baldwin (1983) adopt similar approaches in modeling investment. In their models, shareholders find it optimal to underinvest in fixed assets, because the quasi-rents generated by such investments accrue partly to workers when there is collective bargaining. Similar to the strategic debt policy considered in this chapter, this underinvestment can be interpreted as reducing the cash flows available to the union in bargaining (Bronars and Deere 1991).

[14]Based on similar intuition, Perotti and Spier (1993) develop a model to motivate the LBO wave of 1980s, wheras Sarig (1998) argues that leverage weakens shareholders' bargaining posture vis-à-vis employees who possess firm-specific human capital. Spiegel and Spulber (1994) argue that regulated utilities have an incentive to increase leverage to elicit greater retail prices from regulators. In a sense, the use of debt in "bootstrap acquisitions" can also be interpretted as leverage being used strategically to impact "bargaining" between a corporate raider and target shareholders over takeover gains (Müller and Panunzi 2004).

The timing of the model is depicted in Figure 1-1. There are three periods. At date 0, a firm with wealth $A$ faces an investment opportunity whereby an initial investment $I$ returns $R$ if the project succeeds and 0 if it fails. To finance the fixed investment cost $I$, the firm needs to borrow $I-A$. In exchange for receiving $I-A$ at date 0, the firm agrees to pay investors at least $d$ at date 1 (or face liquidation) and $D$ at date 2 (if the project succeeds). Management of the firm is protected by limited liability, and the financial markets are competitive with investors demanding a rate of return equal to 0. Management, labor, and investors, are risk neutral.

At date 1, the investment yields a deterministic payoff $r \geq 0$. However, the firm also experiences a liquidity shock, $\nu$, that it must withstand in order to continue with the project. The amount of $\nu$ is unknown ex ante and has a cumulative distribution function $F(\nu)$ with density $f(\nu)$ on $[0, \infty)$. This liquidity shock can simply be interpreted as a need for cash to cover operating expenditures and any investment cost overrun, but it can equivalently be considered a shortfall in date 1 earnings (Tirole 2004, p. 280). In addition to paying $\nu$, continuation requires hiring workers. If the firm is unionized, the wage bill, $w(\nu)$, is the outcome of Nash bargaining, described further below, and is paid at date 1 from the firm's current cash flow. Otherwise, the firm pays workers their alternative wage, which is normalized to 0. If $\nu$ is paid and workers are hired, the project continues and a final payoff is realized at date 2. If the firm fails to either reinvest or hire workers, it is liquidated with value equal to 0.

Once the firm withstands the liquidity shock and successfully hires workers, management privately chooses the probability $p$ that the project succeeds, subjecting investment to moral hazard. Management can either put forth high or low effort. If management is diligent, the probability of success is $p_H$; whereas if it slacks, it enjoys a private benefit $B > 0$ but reduces the probability of success to $p_L$, where $p_H - p_L \equiv \Delta p > 0$. Assume the project's net present value is positive if management is diligent but not if it shirks. Given that the firm is cash-rich, a sufficient condition is:

$$p_H R - I > 0 > p_L R - I + r + B$$

The moral hazard drives a wedge between the project's net present value (expected return) and its pledgeable income (expected return to investors), constraining the borrowing opportunities of the firm.

At date 0, the firm contracts with outside investors, specifying the amount that investors will contribute and the distribution of the proceeds of investment, i.e., $d$ and $D$. As in Holmström and Tirole (1996), assume that the project outcome and the liquidity shock are verifiable. For simplicity, also assume $D$ cannot be refinanced at date 1.[15] By constraining cash flows at the intermediate date, the contracted level of current debt, $d$, implicitly also determines the contingencies in which the project will be continued at date 1. Since continuation is attractive to the firm, it will continue whenever it is financially feasible, that is for all $\nu$ such that

$$\nu + w(\nu) \leq \nu^* \equiv r - d$$

In this sense, the firm uses current debt to credibly commit to investors that it will only continue the project beyond date 1 if the costs of continuation are low enough — below some agreed upon cutoff, denoted $\nu^*$.

In addition to determining the continuation cutoff, the cash flow demands of debt service also crucially affect the negotiations between management and labor. When the time comes to hire workers at date 1, it is common knowledge that the firm has $r - d - \nu$ in excess liquidity — cash on hand net of the impending current debt payment — from which to pay workers. If the parties cannot come to terms, the firm liquidates and is worth 0 and workers find work elsewhere (or collect unemployment benefits), also valued at 0.[16] If the parties reach an agreement, the excess liquidity may be divided in any number of ways.

Following Grout (1984) and Baldwin (1983), I adopt a generalized Nash bargaining solution where management aims to maximize the returns to shareholders, labor is interested in maximizing total income, and union and management respective bargaining powers are

[15] Refinancing is infeasible if, for example, investors are dispersed and it is sufficiently costly to dilute their claims through renegotiation. Alternatively if renegotiation were feasible, the firm's soft budget constraint would increase current debt in the optimal capital structure. In fact, assuming workers are unable to hold up lenders (final wage negotiation follows refinancing) and the managerial agency problem is relatively mild ($\frac{B}{\Delta p}$ is sufficiently small relative to $R$), then the firm would strategically exploit refinancing by further increasing current debt, thereby decreasing wages, and relying on opportunities to refinance at date 1 to fund projects.

[16] One might ask whether workers might disagree with the firm to force it into bankruptcy and then attempt to divert some of $d$ from creditors (Bronars and Deere 1991). If breaking off negotiations with the firm is sufficiently costly, it is not in the interest of workers to do so (Dasgupta and Sengupta 1993). Weiss (1990), for example, estimates that the direct costs of bankruptcy average 3 percent of the equity value of the firm, and the costs of a strike may be even greater. Furthermore, even when this may be a beneficial strategy in a one-shot game, it is unlikely to be optimal in a more realistic model which includes repeated interactions.

$z \in [0,1]$ and $1-z$.[17] Consequently, the workers are hired and paid a wage equal to

$$w(\nu) = z(\nu^* - \nu) = z(r - d - \nu) \tag{1.1}$$

Several observations are immediate. The negotiated wage is increasing in the workers' bargaining power, but decreasing in the level of current debt and the realized value of nonlabor operating costs. That is, if nonlabor costs overrun, the union will be less able to elicit a favorable wage. For the project at the continuation margin, $\nu = r - d$, the union is left with zero surplus, $w = 0$. Note that $\nu^*$, the implicitly contracted cutoff in total continuation costs $\nu + w(\nu)$, is then also the firm's cutoff in $\nu$; the project will continue whenever $\nu \leq \nu^*$. Equation (1.1) also implies that, for any $\nu$, current debt reduces the union rent.

The net present value of investment is maximized by continuing the project if and only if

$$\nu \leq \nu_1^* \equiv p_H R - z\frac{F(\nu_1^*)}{f(\nu_1^*)}$$

that is, whenever the expected return $p_H R$ from continuation exceeds the operating cost $\nu$ plus the expected wage $z\frac{F(\nu_1^*)}{f(\nu_1^*)}$. The optimal cutoff is decreasing in union bargaining power, and allows less continuation than what is optimal for a nonunion firm. However, if investors are to ensure that management is diligent (chooses $p_H$), they cannot take the entire project return. Investors must structure their claim at date 2 to provide the appropriate incentives for management:[18]

$$(\Delta p)(R - D) \geq B$$

This moral hazard constraint reduces the project's expected return to investors below its NPV. In fact, pledgeable income is maximized by continuing the project if and only if

$$\nu \leq \nu_0^* \equiv p_H R - p_H \frac{B}{\Delta p} - z\frac{F(\nu_0^*)}{f(\nu_0^*)}$$

[17]The bargaining game can be formulated in any number of equivalent ways. For example, a crude model final offer arbitration allows both parties to propose a take it-it-or-leave-it offer (Blanchard and Tirole 2004). Based on a random draw, the union offer is selected with probability $z$ and the firm's offer is selected with probability $1-z$. Then, the expected wage for any realization of $\nu$ is $z(r-d-\nu)$.

[18]When approaching the capital market to finance the new project, management represents the interests of incumbent shareholders, subject to the shirking temptation. This assumption is most natural when management is the sole proprietor before financing.

which is to say, whenever the expected return $p_H R$ from continuation exceeds the sum of the operating cost $\nu$, the expected wage $z\frac{F(\nu_0^*)}{f(\nu_0^*)}$ paid to workers, and the moral hazard payment $p_H\frac{B}{\Delta p}$ made to the firm.

There are then four cases to consider, summarized in Figure 1-2. To simplify notation, let $\rho_0 \equiv p_H R - p_H\frac{B}{\Delta p}$, and let $P(\nu^*)$ denote the maximal net income pledgeable to investors when the continuation cutoff is $\nu^*$.

$$P(\nu^*) = r + F(\nu^*)\rho_0 - \left[I - A + \int_0^{\nu^*} \nu f(\nu)d\nu + \int_0^{\nu^*} z(\nu^* - \nu)f(\nu)d\nu\right]$$

If $P(\nu_1^*) \geq I - A$, then the firm is able to secure enough funding to implement its optimal cutoff $\nu_1^*$. If $P(\nu_0^*) < I - A$, achieving funding is not feasible, as pledgeable income is insufficient to compensate investors. Let $\overline{P}(\nu^*)$ denote the maximal pledgeable income of a nonunionized firm.[19] Over a subset of the infeasible range, $P(\nu_0^*) < I - A \leq \overline{P}(\rho_0)$, funding would be possible if the firm were not unionized.

The final case is the most interesting both theoretically and empirically.[20] If $P(\nu_1^*) < I - A \leq P(\nu_0^*)$, the firm will obtain some financing but not enough to support the unconstrained-optimal continuation cutoff. With a competitive capital market, the optimal contract for the credit-constrained firm has $d = r - \nu^*$ and $D = R - \frac{B}{\Delta P}$, where the cutoff $\nu^* \in [\nu_0^*, \nu_1^*]$ is given by the investors' breakeven condition:

$$r + F(\nu^*)\rho_0 = I - A + \int_0^{\nu^*} \nu f(\nu)d\nu + \int_0^{\nu^*} z(\nu^* - \nu)f(\nu)d\nu \tag{1.2}$$

With continuation cutoff $\nu^*$, the expected returns accruing to investors, both at date 1 and at date 2, equal their expected investment outlays, including $I - A$ at date 0 and the operating expenses and union rent paid at date 1. Rearranging the terms in equation (1.2) and performing integration by parts yields another implicit expression for $\nu^*$:

$$r + \int_0^{\nu^*} (\rho_0 - \nu) f(\nu)d\nu = I - A + z\int_0^{\nu^*} F(\nu)d\nu \tag{1.3}$$

Equation (1.3) shows that setting $\nu^*$ involves a tradeoff between the two inseparable prod-

[19]Formally, $\overline{P}(\nu^*) = r + F(\nu^*)\rho_0 - \left[I - A + \int_0^{\nu^*} \nu f(\nu)d\nu\right]$.

[20]Extensive empirical evidence of external financing constraints (e.g., Fazzari, Hubbard, and Petersen 1988; Lamont 1997; Rauh 2006) suggests this case is the most relevant for many firms.

ucts of continuation: (1) project surplus (net of moral hazard payment) created at the continuation margin, $\int_0^{\nu^*}(\rho_0 - \nu)\, f(\nu) d\nu$, and (2) labor rent $\int_0^{\nu^*} zF(\nu) d\nu$, which accrues infra-marginally and is increasing in union bargaining power. The strength of the union does not impact the magnitude of project surplus at the margin, because as shown above, the union does not earn rents from the marginal project. Rather, the union's rent accrues infra-marginally, with current debt (and the effective continuation cutoff) influencing the magnitude of its wage demands.

Key comparative static relationships follow directly from equation (1.3). As union bargaining power increases, the union rent increases, the project NPV decreases, the firm continues less often, and therefore carries more current debt.

$$\frac{\partial \nu^*}{\partial z} < 0, \;\; \frac{\partial d}{\partial z} > 0 \tag{1.4}$$

Increasing the variability of the reinvestment need (or "profit shortfall"), by applying a mean preserving spread, has an ambiguous effect on the debt structure and the contingencies for optimal continuation. There is a tradeoff; for a given continuation threshold, increasing variability makes the project succeed less often, but when it does succeed, expected profits are greater. Because the level of debt is being set by the investors' breakeven condition, whether more variability increases or decreases current debt depends on which effect dominates — that is, it relies on the distribution of $\nu$.[21]

$$\frac{\partial \nu^*}{\partial \sigma} \gtrless 0, \;\; \frac{\partial d}{\partial \sigma} \lessgtr 0 \tag{1.5}$$

While the main effect of variability on current debt could be positive or negative, the interaction effect of increasing both union bargaining power and reinvestment cost/profit variability is always positive (regardless of distributional assumptions). Since unions do not earn a rent at the continuation margin, they do not affect the frequency with which the project succeeds, only the level of expected profits when the project is successful.[22] Greater variability exposes more liquidity to union capture. As illustrated in equation (1.3), while the added liquidity does not impact the effect of union power on project surplus at the

[21] While perhaps a slight abuse of notation, I use $\frac{\partial}{\partial \sigma}$ to represent the effect of applying a mean preserving spread.

[22] This claim is consistent with evidence presented by Freeman and Kleiner (1999), which suggests that while unions reduce profits, they do not generally demand so much as to force firms out of busiess.

margin (it is already 0), it does increase the effect of union power on infra-marginal wage costs. The interaction leads investors to require the firm to continue less often and carry more current debt.

$$\frac{\partial^2 \nu^*}{\partial z \partial \sigma} < 0, \quad \frac{\partial^2 d}{\partial z \partial \sigma} > 0 \tag{1.6}$$

Thus, collective bargaining has two effects on the firm's capital structure: a balance sheet effect and a strategic effect. The balance sheet effect is in some sense mechanical; paying higher wages increases the costs of the project, reducing its attractiveness as well as the favorableness of the terms of financing. The strategic effect, on the other hand, comes from debt changing union behavior. Workers' demands in negotiations reflect the cash flow requirements imposed by debt service. Given costs, increasing the level of current debt decreases the negotiated wage, leaving more value on the table for the firm and its investors. While the balance sheet effect is a response to the outcome of labor negotiations, the strategic effect is an attempt to influence these negotiations.

The effects of the interaction of union power and profit variability provide the most direct evidence of firms using debt *strategically* to assuage the wage demands of unionized workers. Whereas the comparative static represented by equation (1.4) reflects a combination of the balance sheet and strategic effects, the interaction effect represented by equation (1.6) isolates the strategic effect. Beginning with Brown and Medoff (1978), some researchers have argued that unions or labor relations improve productivity; for example, they may reduce turnover and improve morale and worker cooperation.[23] If unions or labor relations have positive externalities on project returns, the net balance sheet effect may not be positive, and straight empirical correlations might suggest that unions have little net effect on capital structure. The interaction effect, on the other hand, reflects only the strategic effect. As the combination of union power and profit variability increases the scope for union rent seeking, the interaction effect most closely reflects firms' *strategic* use of debt.

[23]Empirical evidence of union productivity effects is mixed. It seems labor relations, rather than unionization or collective bargaining per se, contribute to productivity. See summaries in Belman (1992) and Freeman (1992) and recent evidence presented by Kleiner, Leonard, and Pilarski (2004), Krueger and Mas (2004), and Mas (2004).

## 1.3 Cross-sectional evidence

While notions of using capital structure to influence collective bargaining are not new, the existing empirical evidence is weak.[24] Identifying a relationship is fraught with challenges. Union bargaining power is not a well-defined empirical construct; estimates of natural proxies — such as firm-level collective bargaining coverage — are rare; and identifying a source of variation in bargaining power presents an additional challenge.

I employ two very different estimation strategies to explore the predictions of the model developed in Section 1.2. Both approaches regress measures of financial debt on proxies for union bargaining power and exploit the interaction between union bargaining power and profit variability to specifically identify the strategic effect.[25] The first approach uses various cross-sections of primarily manufacturing firms from the 1970s, 1980s, and 1990s to establish the correlation of collective bargaining coverage, debt, and profit variability at the firm level. The latter approach uses changes in state labor laws in the 1950s, 1960s, and early 1970s to identify how changes in bargaining power affect debt levels at panels of firms in highly unionized industries.

### 1.3.1 Cross-sectional empirical approach

The degree of union bargaining power in negotiations with a given firm likely increases with the proportion of the firm's employees covered by collective bargaining. At firms with greater coverage, union-organized job actions are more costly, and firm-wide policies are more likely to be affected by bargaining. I use firm-level data on collective bargaining coverage as a proxy for union bargaining power and estimate its effect on the firm's choice of capital structure.

Firm-level estimates of collective bargaining coverage are not widely available. I obtained them from two different sources for cross-sections of firms in 1977, 1987, and 1999. Coverage estimates for 1977 and 1987 are derived primarily from a 1987 survey of manufacturing firms conducted by Barry Hirsch (1991).[26] The data for 1999 were compiled

[24]The existing evidence, described in Section 1.1, consists of cross-sectional correlations that may be affected by omitted variables bias. It also fails to distinguish between strategic and balance sheet effects.

[25]Cross-sectional correlations, presented in Appendix Table 1.A1, support notions of profit variability generating excess liquidity, which is partially captured by workers during collective bargaining. Based on the universe of Compustat firms in 1983, a firm with one standard deviation greater profit variability has, on average, 11 cents greater profits per dollar of assets and pays an 80 percent greater union wage premium.

[26]In addition to data from his own survey, Hirsch augments the 1977 sample using firm coverage data

by Craig Eschuk (2001), mostly from company 10-K annual reports, and were provided to me by Richard Freeman. Each cross-section is supplemented with firm-level financial data from Compustat,[27] and summary statistics are presented in Table 1.1. The sample includes 656 firms in 1977, 368 firms in 1987, and 349 firms in 1999. The 10-K sample may provide less generalizable estimates, since non-unionized firms are not represented.[28] In each sample, a quarter to a third of employees are covered by collective bargaining at a representative manufacturing firm. For 1977 and 1987, these figures are similar to estimates derived from other sources.[29]

I employ regression analysis to examine the correlation of collective bargaining coverage, debt, and profit variability at the firm level. Let $DEBT_{ij}$ be a measure of financial debt at firm $i$ in industry $j$ and $VALUE_{ij}$ represent the market value of the firm.[30]

$$\begin{aligned}\frac{DEBT_{ij}}{VALUE_{ij}} &= \alpha_1 COVERAGE_i + \alpha_2 VARIABILITY_i \\ &\quad + \alpha_3 COVERAGE_i * VARIABILITY_i + X_i\beta + \omega_j + \varepsilon_{ij}\end{aligned}$$

The level of debt as a fraction of the firm's total value is modeled as a function of the proportion of employees covered by collective bargaining ($COVERAGE_i$), a measure of profit variability ($VARIABILITY_i$), the interaction of $COVERAGE_i$ and $VARIABILITY_i$, a set of financial controls $X_i$, and two-digit SIC industry fixed effects $\omega_j$. Profit variability is measured using the standard deviation of the change in earnings before depreciation and amortization, divided by total assets.[31] It is demeaned (with respect to the sample mean)

collected in an independent 1972 Conference Board Survey. See Hirsch (1991) for details.

[27] Estimates of firms' marginal tax rates before interest expense for 1987 and 1999 were provided by John R. Graham and are described in Graham, Lemmon, and Schallheim (1998). Compustat variables are winsorized at the 1% tails.

[28] Information on collective bargaining coverage is not uniformly reported in the 10-K. Only companies for which union relations are material to the firm tend to report the figure. Eschuk contacted non-reporting firms, in total collecting data for about 65 percent of the original sample of manufacturing firms (Eschuk 2001, p.112). As reported in Table 1.1, all firms for which data was collected had at least some collective bargaining coverage.

[29] Estimates from the Current Population Survey (CPS) suggest that approximately 35 percent of manufacturing workers were covered by collective bargaining in 1977, and 25 percent were covered in 1987 (Hirsch 1991). On the other hand, CPS estimates suggest that only 16.6 percent of manufacturing employees were covered by collective bargaining in 1999, and only 10.2 percent of private sector employees overall (Hirsch and Macpherson 2003). Selection in firm reporting of collective bargaining coverage in 10-Ks likely explains the greater rates present in the 1999 sample. See note 28.

[30] The empirical results are robust to using total assets (rather than market values). Results for both types of specifications are reported below.

[31] This measure of profit variability dates back to Brealey, Hodges, and Capron (1976), if not before, and is common in the finance literature. For a discussion of its relative merits, see Chaplinsky (1984). It

before it is interacted with union coverage to allow for a more meaningful interpretation of the $COVERAGE_i$ main effect. The financial controls are those typically included in leverage regressions, specifically the proportion of fixed assets (a proxy for potential collateral), the marginal tax rate before interest expense (nondebt tax shields), the market-to-book ratio (investment opportunities), log sales (firm size), modified Altman's z-score (probability of bankruptcy), and return on assets (profitability).[32] Table 1.1 summaries the financial variables in each of the sample years. On average, five percent of a firm's capitalization is debt due within one year, whereas total debt represents a quarter to a third. Similar equations with different dependent variables are used to investigate the relationship between collective bargaining coverage and other variables.

Theory ascribes a special role in influencing collective bargaining negotiations to any financial (or non-financial) instrument that places demands on current cash flows. Empirically, debt in current liabilities (what I have been calling "current debt") most closely reflects the cash flow demands of debt service payments. It includes both short-term debt and the current maturities of long-term debt.[33] Debt with noncurrent maturity may not require any current period cash expenditure if the interest is paid later, such as with zero coupon bonds. Because the vast majority of debt service payments reflect principal versus interest (approximately 75 percent in my 1950s sample), current debt provides a better measure than total debt of the near-term need for cash that financial leverage imposes on

---

is normalized by its standard deviation to ease the interpretation of the estimates. Measures of profit variability are meant to reflect the product market variability underlying each firm's business. Depreciation and amortization are added back to earnings, because they are noncash charges. The estimates are robust to alternative calculations of variability, including using earnings also before interest and taxes (EBITDA), using cash-basis rather than accrual-basis measures of earnings (i.e., operating cash flows), and using sales (to avoid any bias introduced by including labor costs). Regressions using these alternative measures in the labor law analysis framework are reported in Appendix Table 1.A3. Where available, data for up to the previous 10 years is included in the calculation of profit variability (e.g., 1967-1977 for 1977). Including more years of data does not change the results. An observation is dropped if fewer than 5 years of data are available.

[32] These variables reflect the literature on capital structure, surveyed in Harris and Raviv (1991). They are the variables included in cross-sectional analysis in Rajan and Zingales (1995) plus other variables the authors state they would have included but for lack of data availability across their broad set of countries. Modified Altman's z-score is

$$3.3\frac{\text{EBIT}}{\text{total assets}} + 1.0\frac{\text{sales}}{\text{total assets}} + 1.4\frac{\text{retained earnings}}{\text{total assets}} + 1.2\frac{\text{working capital}}{\text{total assets}}$$

(MacKie-Mason 1990). Note that while these financial variables are known to be correlated with total leverage ratios, there may be no reason to expect a correlation with measures of short-term debt. They also may be endogenous. For these and other reasons, some may worry that the financial variables actually distort measurement of the relationships between debt, union coverage, and the variability of profits. To address these concerns, I also estimate specifications without financial controls.

[33] The empirical results are robust to using the sum of debt in current liabilities and interest expense.

a firm (Plesko 2001). Accordingly, I focus on current debt as the primary debt instrument available to a firm seeking to strategically reduce liquidity. Yet in theory, any cash flow demands imposed by debt before the project reaches completion may influence bargaining, and in practice, even debt with a long maturity can impose cash flow demands through covenant restrictions. Therefore, it is likely that firms also boost noncurrent debt to influence collective bargaining. Where available, I examine debt due within 2-5 years and total leverage in addition to current debt.

While this sort of cross-sectional approach provides important evidence, it is limited. When measuring the effect of collective bargaining using comparisons across firms, there will always be a suspicion that the controls included in the analysis are not exhaustive. These concerns are mitigated by identification of the strategic effect coming from the interaction with profit variability. Nevertheless, if an omitted firm characteristic differentially affects both the degree of unionization and capital structure determination at firms with greater profit variability, then it would be inappropriate to assign the estimates a causal interpretation. I address this important concern with a second empirical approach.

### 1.3.2 Cross-sectional estimates

Cross-sectional evidence of the effect of unionization on current borrowing is presented in Table 1.2. Each panel corresponds to analysis conducted for a cross-section of firms in 1977, 1987, and 1999, respectively. Raw correlations at the firm level provide inconsistent evidence on the relationship between collective bargaining coverage and the use of current debt (Column 1).[34] However, further analyses show that this obfuscates the importance of the relationship at firms with highly variable profits (Columns 2-4). The interaction term is positive and both statistically and economically significant. Greater variability exposes relatively more liquidity to union capture at firms with greater union coverage, apparently leading these firms to increase current debt. The interaction provides a margin for detecting strategic responses to collective bargaining, even in many empirical models where the causal union main effect is not independently identified.

The magnitude of the interaction effect is also significant. Consider two firms — one

[34] While the estimated 1977 and 1999 coefficients are close to zero, it is difficult to interpret these estimates, because they may be affected by omitted variables bias. Because its identification comes from an additional contrast, the interaction effect is arguably less susceptible to omitted variables bias.

with one standard deviation greater profit variability than the other. An increase in a firm's union coverage by 10 percentage points is associated with approximately a 30 to 50 basis point greater increase in current debt for the more exposed firm. This effect is sizeable, measuring 5 to 8 percent of mean current debt among sample firms. Stated differently using the main effect estimate for 1987, for firms with profit variability one standard deviation above the mean, the ratio of current debt to firm value is 5 to 10 percent greater when an additional 10 percent of employees bargain collectively.[35] In contrast, for firms with little profit variability, differences in union coverage rates seem to have little effect.

Analyses of other measures of near-term debt yield comparable estimates. Because debt due within 2 to 5 years also affects near-term cash flows, it, too, may be used strategically by firms to affect collective bargaining. As with the analysis of current debt, I interpret positive estimates of the interaction of collective bargaining coverage and profit variability as evidence of firms' employing such a strategy. To account for the differing magnitudes of the various debt measures, I convert the interaction coefficients to "elasticities," which are reported in Table 1.3. (More precisely, these are derivatives, with respect to profit variability, of the union coverage elasticity of debt, evaluated at the mean.) Each elasticity is from a separate regression, where the rows correspond to different dependent variables and the columns represent the cross-sections analyzed. Collective bargaining seems to influence the maturity structure of debt in favor of current borrowing. Its impact is greatest on debt due within one year and declines with each broader classification of near-term debt.

Cross-sectional analysis also suggests that collective bargaining may increase total debt. As reported in Table 1.4, the variables are highly positively correlated (Column 1). Total debt is on average 60 to 160 basis points (2 to 6 percent) greater at firms where an additional 10 percent of employees bargain collectively. However, financial controls seem to account for much, if not all, of the correlation (Column 3), complicating the interpretation of the correlation between unionization and total debt.[36] From this evidence, it is difficult to discern to what extent the financial variables control for selection in union organizing versus represent a causal mechanism through which union presence affects capital structure. Regardless,

[35] For these firms, the marginal effect is 50 to 70 basis points, and the average ratio of current debt to firm value is about 7 percent.

[36] In fact, the main effect of union coverage on debt divided by assets is negative and statistically significant in the 1999 sample (Panel C, Column 4). Given the selected nature of the 1999 sample (see note 28), this estimate should be interpreted carefully. Regardless of the magnitude it reports, that a firm includes union coverage at all in its 10-K suggests that the union has significant bargaining power.

the cross-sectional interaction of collective bargaining and profit variability provides little support for a strategic effect on *total* borrowing (Columns 2-4).

Boosting the cash flow demands of capital structure is not the only tactic available to firms seeking to improve their bargaining position vis-à-vis workers. My empirical evidence of the strategic use of current debt relies heavily on the importance of profit variability in increasing expected excess liquidity and the degree of union rent seeking. Evidence of firms adopting other anti-union strategies specifically when profits are more variable provides a useful robustness check.[37] In search of such evidence, I examine inventory policy.

Ultimately, much of workers' bargaining power in collective bargaining negotiations is derived from credible threats to withhold labor services. To mitigate these threats, firms may strategically maintain costly "buffer" inventories, which increase the costs of a strike borne by workers relative to those borne by the firm (Christenson 1953).[38] Compared to materials and goods in earlier stages of the production process, inventories of finished goods provide the most effective insurance and deterrence against employee job actions. In fact, it would be surprising if inventories of raw materials had any correlation with worker-firm bargaining power.

Analyses of these inventory hypotheses are presented in Table 1.5, based on the cross-section of manufacturing firms in 1977. Unfortunately, inventory data for my later, smaller samples of firms are not sufficiently complete to provide meaningful analysis. In 1977, firms appear to have been using inventories of finished goods strategically in conjunction with collective bargaining negotiations. Although raw correlations suggest no relationship between collective bargaining and inventories (Column 1), the effect of the interaction between union coverage and profit variability suggests there is a strategic component in inventory policy — similar current debt.[39] Whereas profit variability is negatively corre-

[37]The correlations presented in Appendix Table 1.A1 provide an additional robustness check. See note 25.

[38]In a study of the determinants of U.S. strike activity, Tracy (1986) finds that *total* inventories have no effect on strike activity. However, as he argues and as is demonstrated below, total inventories, which include raw materials and work-in-progress in addition to finished goods, provide a relatively poor measure of buffer stock.

[39]As with the other cross-sectional models, the union main effect measured here should not be assigned a causal interpretation. Even if firms boost inventories in response to unionization, the variables may not be correlated if, for example, worker organization drives are more successful at firms with otherwise lower inventory levels. Threat effects on nonunionized firms may also prevent detection. These straightforward explanations are much less trouble for interpretations of the interaction effect. Nevertheless, noncausal explanations of the interaction effect cannot be unambiguously refuted using cross-sectional estimates.

lated with inventories, perhaps because these firms are more liquidity constrained (Gertler and Gilchrist 1994), the opposite is true at majority-unionized firms (Columns 2-3). As developed in Section 1.2, the scope for union rent seeking increases with the interaction between unionization and profit variability. Regression analysis finds evidence of a sizeable strategic effect: given two sample firms, one with one standard deviation greater profit variability than the other, an increase in union coverage by 10 percentage points is associated with a greater increase in total inventories for the more exposed firm by 41 basis points (as a percent of sales), or 2.2 percent of mean inventories. As expected, the effect appears to be driven most by a build-up of finished goods (3.4 percent) and, to a lesser extent, increased work-in-progress inventories (2.8 percent). The estimates find no relationship with inventories of raw materials. Although these results are persuasive, corroboratory evidence from the labor law analysis would be ideal. Unfortunately, data breaking down inventories by stage-of-production is not readily available for the period of that analysis, and evidence based on total inventories is imprecise.

## 1.4 Exploiting state changes in labor law

The second set of empirical analyses uses state-specific changes in labor law to identify changes in union bargaining power. Over time, state policymakers have used legislation and public subsidies to influence the costs of union organizing and activism, altering workers' relative bargaining position. Two important policies in this context are right-to-work laws and unemployment insurance work stoppage provisions.

### 1.4.1 Right-to-work laws

Federal collective bargaining law was established by the National Labor Relations Act (the Wagner Act) in 1935. It set up the National Labor Relations Board (NLRB), an independent federal agency, which administers union elections and ensures that a union represents its constituent employees. Once a union is certified by the NLRB, the employer is required to bargain with the union in good faith. By preventing employers from discriminating against workers who join unions or participate in a strike, labor law confers significant market power to certified unions. The Wagner Act also allowed the parties to agree (through bargaining) to require anyone hired to join and financially support the union.

However, Republican Party gains in the 1946 mid-term Congressional elections and strong anti-labor sentiment following World War II, resulted in the Labor-Management Relations Act (the Taft-Hartley Act), which was passed over President Harry Truman's veto in 1947. Among other provisions which were broadly construed as anti-union, the Taft-Hartley Act granted states the power to pass so-called "right-to-work" (RTW) laws. RTW laws outlaw employment contract provisions that require employees to join or financially support a union. As such, the laws expose unions to a free rider problem, whereby non-union employees benefit from collective bargaining without paying dues. Figure 1-3, Panel A, describes the history of state RTW legislation.

I use state adoption of RTW laws in the 1950s as a source of geographic changes in union bargaining power.[40] As surveyed by Moore (1998), the existing empirical evidence indicates that RTW laws impact union organizing activity and industrial development. By examining flows into unionization, Ellwood and Fine (1987) demonstrate that RTW laws have a sizeable impact on union organizing. A similar analysis is presented in Appendix Table 1.A2. Using state-level data, I regress the log number of members of newly elected bargaining units on a RTW indicator, a pre-adoption indicator, and various controls. While organizing activity was comparable across states before adoption of RTW laws, it decreased by approximately 30 percent in adopting states after the laws were passed. These estimates are consistent with RTW laws ultimately diminishing union membership by 5 to 10 percent, or 1 to 3 percentage points (Ellwood and Fine 1987, p.266).[41] RTW laws may also encourage industrial development; Holmes (1998) uses comparisons across state borders to show that relative manufacturing employment is about one-third greater in states with RTW laws than in other states.

In addition to reducing the threat of new union organizing, RTW laws likely directly affect collective bargaining at firms with existing unions. Ellwood and Fine (1987, p.270) argue:

> The most obvious explanation is simply that passage of an RTW law makes union membership less economically attractive to workers. Without the ability to enforce payment of dues or to fine those who cross the picket line, unions

[40]Limitations in the availability of Compustat data restricts my analysis to this period.

[41]These decreases in union membership are in addition to the losses that might occur if any members of existing bargaining units choose to not be union members when union shop rules are eliminated. Unfortunately, the dearth of comparable state-level data on the stock of union membership before and after the passage of RTW laws interferes with estimation of these losses. For a detailed discussion of data availablity, see Ellwood and Fine (1987, p.253-4).

> may prove less powerful. Their strike threats are diminished both by reduced financial resources and by less certain participation.

As a symbol of union defeat, the passage of RTW laws may also have a psychological effect on a union's appeal to workers (Ellwood and Fine 1987). Both economic and psychological channels weaken the union's bargaining position, thereby reducing the expected benefits of union membership, the marginal benefit of organization, and the supply of union jobs (Farber 1984).[42]

### 1.4.2 Unemployment insurance work stoppage provisions

> *I believe that the work stoppage portion of the labor dispute disqualification provision is the most significant [feature of the unemployment insurance system] in affecting behavior of the parties in collective bargaining or industrial relations.*
>
> Sandra D. Dragon, Commissioner
> Vermont Department of Employment and Training[43]

Another plausibly exogenous source of variation in union bargaining power comes from changes in the unemployment insurance system. The United States unemployment insurance system invests states with considerable autonomy to establish rules governing claimant eligibility for benefits. State autonomy results in considerable variation across states and over time in the conditions under which workers unemployed because of a labor dispute qualify for unemployment compensation. While eligibility rules generally exclude striking workers, a majority of states allow those unemployed because of a labor dispute to collect unemployment insurance benefits under specific (but not usual) conditions.

I focus on one such eligibility rule that has been shown empirically to be of particular importance: the work stoppage provision (WSP). In 1960, 35 states permitted strikers to collect unemployment benefits during a labor dispute if their employer continued to operate at or near normal levels. In these states, an eligible striker could collect benefits after the normal waiting period (generally one week after filing for benefits). In a sense, a WSP provided strikers with insurance for a failed strike. Striking workers could collect benefits only if employers succeeded in weathering the strike and continued to operate at or near normal levels.

[42] Theoretically, the effect of RTW laws on wages is ambiguous. While eroded union bargaining power may decrease wages, enhanced industrial development may boost both labor demand and wages. Empirical evidence on RTW wage effects varies widely (Farber 1984; Reed 2003).

[43] Letter, dated December 28, 1981, quoted in Hutchens, Lipsky, and Stern (1992, p.340).

WSPs have been shown to affect collective bargaining. Unions accrue bargaining power in negotiations from an implicit, if not explicit, threat to withhold labor services. While a variety of theories explain strike activity, it is relatively uncontroversial that workers' bargaining position is improved when striking is less costly (Kennan 1986). One theory of strike activity predicts that strikes are a decreasing function of the combined cost borne by workers and management (Reder and Neumann 1987). Because unemployment insurance premiums are only imperfectly experience-rated, joint cost theory predicts that paying benefits to strikers not only improves their bargaining position, but also increases strike activity. In an analysis of the influence of various government transfer programs on strikes, Hutchens, Lipsky, and Stern (1989) find that the repeal of unemployment insurance WSPs is associated with less frequent strike activity in states with relatively generous unemployment insurance programs.[44] Figure 1-3, Panel B, depicts the history of state WSPs. Seven states repealed WSPs between 1960 and 1973.[45] I use this legal variation to identify changes in union bargaining power.

### 1.4.3 Labor law empirical approach

I estimate the reduced form effect of RTW laws and WSPs on firms located in the affected states, and interpret the results as indicating the effects of changes in union bargaining power. To focus on sectors where these laws are most relevant, I restrict the sample to industries known to have relatively high union coverage. Included industries are listed in Appendix Table 1.A3.[46] Let $LAW_{st}$ indicate the presence of a RTW law or absence of a WSP in state $s$ at time $t$.

$$\frac{DEBT_{ijst}}{VALUE_{ijst}} = \alpha_1' LAW_{st} + \alpha_2' LAW_{st} * VARIABILITY_i + \eta_i + \tau_{jt} + \xi_{ijst} \tag{1.7}$$

[44]Hutchens, Lipsky, and Stern (1989) also find that innocent bystander rules increase strike activity in generous states. Changes in these rules are highly correlated with the repeal of WSPs, but they are less frequent. The authors find no evidence of a link between strike activity and other unemployment insurance rule changes, including those related to lockouts and interim employment, AFDC, food stamps, or general assistance.

[45]The empirical estimates are robust to excluding New Jersey, which readopted a WSP shortly after repealing it in 1967.

[46]The sample includes two-digit SIC industries with greater than 25 percent union coverage in 1983 as measured by the CPS. While selecting the sample in this regard limits us from generalizing the results to other firms, it does not bias estimates of strategic behavior *within* the sample. In fact, similar analyses of firms in scarcely unionized industries will provide a falsification test.

The specification includes firm fixed effects $\eta_i$, as well as industry-by-year fixed effects $\tau_{jt}$.[47] Standard errors are clustered at the state level, allowing for unspecified within-state correlation over time.

Firms set capital structure based on a number of factors, many of which are at least partially unobservable. For example, firms expecting better future investment opportunities likely use less leverage. As long as these and any other unobservables which comprise $\xi_{ijst}$ are not correlated with legislative changes to RTW laws or WSPs, the estimates of $\alpha_1'$ and $\alpha_2'$ in equation (1.7) have causal interpretations. This assumption of uncorrelation is the principal identification assumption of this approach.

Labor law almost certainly responds to economic conditions and trends in industrial relations, some of which may be correlated with the use of debt. Some of these factors may induce omitted variables bias in estimates of the effect of RTW laws and WSPs.[48] At the same time, a number of factors support a causal interpretation. The states changing RTW laws in the 1950s and WSPs in the 1960s are not restricted to any particular geographic region (Figure 1-3). Furthermore, analyses of pre-existing trends show that decreases in union organizing do not precede the passage of RTW laws (Ellwood and Fine 1987; Appendix Table A2), and the decreases in debt levels are greatest after changes to RTW laws and WSPs (see Section 1.4.4).

I use both the adoption of RTW laws during 1950-1960 and the repeal of WSPs during 1960-1973 to identify decreases in union bargaining power. Each has its strengths and weaknesses for this research. As compared to WSPs, RTW laws likely have greater impact on union power and industrial relations. RTW laws weaken unions both financially and organizationally. In addition to diminishing the financial resources available to unions, RTW laws reduce strike threats by rendering unions less able to discipline those who cross

[47] Firm fixed effects ensure that changes in debt levels are estimated from a consistent sample of firms. Since Compustat data is not available before 1950 and the sample size grows rapidly in the early 1960s, measures of profit variability are computed over each contemporaneous period, 1950-1960 and 1960-1973, and do not vary over time for a given firm. (The main profit variability effects are absorbed by the firm fixed effects.) Firms are assigned to a state and an industry based on Compustat header information relating to the company's last reported location and industry of primary operation. To the extent that firms may have moved locations or changed industries since the 1950s, these may represent noisy measures of the historical variables. While point estimates may be less precise, there is no obvious reason why this mismeasurement would bias the results. Manual checks against historical 10-Ks also suggest the location and industry information is generally accurate.

[48] If anything, states are more likely to repeal a WSP following a protracted, high-profile labor dispute. As firms affected by a work stoppage are more likely to be in financial distress, this proposed relationship works against finding that the repeal of WSPs are associated with decreases in leverage.

a picket line. Due to their potent impact on organizing, RTW laws also significantly impact the threat of new unionization at both partially-unionized and nonunionized firms. While the repeal of WSPs also diminishes strike threats, the magnitude of the effect on bargaining power is likely relatively modest.

On the other hand, the sample of firms in the WSP analysis is much larger than the RTW sample on two dimensions, enabling more precise estimation. Key features of each sample are presented in Table 1.6, Panel B. First, Compustat has far greater firm coverage in the later period.[49] The WSP analysis is based on almost four times as many firms (1,273) as the RTW analysis (326). Second, while seven states changed laws in each period, the states that changed WSPs happened to be larger than those that changed RTW laws (in the respective sample periods). Whereas only 2.3 percent of observations in the 1950-1960 sample are in states adopting RTW laws, 21.4 percent of observations in the 1960-1973 sample are in states repealing WSPs. For these reasons, I rely on both sources of variation in union bargaining power to present evidence on the strategic use of debt.

Table 1.6 also presents summary statistics for several relevant firm financial variables. The sample of firms in the WSP analysis (1960-1973) has similar average debt levels to the cross-sections of firms analyzed in Section 1.3 (1977, 1987, 1999), while the RTW firms (1950-1960) have slightly lower levels of average total debt. While the firms in the earlier periods are smaller on average (in terms of total assets and sales), a greater percentage of their assets are fixed and they tend to earn a greater return.

It is tempting to interpret the changes in labor laws as an instrument for collective bargaining coverage. Such a calculation is infeasible without firm-level (or even state-level) unionization data for the earlier periods. Regardless, assigning such an interpretation is conceptually unjustified. RTW laws and WSPs affect union bargaining power through channels other than just the percentage of employees covered. As explained above, the laws have both financial and psychological consequences in addition to any effects on participation. As such, the laws provide a proxy for union bargaining power, but not an

[49] Unlike in more recent years, the Compustat database does not include all firms with SIC filings in the period of these analyses. Screening for nonmissing total assets, Compustat includes 626 firms in 1950, 1000 in 1959, 1619 in 1960, and 4522 in 1973. According to information provided by Standard & Poor's, the product's vendor, the 1950s sample primarily includes companies in the S&P 425, and the 1960s sample also includes firms listed in the NYSE and ASE. The sample for the RTW analysis includes firms in 30 states, three of which adopted RTW laws in the 1950s (Indiana, Kansas, and Utah). The sample for the WSP analysis includes firms in all 50 states, seven of which repealed WSPs during the sample period (see Figure 1-3, Panel B).

instrument for collective bargaining coverage. I estimate reduced form effects and interpret the results as effects of changes in union bargaining power.

### 1.4.4 Labor law estimates

Using state changes in right-to-work laws and unemployment insurance work stoppage provisions to identify changes in bargaining power confirms that incentives from union bargaining have a substantial impact on capital structure determination. Figures 1-4 and 1-5 present a graphical overview of these results. Focusing on firms located in states adopting RTW laws during 1950-1960 (Figure 1-4) or repealing WSPs during 1960-1973 (Figure 1-5), I graph both current and total debt in the four years before and after the law was changed in their states. Each panel presents the graph for a different sample of firms. Figure 1-4, Panel A, shows that while there is apparently no pre-existing trend, average debt levels at firms in densely unionized industries decrease after a RTW law is adopted. In contrast, Figure 1-4, Panel B, shows that debt levels do not decrease in less unionized industries.[50] Both panels in Figure 1-5 focus on firms in densely unionized industries, but Panel A includes firms with profit variability in the top quartile, whereas Panel B includes firms in the bottom quartile. While there is no apparent pre-existing trend, average debt levels decrease at firms with relatively variable profits after a WSP is repealed (Panel A). Yet the repeal of a WSP does not seem to affect the capital structure of firms with less variable profits (Panel B). The differential impact at firms with the excess cash flow exposure associated with greater profit variability suggests that these changes in labor law, which erode union bargaining power, apparently diminish a firm's strategic incentive to carry debt. Comparing magnitudes across figures, the effect appears to be greater in magnitude for RTW laws than WSPs (as expected). Although they are strongly suggestive, these figures merely present unconditional means. They do not control for macroeconomic year effects or industry-wide trends. For those tests, I turn to the multivariate regression analysis described in Section 1.4.3.

Table 1.7 presents the effect of collective bargaining on current debt. Evidence from changes in RTW laws is presented in Panel A, and WSP evidence is in Panel B. While the

[50] The graph suggests total debt levels in these industries may actually increase after a RTW law is adopted. Regression analysis that controls for contemporaneous changes in these industries in other states finds this apparent effect is not statistically significant (Table 1.9, Panel III, Column 2).

direction of the effects of the laws on debt is the same, the magnitude of RTW law effects is substantially greater (as expected). While not statistically significant, the point estimates measuring the main effect of RTW laws on current debt are sizeable — approximately equal to the sample mean. The WSP main effect is much smaller in magnitude; while the estimates are close to zero, modest effects (on the order of 20 percent of the sample mean) cannot be ruled out. The econometric tests likely have insufficient power to identify the effect, as only a handful of states modify each law during the sample periods. Although I suspect additional precision would reveal an effect, there may in fact be no main effect. Such an interpretation suggests that unionization or amicable labor relations have positive externalities on profits — such as the productivity effects proposed by Brown and Medoff (1978) — in addition to the negative balance sheet effects described by the theory. While the main effects of these laws on debt are interesting, they do not specifically address the *strategic* use of debt to influence collective bargaining.

Evidence on the strategic use of debt is presented in the differential impact of the laws on firms with more variable profits. Since greater profit variability implies that more liquidity is subject to union capture, these firms have the greatest incentive to alter the maturity structure of their debt to reduce the union's ability to expropriate quasi-rents. As found with the cross-sectional analyses, union bargaining power has a large and statistically significant differential effect on firms with highly variable profits. The inclusion of both firm and industry-by-year fixed effects ensures that the estimates measure within-industry-year comparisons of within-firm changes in debt across states with different legislative patterns. The estimated effects are robust to including financial controls (the proportion of fixed assets, log sales, and the return on assets; Column 2) and measuring debt as a proportion of total assets (rather than the market value of the firm; Column 3).

By decreasing labor's bargaining power, RTW laws and WSPs seem to lead firms with more variable profits to strategically employ less current debt. Consider two firms — one with profit variability one standard deviation greater than the other. As compared to the firm with less profit variability, the more exposed firm decreases the ratio of current debt to firm value by approximately 5 more percentage points after a RTW law is passed and by approximately 1 more percentage point after a WSP is repealed. Both effects are sizeable, but as expected, the WSP effect is more modest: for firms with profit variability one standard deviation above the mean, the ratio of current debt to total firm value decreases by up to two-thirds after a RTW law is adopted, and by one-fifth after a WSP is repealed.[51] In contrast, these changes in labor law have little effect on firms with little profit variability.

Labor law evidence suggests that union power also leads firms to increase longer-term debt.[52] RTW estimates definitively suggest that collective bargaining increases the total leverage employed by affected firms. As reported in Table 1.8, Panel A, firms in states adopting RTW laws decrease their total leverage by 11 percentage points — over half of the sample mean. Including financial controls and normalizing debt using total assets, rather than the market value of the firm, have little effect on the estimates (Columns 2 and 3). Table 1.8, Panel B, presents estimates of the effect of WSPs on total debt. Similar to the effect on current debt, the WSP main effect is much smaller in magnitude and close to zero.

The differential effect of WSPs on firms with more variable profits provides evidence of the strategic use of total debt to influence collective bargaining negotiations. As before, consider two firms — one with profit variability one standard deviation greater than the other. When a WSP is repealed, the more exposed firm decreases current debt by approximately 3 percentage points more than the less exposed firm — approximately 10 percent of the sample mean. The RTW interaction coefficients are slightly greater in magnitude (relative to average debt levels), but they are not statistically significant. In this sample, the approximately 3 percentage point differential effect between firms with different profit variabilities (by one standard deviation) corresponds to approximately 15 percent of the sample mean. While the point estimates are comparable to the WSP estimates, the standard errors of the RTW estimates are much greater due to the limited size of both the sample and the states adopting RTW laws during the sample period.

Although additional precision may reveal an effect, there may be suspicion that the observed effect of RTW laws on total leverage represents only a balance sheet effect of unionization and *not* a strategic effect. After all, cash flow is a well-known predictor of debt levels in that firms tend to pay down debt when cash flow increases. Unionization, therefore, "mechanically" increases debt by reducing profits. The robustness of the RTW main effect to adding financial controls addresses this concern. In improving the firm's

[51]For these firms, the marginal effect of RTW adoption (WSP repeal) is 9 to 10 (1.5 to 2.0) percentage points, and the average ratio of current debt to firm value is about 14 (7) percent.

[52]The labor law results support the causal interpretation of the cross-sectional unconditional correlations reported in Table 1.4, Column 1. That is, they suggest that unions organize selectively (for example, at more profitable firms) generating a cross-sectional correlation between unionization and the financial control variables that seems to obfuscate a direct relationship between unionization and leverage.

bargaining position with labor, RTW laws should increase profits, improving the balance sheet of the firm, enabling additional continuation, and leading to a decrease in leverage. However, the estimates are robust to financial controls (Column 2), including return on assets — a proxy for profitability.[53] Although the controls are imperfect, the fact that point estimates are unaffected strongly suggests the primary mechanism is actually the strategic effect.

These analyses of both current and total debt support the notion that firms use debt policy strategically to affect bargaining with suppliers. Recognizing that changing labor laws should have a greater impact on firms subject to greater union rent seeking motivates analyzing the profit variability interaction. Similar reasoning suggests that adopting a RTW law or repealing a WSP will have a stronger effect on firms with greater operating income. As a robustness check, I regress the various financial policy measures on a RTW or WSP indicator variable, an interaction of the law indicator and the firm's average (pre-period) operating income, and a set of controls.[54] Operating income is before interest expense, payment of current debt maturities, taxes, depreciation, and amortization, is divided by lagged assets, and is normalized by its standard deviation to ease the interpretation of the estimates.[55] The results, reported in Table 1.9, Panel I, support the strategic use of capital structure. Evidence from changes in RTW laws is presented on the left side of the table, and WSP evidence is on the right. Following the adoption of a RTW law, a standard deviation increase in operating income is associated with a 5 to 6 percentage point greater decrease in both current and total debt ($p < 0.01$). As expected, WSPs have a more moderate effect. A standard deviation increase in operating income is associated with a 1 to 2 percentage point greater decrease in both current and total debt following the repeal of a WSP ($p < 0.01$).[56]

[53] Coefficient estimates for the controls are not reported to conserve space. The financial controls – the proportion of fixed assets, log sales, and ROA – are statistically significant with the expected signs ($p < 0.01$).

[54] The operating income main effect is absorbed by a firm fixed effect. Controls include firm fixed effects, industy-by-year fixed effects, and financial controls: the proportion of fixed assets, log sales, and ROA. The industry fixed effects are at the two-digit SIC level.

[55] This measure would ideally also include the union rent (the portion of wages paid in excess of the workers' alternative wages). Unfortunately, such data is not available. The results are robust to including total labor related expenses, which reduces the sample almost by half. Although neither measure is perfect, they provide a useful robustness check.

[56] Conceptually, inserting an average operating income interaction in the cross-sectional approach could provide a parallel robustness test. Such estimates are noisy. Furthermore, obvious simultaneity issues regarding collective bargaining coverage, operating income, and financial policy make any such estimates difficult to interpret.

Tests for pre-existing trends also support a causal interpretation of the observed labor law effects on debt. In general, the level of debt changes at firms affected by the laws only after the laws are adopted. Regression estimates are presented in Table 1.9, Panel II. In addition to an indicator variable for the presence of a RTW law or WSP (and the related interaction term), I include an indicator variable for the two years prior to the adoption of the law. While the pre-period coefficients are negative, they are generally small in magnitude and not statistically significant. Two exceptions are the pre-period interaction coefficients in both the RTW and WSP *total* debt regressions. In the RTW regression, the pre-period interaction coefficient is similar in magnitude to the interaction coefficient in the adoption period, but the estimates are extremely imprecise (Column 2). In the WSP regression, the pre-period interaction coefficient is statistically significant, suggesting presence of a pre-existing trend. However, the point estimate of the interaction coefficient in the adoption period is much greater in magnitude than that for the pre-period, suggesting WSPs have impact (Column 4). In all, tests for pre-existing trends support attributing the observed capital structure effects to the changes in labor law, especially for current debt.

A final robustness test is also possible. Each of the preceding analyses demonstrates that RTW laws and WSPs affect the financial policy of firms in industries with concentrated labor markets. Because only firms in industries with high union presence should be affected by these laws, estimating the effect of the laws on firms in industries with low union presence provides an important falsification test. Regression estimates comparable to those in Tables 7 and 8 are presented in Table 1.9, Panel III. Adopting a RTW law or WSP is not associated with a statistically significant change in levels of current or total debt for these firms.[57]

All in all, a variety of empirical evidence suggests that union bargaining power (and supplier market power more broadly) has far-reaching effects on firm financial strategy. The labor laws analyzed seem to have an acute impact on firms' current debt and total leverage. Cross-sectional evidence also suggests that firms use inventories strategically when engaged in collective bargaining with workers.[58]

[57]Although it is not statistically different than zero, the RTW-profit variability interaction point estimate for total debt is negative. Since RTW laws affect organizing, this may be because even some firms in industries with 25 percent union coverage are sensitive to the adoption of RTW laws. Consistent with this interpretation, the point estimate decreases in magnitude in regressions with lower "low union" cutoffs (not reported). WSPs are less likely to affect less-unionized firms and, in fact, point estimates for the WSP-profit variability interaction are positive, close to zero, and not statistically significant.

[58]Unfortunately, limited data availability interferes with the analysis of buffer inventory during the RTW and WSP periods. Inventory figures are not broken down by stage-of-production in Compustat for most

## 1.5 What about dividends?

Given its effect on optimal capital structure, it is natural to ask whether supplier market power also affects payout policy. Agency cost explanations of dividends argue that expected, continuing dividends discipline managers (Easterbrook 1984). In addition to boosting the degree of financial leverage, dividends compel managers to raise new money more frequently and thereby undergo more intensive monitoring. Consequently, stockholders often penalize firms making abrupt reductions in dividends (Lintner 1956).[59]

While in theory dividends may be used strategically to commit managers to be tough in union negotiations, such a role is naturally limited by the degree to which dividends can be affected by labor through bargaining. Bronars and Deere (1991) implicitly assume that dividends are fully subject to union capture, and DeAngelo and DeAngelo (1991) present case studies showing that dividend reductions are correlated with union negotiations. Using the empirical framework employed above to study strategic capital structure, I confirm that dividends appear not to be used as a commitment device vis-à-vis organized labor.

In a recent paper, Ramírez-Verdugo (2005) also analyzes payout policy in the presence of a unionized workforce. Whereas I focus on dividends as a commitment device, he analyzes their role as a signal. He argues that unions interfere with firms ability to use payout policy to convey earnings information to investors. My finding that union bargaining power decreases dividend payments is broadly consistent with this interpretation. It is also consistent with collective bargaining simply reducing profits and therefore also reducing payouts to shareholders.

Evidence on the cross-sectional relationship between collective bargaining and dividends is mixed. As shown in Table 1.10, Panel A, there is a strong correlation in 1977 (Column

firms in the 1950s and 1960s. Data on total inventories is generally available, but it is less representative of a firm's strategic use of inventories to insure and deter against union job action. Although regression estimates (not reported) are not sufficiently precise to be distinguished from zero, the signs of both the main RTW effect and its interaction with profit variability are consistent with strategic motivations. The point estimates suggest firms may decrease total inventories by 10 percent following RTW law adoption. Presumably, this masks a greater reduction in finished goods inventories. In contrast, the WSP point estimates are essentially zero.

[59] While stock repurchases perform a similar financial role to dividends, being irregular in nature they are less likely to serve as a commitment device. Evidence suggests that while repurchases have substituted for dividend payments over the last 10-15 years, regulatory constraints inhibited firms from aggressively repurchasing shares before 1983 (Grullon and Michaely 2002). Data on repurchases is not available in Compustat for the period of RTW analysis. Including the total value of repurchases in the cross-sectional regressions for the later periods provides qualitatively similar results.

1). In that year, manufacturing firms with an additional 10 percent of employees covered by collective bargaining paid, on average, almost 1 cent greater dividends (per dollar of total capitalization), corresponding to 3.5 percent of mean dividends among sample firms. However, this correlation is not robust to normalizing dividends using the book value of assets (not reported), nor is it present in the 1987 or 1999 samples (Columns 2 and 3). The correlation in 1977 may be explained by either the targeting of union organizing efforts on the most profitable firms, or as a "mechanical" effect whereby unionization decreases profitability and the market value of the firm. Firms with more variable profits seem to pay smaller dividends, likely to provide greater financial flexibility, but the interaction of union coverage and profit variability is not statistically significant (Columns 1-3).

The effect of RTW laws, presented in Panel B, suggests that collective bargaining may moderately reduce dividends. The main empirical specification suggests that RTW laws increase dividend payments by approximately 35 basis points — 9 percent of the sample mean (Column 1). However, the estimates are not precise enough to rule out the possibly that these results are driven by either the decrease in debt levels documented above (Column 3) or a pre-existing trend (not reported). These results suggest that omitted variables, such as an unobserved dimension of firm quality, bias upward the cross-sectional estimate for 1977 reported in Panel A. Estimates of the effect of WSPs on dividends are presented in Panel C. While not statistically significant, point estimates on the WSP main effect are also positive and smaller in magnitude than the RTW estimates.

At least two distinct factors may explain a negative effect of union bargaining power on dividends. First, collective bargaining may directly influence dividend payments. Whether negotiated explicitly or implicitly, case studies find "substantial and pervasive" dividend reductions clustered during union negotiations (DeAngelo and DeAngelo 1991). Collective bargaining may also dampen the use of dividends as an earnings signal (Ramirez-Verdugo 2005). Second, the dividend changes may simply be reflecting the financial position of the firm. For example, to the extent that RTW laws improve current or future earnings, managers likely feel compelled to share this financial windfall with shareholders (Lintner 1956). Either way, the empirical evidence is inconsistent with firms using *dividends* as a commitment device vis-à-vis workers.

## 1.6 Conclusion

In the past 20 years, many large unionized firms have filed for bankruptcy. Examples include at least eight major airlines and most recently auto-parts-maker Delphi. A natural question is whether collective bargaining led these firms to adopt a capital structure that made them more vulnerable to negative cash flow shocks. The results in this chapter suggest the answer to this question may be yes. As a supplier with market power, a union can demand a share of a firm's liquidity, which the firm maintains primarily to insure against negative shocks. To reduce the impact of collective bargaining on profits, the firm has the incentive to undertake costly actions that reduce its expropriable liquidity. Consequently, even efficient bargaining in the labor market terms yields outcomes that are not Pareto optimal.

While it is in the interest of both management and labor to produce institutional arrangements that lead to efficient contracts, this chapter demonstrates a dimension in which they come up short. If collective bargaining leads firms to distort their capital structure, then labor market outcomes will be inefficient even if employment levels are set optimally.[60] Greater than efficient levels of debt are also likely to distort product-market competition (Brander and Lewis 1986) and investment (Myers 1974). For example, debt buildup may explain part of the negative effect of collective bargaining on investment (demonstrated empirically by Connolly, Hirsch, and Hirschey 1986, and Fallick and Hassett 1999).

Previous studies have suggested that firms use debt to counter union power, but they have not (1) made a strong case for causality or (2) distinguished between strategic and balance sheet effects. This distinction is important; while the balance sheet effect is a response to the outcome of labor negotiations, the strategic effect is an attempt to influence these negotiations. In this chapter, I develop a testable implication for the *strategic* use of debt at unionized firms. An increase in the profit variability underlying a firm's operations

[60]The empirical literature on the efficiency of union contracting has focused predominantly on employment determination (see Farber 1986 for a survey). The implication of efficiency generally tested is whether employment levels are set optimally (for example, so as to equate the marginal revenue product of workers to their alternative wage). Using careful assumptions about union objective functions and data on membership and contract wages for the International Typographer's Union (ITU), MaCurdy and Pencaval (1986) find that employment levels are affected by bargaining. Nevertheless, using similar data but a different empirical approach, Brown and Ashenfelter (1986) conclude that the ITU labor agreements are not strongly efficient. Analyzing the debt structure of unionized firms expands this discussion by demonstrating another dimension in which collective bargaining may introduce inefficiency.

increases its expected excess liquidity, leaving a firm with more variable profits more vulnerable to union rent seeking. Evidence that collective bargaining more dramatically affects the capital structure of firms with greater profit variability suggests a strategic motive on the part of firms and investors.

Using right-to-work laws and unemployment insurance work stoppage provisions as sources of plausibly exogenous variation in union bargaining power, I find that collective bargaining increases total financial leverage and likely leads firms to reduce the maturity structure of debt. Furthermore, firms with relatively variable profits, and thereby a greater exposure to union rent seeking, respond with the greatest increases in current debt. In contrast, dividends unsurprisingly do not appear to serve as a commitment device vis-à-vis workers to protect liquidity from organized labor. These findings complement previous studies (Chevalier 1995a,b; Chevalier and Scharfstein 1996; Zingales 1998) that demonstrate the real effects of apparently exogenous changes in capital structure. The evidence presented in this paper suggests that these sort of real-side strategic incentives have a substantial impact on financing decisions.

# References


Baldwin, C. Y. (1983). Productivity and labor unions: An application of the theory of self-enforcing contracts. *Journal of Business 56*(2), 155–1985.

Belman, D. (1992). Unions, the quality of labor relations, and firm performance. In L. Mishel and P. B. Voos (Eds.), *Unions and Economic Competitiveness*, pp. 41–107. Armonk, NY: Sharpe.

Blanchard, O. and J. Tirole (2004). The optimal design of unemployment insurance and employment protection: A first pass. MIT Department of Economics Working Paper 04-15.

Brander, J. A. and T. R. Lewis (1986). Oligopoly and financial structure: The limited liability effect. *American Economic Review 76*(5), 956–70.

Brealey, R. A., S. D. Hodges, and D. Capron (1976). The return on alternative sources of finance. *Review of Economics and Statistics 58*(4), 469–77.

Bronars, S. G. and D. R. Deere (1991). The threat of unionization, the use of debt, and the preservation of shareholder wealth. *Quarterly Journal of Economics 106*(1), 231–54.

Brown, C. and J. Medoff (1978). Trade unions in the production process. *Journal of Political Economy 86*(3), 355–78.

Brown, D. T., C. E. Fee, and S. E. Thomas (2005). Financial leverage and bargaining power with suppliers: Evidence from leveraged buyouts. University of Pittsburgh mimeo.

Brown, J. N. and O. Ashenfelter (1986). Testing the efficiency of employment contracts. *Journal of Political Economy 94*(3, Part 2), S40–S87.

Campello, M. (2003). Capital structure and product markets interactions: Evidence from business cycles. *Journal of Financial Economics 68*(3), 353–78.

Campello, M. and Z. Fluck (2004). Product market performance, switching costs, and liquidation values: The real effects of financial leverage. University of Illinois mimeo.

Cavanaugh, J. K. and J. Garen (1997). Asset specificity, unionization and the firm's use of debt. *Managerial and Decision Economics 18*(3), 255–69.

Chaplinsky, S. (1983). *The Economic Determinants of Leverage: Theories and Evidence.* Ph. D. thesis, University of Chicago.

Chevalier, J. A. (1995a). Capital structure and product-market competition: Empirical evidence from the supermarket industry. *American Economic Review 85*(3), 415–35.

Chevalier, J. A. (1995b). Do LBO supermarkets charge more? An empirical analysis of the effects of LBOs on supermarket pricing. *Journal of Finance 50*(4), 1095–1112.

Chevalier, J. A. and D. S. Scharfstein (1996). Capital-market imperfections and countercyclical markups: Theory and evidence. *American Economic Review 86*(4), 703–25.

Christenson, C. L. (1953). The theory of the offset factor: The impact of labor disputes upon coal production. *American Economic Review 43*(4), 513–47.

Connolly, R. A., B. T. Hirsch, and M. Hirschey (1986). Union rent seeking, intangible capital, and market value of the firm. *Review of Economics and Statistics 68*(4), 567–77.

Dasgupta, S. and K. Sengupta (1993). Sunk investment, bargaining and choice of capital structure. *International Economic Review 34*(1), 203–20.

DeAngelo, H. and L. DeAngelo (1991). Union negotiations and corporate policy: A study of labor concessions in the domestic steel industry during the 1980s. *Journal of Financial Economics 30*(1), 3–43.

DiNardo, J. and D. S. Lee (2004). Economic impacts of new unionization on private sector employers: 1984-2001. *Quarterly Journal of Economics 119*(4).

Easterbrook, F. H. (1984). Two agency-cost explanations of dividends. *American Economic Review 74*(4), 650–59.

Ellwood, D. T. and G. Fine (1987). The impact of right-to-work laws on union organizing. *Journal of Political Economy 95*(2), 250–73.

Eschuk, C. A. (2001). *Unions and Firm Behavior: Profits, Investment, and Share Prices.* Ph. D. thesis, University of Notre Dame.

Fallick, B. C. and K. A. Hassett (1999). Investment and union certification. *Journal of Labor Economics 17*(3), 570–82.

Farber, H. S. (1984). Right-to-work laws and the extent of unionization. *Journal of Labor Economics 2*(3), 319–52.

Farber, H. S. (1986). The analysis of union behavior. In O. Ashenfelter and R. Layard (Eds.), *The Handbook of Labor Economics*, Volume 2 of *Handbooks in Economics*, pp. 1039–89. New York: Elsevier Science.

Fazzari, S. M., R. G. Hubbard, and B. C. Petersen (1988). Financing constraints and corporate investment. *Brookings Papers on Economic Activity 0*(1), 141–95.

Franck, T. and N. Huyghebaert (2004). On the interactions between capital structure and product markets: A survey of the literature. *Tijdschrift voor Economie en Management 49*(4), 727–87.

Freeman, R. (1992). Is declining unionization of the U.S. good, bad, or irrelevant? In L. Mishel and P. B. Voos (Eds.), *Unions and Economic Competitiveness*, pp. 143–69. Armonk, NY: Sharpe.

Freeman, R. B. (1986). The effect of the union wage differential on management opposition and union organizing success. *American Economic Review 76*(2), 92–96.

Freeman, R. B. and M. M. Kleiner (1999). Do unions make enterprises insolvent? *Industrial and Labor Relations Review 52*(4), 510–27.

Gertler, M. and S. Gilchrist (1994). Monetary policy, business cycles, and the behavior of small manufacturing firms. *Quarterly Journal of Economics 109*(2), 309–40.

Gorton, G. and F. Schmid (2004). Capital, labor, and the firm: A study of German codetermination. *Journal of the European Economic Association 2*(5), 863–905.

Graham, J. R., M. L. Lemmon, and J. S. Schallheim (1998). Debt, leases, taxes, and the endogeneity of corporate tax status. *Journal of Finance 53*(1), 131–62.

Grout, P. A. (1984). Investment and wages in the absence of binding contracts: A nash bargining approach. *Econometrica 52*(2), 449–60.

Grullon, G. and R. Michaely (2002). Dividends, share repurchases, and the substitution hypothesis. *Journal of Finance 57*(4), 1649–84.

Hanka, G. (1998). Debt and the terms of employment. *Journal of Financial Economics 48*(3), 245–82.

Harris, M. and A. Raviv (1991). The theory of capital structure. *Journal of Finance 46*(1), 297–355.

Hirsch, B. T. (1991). *Labor Unions and the Economic Performance of Firms.* Kalamazoo, Michigan: W. E. Upjohn Institute for Employment Research.

Hirsch, B. T. and D. A. Macpherson (2003). Union membership and coverage database from the current population survey: Note. *Industrial and Labor Relations Review 56*(2), 349–54.

Holmes, T. J. (1998). The effect of state policies on the location of manufacturing: Evidence from state borders. *Journal of Political Economy 106*(4), 667–705.

Hölmstrom, B. and J. Tirole (1996). Modeling aggregate liquidity. *American Economic Review 86*(2), 187–91.

Hutchens, R., D. Lipsky, and R. Stern (1989). *Strikers and Subsidies: The Influence of Government Transfer Programs on Strike Activity.* Kalamazoo, Michigan: W. E. Upjohn Institute for Employment Research.

Ippolito, R. A. (1985). The economic function of underfunded pension plans. *Journal of Law and Economics 28*(3), 611–51.

Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review 76*(2), 323–29.

Kale, J. R. and H. Shahrur (2004). Capital structure and characteristics of supplier and customer markets. Georgia State University mimeo.

Kennan, J. (1986). The economics of strikes. In O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics*, Volume 2 of *Handbooks in Economics*, pp. 1091–1137. New York: Elsevier Science.

Kleiner, M. M., J. S. Leonard, and A. M. Pilarski (2002). How industrial relations affects plant performance: The case of commercial aircraft manufacturing. *Industrial and Labor Relations Review 55*(2), 195–218.

Kovenock, D. and G. M. Phillips (1997). Capital structure and product market behavior: An examination of plant exit and investment decisions. *Review of Financial Studies 10*(3), 767–803.

Krueger, A. B. and A. Mas (2004). Strikes, scabs, and tread separations: Labor strife and the production of defective Bridgestone/Firestone tires. *Journal of Political Economy 112*(2), 253–89.

Lamont, O. (1997). Cash flow and investment: Evidence from internal capital markets. *Journal of Finance 52*(1), 83–109.

Lewis, H. G. (1986). Union relative wage effects. In O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics*, Volume 2 of *Handbooks in Economics*, pp. 1139–81. New York: Elsevier Science.

Lintner, J. (1956). Distribution of incomes of corporations among dividends, retained earnings, and taxes. *American Economic Review 46*(2), 97–113.

MacKie-Mason, J. K. (1990). Do taxes affect corporate financing decisions? *Journal of Finance 45*(5), 1471–93.

MaCurdy, T. E. and J. H. Pencavel (1986). Testing between competing models of wage and employment determination in unionized markets. *Journal of Political Economy 94*(3, Part 2), S3–S39.

Mas, A. (2004). Labor unrest and the quality of production: Evidence from the construction equipment resale market. U.C. Berkeley mimeo.

Modigliani, F. and M. H. Miller (1958). The cost of capital, corporation finance, and the theory of investment. *American Economic Review 48*, 261–97.

Moore, W. J. (1998). The determinants and effects of right-to-work laws: A review of the recent literature. *Journal of Labor Research 19*(3), 445–69.

Müller, H. M. and F. Panunzi (2004). Tender offers and leverage. *Quarterly Journal of Economics 119*(4), 1217–48.

Myers, S. C. (1974). Interactions of corporate financing and investment decisions-implications for capital budgeting. *Journal of Finance 29*(1), 1–25.

Opler, T. C. and S. Titman (1994). Financial distress and corporate performance. *Journal of Finance 49*(3), 1015–40.

Pagano, M. and P. F. Volpin (2005). Managers, workers, and corporate control. *Journal of Finance 60*(2), 841–68.

Perez, E. (2004, October 29). Cross winds: How delta's cash cushion pushed it onto wrong course. *The Wall Street Journal*, A1.

Perotti, E. C. and K. E. Spier (1993). Capital structure as a bargaining tool: The role of leverage in contract renegotiation. *American Economic Review 83*(5), 1131–41.

Phillips, G. M. (1995). Increased debt and industry product markets: An empirical analysis. *Journal of Financial Economics 37*(2), 189–238.

Plesko, G. A. (2001). The role of short-term debt in capital structure. In J. Hines, James R. (Ed.), *Proceedings: Ninety-Third Annual Conference on Taxation, Santa Fe, New Mexico, November 9-11, 2000*, Washington, D.C., pp. 135–40. National Tax Association.

Rajan, R. G. and L. Zingales (1995). What do we know about capital structure? Some evidence from international data. *Journal of Finance 50*(5), 1421–60.

Ramirez-Verdugo, A. (2005). Dividend signaling and unions. MIT mimeo.

Rauh, J. D. (forthcoming). Investment and financing constraints: Evidence from the funding of corporate pension plans. *Journal of Finance*.

Reder, M. W. and G. R. Neumann (1987). Conflict and contract: The case of strikes. *Journal of Political Economy 88*(5), 867–86.

Reed, W. R. (2003). How right-to-work laws affect wages. *Journal of Labor Research 24*(4), 713–30.

Sarig, O. H. (1998). The effect of leverage on bargaining with a corporation. *Financial Review 33*(1), 1–16.

Skatun, J. D. F. (1997). Delegation, union leaders and capital allocation. *Labour 11*(2), 249–64.

Spiegel, Y. and D. F. Spulber (1994). The capital structure of a regulated firm. *RAND Journal of Economics 25*(3), 424–40.

Tirole, J. (2004). The theory of corporate finance. Unpublished manuscript.

Tracy, J. S. (1986). An investigation into the determinants of U.S. strike activity. *American Economic Review 76*(3), 423–36.

Ulph, A. (1989). The incentives to make commitments in wage bargains. *Review of Economic Studies 56*(3), 449–65.

Weiss, L. A. (1990). Bankruptcy resolution: Direct costs and violation of priority of claims. *Journal of Financial Economics 27*(2), 285–314.

Wilson, L. (2004). Hard debt, soft CEOs, and union rents. Oxford University mimeo.

Zingales, L. (1998). Survival of the fittest or the fattest? Exit and financing in the trucking industry. *Journal of Finance 53*(3), 905–38.

Figure 1-1: Model of optimal capital structure determination

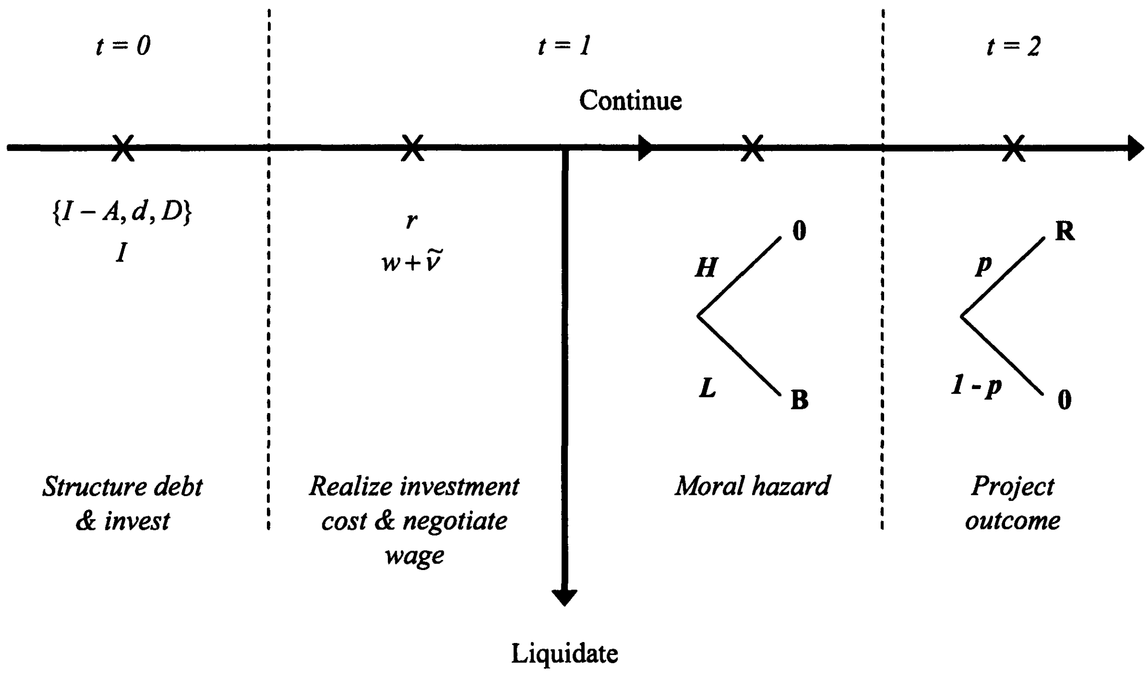



Figure 1-2: Impact of borrowing need on optimal capital structure determination

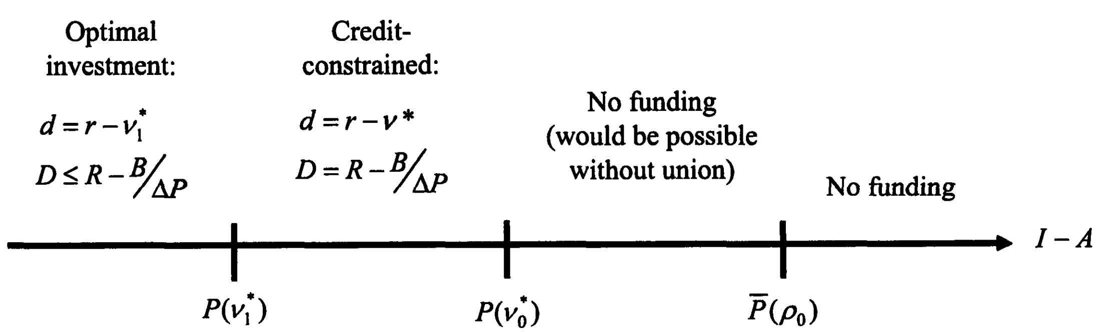
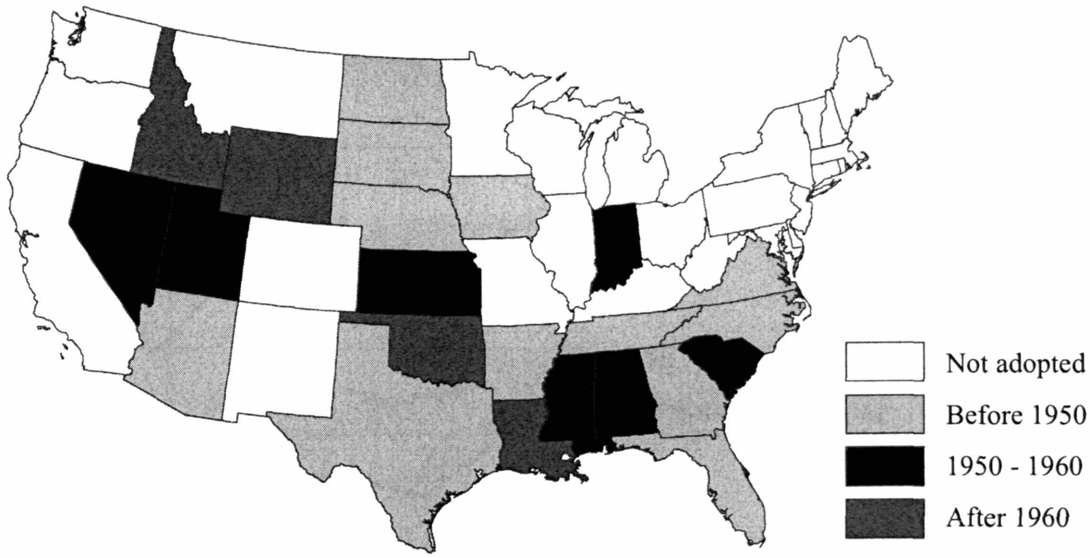
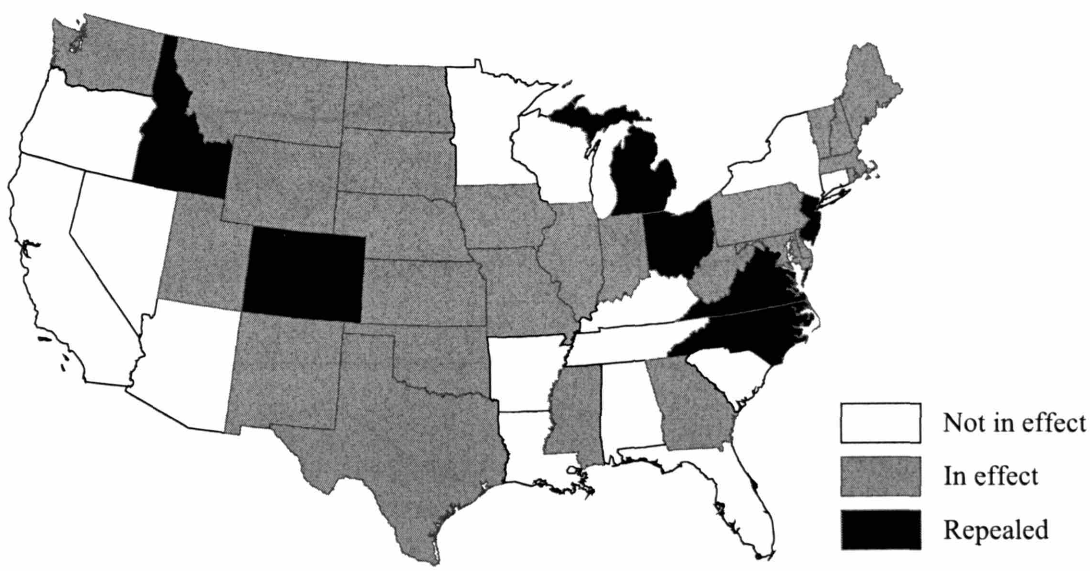


Figure 1-3: Legislative history of select state labor laws

A. Adoption of right-to-work laws, 1947-2005




B. Unemployment insurance work stoppage provisions, 1960-1973

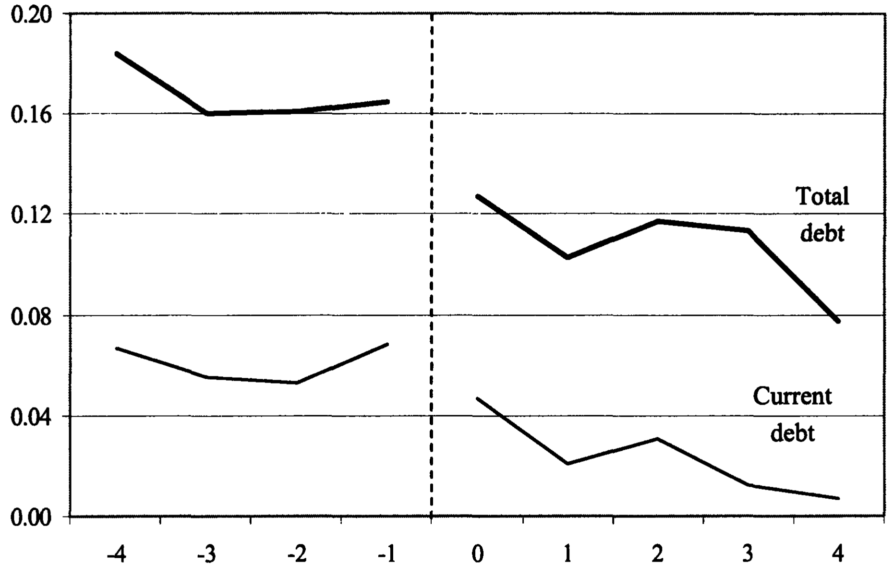



*Note*: The Indiana RTW law was later repealed, and the New Jersey WSP was re-adopted soon after it was abolished.

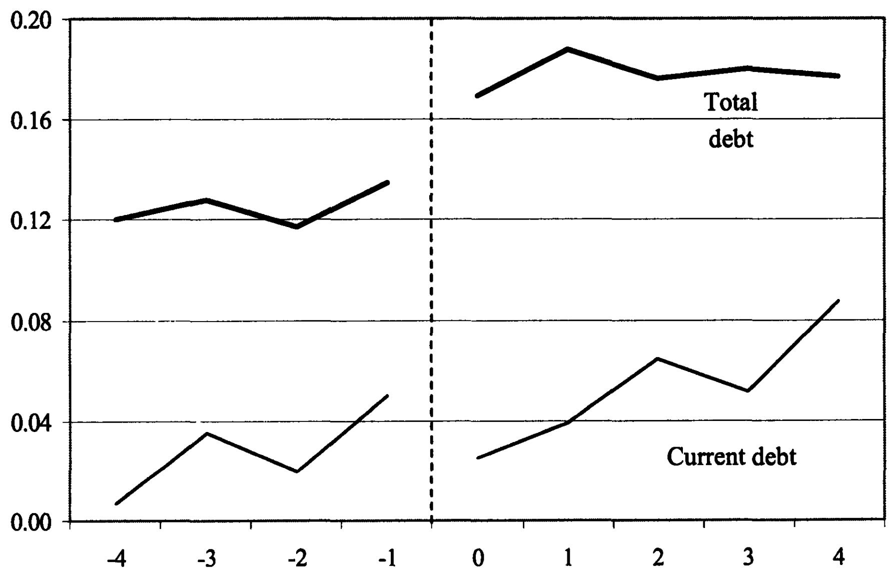
*Source*: Ellwood and Fine (1987); Hutchens, Lipsky, and Stern (1989)

Figure 1-4: Debt around right-to-work law adoption, 1950-1960

A. Firms in industries with high union presence



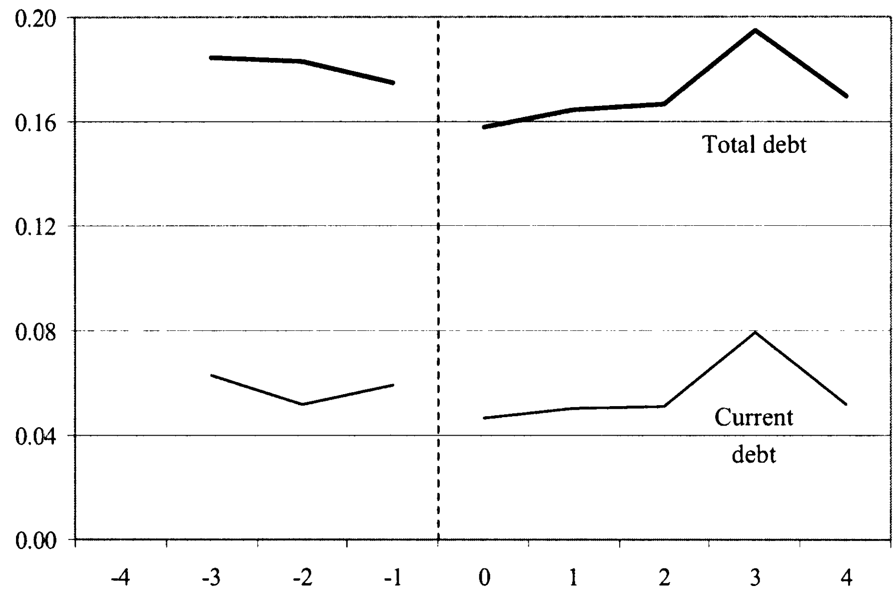

B. Firms in industries with low union presence



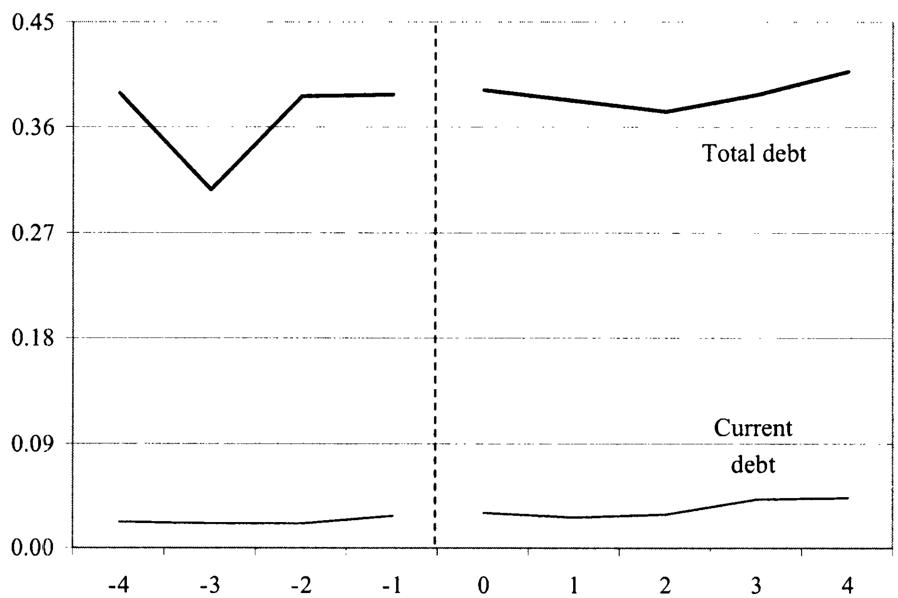

*Note*: These figures depict average debt divided by total assets in the four years before and after the adoption of RTW laws. Panel A includes firms in densely unionized industries (listed in Appendix Table A3), and Panel B includes firms in industries with low union presence.

Figure 1-5: Debt around repeal of UI work stoppage provisions, 1960-1973

A. Firms with more variable profits




B. Firms with less variable profits




*Note*: These figures depict average debt divided by total assets in the four years before and after adoption of the indicated labor law. Panel A includes firms with profit variability in the top quartile, and Panel B includes firms with profit variability in the bottom quartile.

Table 1.1: Cross-sectional analysis -- Summary statistics

| | 1977 | 1987 | 1999 |
|---|---|---|---|
| | *A. Ratio of debt to market value of firm* | | |
| Debt due within... | | | |
| 1 year | 0.060 | 0.063 | 0.053 |
| | (0.074) | (0.080) | (0.055) |
| 2 years | 0.087 | 0.090 | 0.089 |
| | (0.093) | (0.104) | (0.089) |
| 3 years | 0.113 | 0.112 | 0.124 |
| | (0.110) | (0.119) | (0.108) |
| 4 years | 0.137 | 0.134 | 0.171 |
| | (0.123) | (0.134) | (0.145) |
| 5 years | 0.161 | 0.159 | 0.200 |
| | (0.138) | (0.151) | (0.157) |
| Anytime | 0.328 | 0.279 | 0.360 |
| | (0.218) | (0.206) | (0.216) |
| | *B. Other key variables* | | |
| Union coverage | 0.333 | 0.254 | 0.326 |
| | (0.280) | (0.258) | (0.237) |
| % sample with any union coverage | 77.3 | 71.2 | 100.0 |
| Profit variability | 1.224 | 1.159 | 0.848 |
| | (1.000) | (1.000) | (1.000) |
| Assets ($ Mil 1999) | 2,817 | 4,077 | 8,111 |
| | (7,727) | (13,731) | (17,589) |
| | [592] | [644] | [2,761] |
| | *C. Financial control variables* | | |
| Fixed assets (%) | 0.343 | 0.352 | 0.414 |
| | (0.148) | (0.164) | (0.208) |
| Marginal tax rate | | 0.364 | 0.316 |
| | | (0.065) | (0.081) |
| Market-to-book | 1.108 | 1.708 | 2.822 |
| | (0.667) | (1.401) | (3.750) |
| Ln sales ($ Mil) | 5.833 | 6.429 | 7.669 |
| | (1.623) | (1.801) | (1.613) |
| Z-score | 2.819 | 2.426 | 1.742 |
| | (0.727) | (0.885) | (0.952) |
| ROA | 0.078 | 0.055 | 0.048 |
| | (0.054) | (0.071) | (0.067) |
| Observations | 656 | 368 | 349 |

*Note:* Means are presented with standard deviations in parentheses. Median total assets are in brackets. The number of observations listed is for total and current debt (due in 1 year); some of the other debt variables have fewer observations. The samples consist of firms with at least five years of data underlying the measure of profit variability. Data for 1977 and 1987 include only manufacturing firms.

Table 1.2: Unionization and current debt -- Cross-sectional evidence

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | *A. Manufacturing firms, 1977 (n = 656)* | | | |
| Union coverage | 0.001 | 0.011 | -0.001 | -0.009 |
| | (0.012) | (0.011) | (0.011) | (0.007) |
| Union coverage * Profit variability | | 0.036 | 0.027 | 0.016 |
| | | (0.010) | (0.009) | (0.006) |
| Profit variability | | 0.009 | 0.005 | 0.002 |
| | | (0.004) | (0.004) | (0.002) |
| $R^2$ | 0.05 | 0.11 | 0.28 | 0.21 |
| | *B. Manufacturing firms, 1987 (n = 368)* | | | |
| Union coverage | 0.042 | 0.035 | 0.027 | 0.009 |
| | (0.019) | (0.019) | (0.017) | (0.014) |
| Union coverage * Profit variability | | 0.040 | 0.051 | 0.040 |
| | | (0.016) | (0.014) | (0.011) |
| Profit variability | | -0.003 | -0.021 | -0.017 |
| | | (0.006) | (0.005) | (0.004) |
| $R^2$ | 0.09 | 0.12 | 0.30 | 0.21 |
| | *C. Manufacturing and non-manufacturing firms, 1999 (n = 349)* | | | |
| Union coverage | -0.002 | -0.005 | -0.004 | -0.025 |
| | (0.015) | (0.015) | (0.015) | (0.016) |
| Union coverage * Profit variability | | 0.045 | 0.045 | 0.011 |
| | | (0.015) | (0.015) | (0.015) |
| Profit variability | | -0.015 | -0.017 | 0.001 |
| | | (0.007) | (0.007) | (0.007) |
| $R^2$ | 0.11 | 0.17 | 0.23 | 0.25 |
| Financial controls | | | X | X |
| Book value | | | | X |

*Note:* Reported coefficients are estimated from regressions of debt in current liabilities divided by the market value of the firm (divided by assets in Column 4). Debt is regressed on the fraction of a firm's workforce covered by collective bargaining, the variability of the firm's profits, an interaction of those variables, and a set of controls. Profit variability is measured in units of standard deviations of sd($\Delta$earnings)/assets, where earnings is before depreciation and amortization. When uninteracted, the collective bargaining coverage coefficient measures the effect of the law at the mean of profit variability, and the profit variability coefficient measures the effect for non-unionized firms. Controls in all regressions include industry fixed effects at the two-digit SIC level. Where indicated, controls also include financial controls: the proportion of fixed assets, the before interest marginal tax rate (1987, 1999 only), the market-to-book ratio, log sales, modified Altman's z-score, and ROA. Standard errors are reported in parentheses. Sample consists of manufacturing firms with at least five years of pre-period data. Compustat variables are winsorized at the 1% tails. In Column (4), the mean (standard deviation) of the dependant variable is 0.042 (0.048).

Table 1.3: Unionization and other near-term debt measures -- Cross-sectional evidence

| | 1977 | 1987 | 1999 |
|---|---|---|---|
| Debt due within... | | | |
| 1 year | 0.150 | 0.206 | 0.278 |
| | (0.061) | (0.057) | (0.093) |
| 2 years | 0.107 | 0.198 | 0.124 |
| | (0.046) | (0.058) | (0.094) |
| 3 years | 0.097 | 0.161 | 0.086 |
| | (0.038) | (0.051) | (0.080) |
| 4 years | 0.071 | 0.167 | -0.025 |
| | (0.037) | (0.048) | (0.072) |
| 5 years | 0.066 | 0.140 | -0.072 |
| | (0.033) | (0.047) | (0.069) |

*Note:* Each estimate represents a separate regression and is the derivative, with respect to profit variability, of the union coverage elasticity of debt, evaluated at the mean. Reported elasticity effects are estimated from regressions similar to those reported in Table 1.2, Column 3, but for different dependent variables. These regressions are of debt due within the indicated number of years divided by the market value of the firm.

Table 1.4: Unionization and total debt -- Cross-sectional evidence

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | *A. Manufacturing firms, 1977 (n = 656)* | | | |
| Union coverage | 0.062<br>(0.032) | 0.085<br>(0.032) | 0.000<br>(0.023) | -0.015<br>(0.016) |
| Union coverage<br>* Profit variability | | 0.026<br>(0.029) | -0.011<br>(0.020) | -0.016<br>(0.014) |
| Profit variability | | 0.029<br>(0.010) | 0.023<br>(0.008) | 0.015<br>(0.005) |
| $R^2$ | 0.17 | 0.19 | 0.62 | 0.47 |
| | *B. Manufacturing firms, 1987 (n = 368)* | | | |
| Union coverage | 0.168<br>(0.047) | 0.162<br>(0.047) | 0.094<br>(0.041) | 0.019<br>(0.030) |
| Union coverage<br>* Profit variability | | 0.028<br>(0.039) | 0.065<br>(0.033) | 0.037<br>(0.025) |
| Profit variability | | 0.026<br>(0.014) | -0.018<br>(0.013) | -0.010<br>(0.010) |
| $R^2$ | 0.14 | 0.16 | 0.41 | 0.34 |
| | *C. Manufacturing and non-manufacturing firms, 1999 (n = 349)* | | | |
| Union coverage | 0.157<br>(0.053) | 0.147<br>(0.054) | 0.076<br>(0.044) | -0.089<br>(0.037) |
| Union coverage<br>* Profit variability | | -0.008<br>(0.055) | 0.000<br>(0.043) | -0.060<br>(0.037) |
| Profit variability | | 0.016<br>(0.026) | -0.005<br>(0.021) | 0.031<br>(0.018) |
| $R^2$ | 0.25 | 0.26 | 0.58 | 0.40 |
| Financial controls | | | X | X |
| Book value | | | | X |

*Note:* Reported coefficients are estimated from regressions similar to those reported in Table 1.2, but for different dependent variables. These regressions are of total debt divided by the market value of the firm (divided by assets in Column 4). In Column (4), the mean (standard deviation) of the dependant variable is 0.225 (0.131).

Table 1.5: Unionization and inventory policy -- Cross-sectional evidence (manufacturing firms, 1977)

| | (1) | (2) | (3) | (1) | (2) | (3) |
|---|---|---|---|---|---|---|
| | *A. Total inventories / sales (mean = 0.188, n = 651)* | | | *B. Raw materials / sales (mean = 0.068, n = 463)* | | |
| Union coverage | 0.008<br>(0.012) | 0.000<br>(0.012) | 0.007<br>(0.010) | -0.007<br>(0.007) | -0.004<br>(0.007) | -0.001<br>(0.007) |
| Union coverage<br>* Profit variability | | 0.043<br>(0.011) | 0.041<br>(0.009) | | 0.005<br>(0.007) | 0.003<br>(0.007) |
| Profit variability | | -0.015<br>(0.004) | -0.025<br>(0.004) | | 0.002<br>(0.002) | -0.003<br>(0.002) |
| $R^2$ | 0.29 | 0.31 | 0.49 | 0.22 | 0.23 | 0.31 |
| | *C. Work-in-progress / sales (mean = 0.053, n = 376)* | | | *D. Finished goods / sales (mean = 0.076, n = 422)* | | |
| Union coverage | -0.001<br>(0.008) | -0.004<br>(0.009) | -0.004<br>(0.008) | 0.011<br>(0.009) | 0.003<br>(0.009) | 0.001<br>(0.009) |
| Union coverage<br>* Profit variability | | 0.017<br>(0.009) | 0.015<br>(0.009) | | 0.029<br>(0.008) | 0.026<br>(0.008) |
| Profit variability | | -0.006<br>(0.003) | -0.009<br>(0.003) | | -0.013<br>(0.003) | -0.013<br>(0.003) |
| $R^2$ | 0.38 | 0.39 | 0.44 | 0.13 | 0.18 | 0.24 |
| Financial controls | | | X | | | X |

*Note:* Reported coefficients are estimated from regressions similar to those reported in Table 1.2, but for different dependent variables.

Table 1.6: Labor law analysis -- Summary statistics

| | Right-to-Work Laws 1950-1960 | UI Work Stoppage Provisions 1960-1973 |
|---|---|---|
| | *A. Debt* | |
| Current debt / Market value | 0.048 | 0.061 |
| | (0.089) | (0.086) |
| Total debt / Market value | 0.195 | 0.306 |
| | (0.181) | (0.215) |
| Current debt / Book value | 0.036 | 0.055 |
| | (0.064) | (0.070) |
| Total debt / Book value | 0.157 | 0.287 |
| | (0.137) | (0.182) |
| | *B. Other key variables* | |
| States adopting/repealing law | 7 | 7 |
| % sample in adopting states | 2.3 | 21.4 |
| Profit variability | 1.574 | 1.012 |
| | (1.000) | (1.000) |
| Assets ($ Mil 1999) | 2,003 | 1,937 |
| | (6,879) | (7,580) |
| | [487] | [398] |
| | *C. Financial control variables* | |
| Fixed assets (%) | 0.378 | 0.480 |
| | (0.188) | (0.269) |
| Ln sales ($ Mil) | 4.853 | 4.546 |
| | (1.455) | (1.574) |
| ROA | 0.122 | 0.107 |
| | (0.060) | (0.066) |
| Observations | 3,277 | 14,150 |
| Firms | 326 | 1,273 |

*Note:* Means are presented with standard deviations in parentheses. Median total assets in brackets. Sample consists of firms in industries with high union coverage (listed in Appendix Table 1.A4) and with at least five years of data.

Table 1.7: Effect of changes in labor law on current debt

| | (1) | (2) | (3) |
|---|---|---|---|
| | *A. Right-To-Work Laws, 1950-1960* | | |
| RTW law in effect | -0.045 | -0.050 | -0.041 |
| | (0.036) | (0.041) | (0.030) |
| RTW law in effect * Profit variability | -0.048 | -0.056 | -0.034 |
| | (0.010) | (0.011) | (0.006) |
| Observations | 3,277 | 2,976 | 2,976 |
| $R^2$ | 0.66 | 0.69 | 0.73 |
| | *B. Work Stoppage Provisions, 1960-1973* | | |
| No WSP in effect | -0.006 | -0.003 | -0.004 |
| | (0.005) | (0.007) | (0.005) |
| No WSP in effect * Profit variability | -0.013 | -0.011 | -0.009 |
| | (0.005) | (0.007) | (0.004) |
| Observations | 14,150 | 13,705 | 13,705 |
| $R^2$ | 0.63 | 0.66 | 0.66 |
| Financial controls | | X | X |
| Book value | | | X |

*Note:* Reported coefficients are estimated from regressions of debt in current liabilities divided by the market value of the firm (divided by assets in Column 3). Debt is regressed on a RTW law or WSP indicator variable, an interaction of that variable with the variability of the firm's profits, and a set of controls. (The profit variability main effect is absorbed by a firm fixed effect.) Profit variability is measured in units of standard deviations of sd(Δearnings)/assets, where earnings is before depreciation and amortization. When uninteracted, the RTW indicator measures the effect of the law at the mean of profit variability. Controls in all regressions include firm and industy-by-year fixed effects. Where indicated, controls also include financial controls: the proportion of fixed assets, log sales, and ROA. Industry fixed effects are at the two-digit SIC level. Standard errors, clustered at state level, are reported in parentheses. Compustat variables are winsorized at the 1% tails. The sample includes firms in industries with high union coverage (listed in Appendix Table 1.A4).

Table 1.8: Effect of changes in labor law on total debt

| | (1) | (2) | (3) |
|---|---|---|---|
| | *A. Right-To-Work Laws, 1950-1960* | | |
| RTW law in effect | -0.107 | -0.108 | -0.077 |
| | (0.053) | (0.052) | (0.024) |
| RTW law in effect | -0.030 | -0.037 | -0.013 |
| * Profit variability | (0.026) | (0.023) | (0.009) |
| Observations | 3,277 | 2,976 | 2,976 |
| $R^2$ | 0.78 | 0.82 | 0.83 |
| | *B. Work Stoppage Provisions, 1960-1973* | | |
| No WSP in effect | -0.013 | -0.007 | -0.005 |
| | (0.007) | (0.007) | (0.007) |
| No WSP in effect | -0.030 | -0.026 | -0.013 |
| * Profit variability | (0.007) | (0.007) | (0.006) |
| Observations | 14,150 | 13,705 | 13,705 |
| $R^2$ | 0.80 | 0.83 | 0.86 |
| Financial controls | | X | X |
| Book value | | | X |

*Note:* Reported coefficients are estimated from regressions similar to those reported in Table 1.6, but for different dependent variables. These regressions are of total debt divided by the market value of the firm (divided by assets in Columns 3).

Table 1.9: Additional robustness checks -- Effects of changes in labor law

| | *A. Right-To-Work Laws, 1950-1960* | | | *B. Work Stoppage Provisions, 1960-1973* | |
|---|---|---|---|---|---|
| | (1) Current Debt / Market Value | (2) Total Debt / Market Value | | (3) Current Debt / Market Value | (4) Total Debt / Market Value |
| | *I. Operating income as a alternative proxy for threat of union rent-seeking* | | | | |
| RTW law in effect | -0.012 | -0.077 | No WSP in effect | -0.002 | -0.011 |
| | (0.017) | (0.043) | | (0.006) | (0.010) |
| RTW law in effect * Average (pre-period) operating income | -0.059 | -0.046 | No WSP in effect * Average (pre-period) operating income | -0.016 | -0.023 |
| | (0.009) | (0.016) | | (0.004) | (0.005) |
| Observations | 2,488 | 2,488 | Observations | 7,787 | 7,787 |
| $R^2$ | 0.70 | 0.82 | $R^2$ | 0.63 | 0.84 |
| | *II. Operating income as a alternative proxy for threat of union rent-seeking* | | | | |
| RTW law in effect | -0.050 | -0.111 | No WSP in effect | -0.009 | -0.019 |
| | (0.045) | (0.067) | | (0.006) | (0.011) |
| RTW law in effect * Profit variability | -0.051 | -0.040 | No WSP in effect * Profit variability | -0.017 | -0.041 |
| | (0.009) | (0.031) | | (0.006) | (0.013) |
| 2 years prior to adoption | -0.018 | -0.019 | 2 years prior to repeal | -0.005 | -0.012 |
| | (0.027) | (0.044) | | (0.004) | (0.012) |
| 2 years prior to adoption * Profit variability | -0.011 | -0.037 | 2 years prior to repeal * Profit variability | -0.008 | -0.021 |
| | (0.008) | (0.024) | | (0.005) | (0.011) |
| Observations | 3,277 | 3,277 | Observations | 14,150 | 14,150 |
| $R^2$ | 0.66 | 0.78 | $R^2$ | 0.63 | 0.80 |
| | *III. Falsification test: Industries with low union presence* | | | | |
| RTW law in effect | -0.005 | 0.030 | No WSP in effect | 0.001 | 0.008 |
| | (0.016) | (0.030) | | (0.004) | (0.012) |
| RTW law in effect * Profit variability | 0.010 | -0.043 | No WSP in effect * Profit variability | 0.006 | 0.014 |
| | (0.030) | (0.046) | | (0.004) | (0.009) |
| Observations | 2,381 | 2,381 | Observations | 11,989 | 11,989 |
| $R^2$ | 0.68 | 0.75 | $R^2$ | 0.77 | 0.79 |

*Note:* In Panel I, reported coefficients are estimated from regressions similar to those reported in Tables 1.6 and 1.7, Column 2, but the law indicator variable is interacted with the firm's average (pre-period) operating income rather than with profit variability. Operating income is before interest expense, payment of current debt maturities, taxes, depreciation, and amortization, is divided by assets, and is normalized by its standard deviation (0.091). In Panel II, reported coefficients are estimated from regressions similar to those reported in Tables 1.6 and 1.7, Column 1, but they also include an indicator variable for the 2 years before the legal change and an interaction of that variable with profit variability. In Panel III, reported coefficients are estimated from regressions similar to those reported in Tables 1.6 and 1.7, Column 2, but on a different sample of firms. The sample includes observations of firms in industries with low rates of union coverage (less than 25 percent of the workforce covered by collective bargaining).

Table 1.10: Union bargaining power and dividends

*A. Cross-sections of firms*

| | 1977 | 1987 | 1999 |
|---|---|---|---|
| RTW law in effect | 0.0094 | 0.0000 | 0.0039 |
| | (0.0021) | (0.0029) | (0.0029) |
| RTW law in effect | 0.0009 | -0.0010 | 0.0000 |
| * Profit variability | (0.0018) | (0.0024) | (0.0028) |
| Profit variability | -0.0021 | -0.0030 | -0.0015 |
| | (0.0007) | (0.0009) | (0.0014) |
| Observations | 651 | 368 | 348 |
| $R^2$ | 0.36 | 0.29 | 0.44 |

*B. Right-To-Work Laws, 1950-1960*

| | (1) | (2) | (3) |
|---|---|---|---|
| RTW law in effect | 0.0035 | 0.0034 | 0.0025 |
| | (0.0016) | (0.0016) | (0.0028) |
| RTW law in effect | 0.0019 | 0.0033 | 0.0016 |
| * Profit variability | (0.0010) | (0.0012) | (0.0015) |
| Observations | 3,277 | 2,976 | 2,976 |
| $R^2$ | 0.74 | 0.77 | 0.83 |
| Financial controls | | X | X |
| Book value | | | X |

*C. Work Stoppage Provisions, 1960-1973*

| | (4) | (5) | (6) |
|---|---|---|---|
| No WSP in effect | 0.0012 | 0.0011 | 0.0006 |
| | (0.0009) | (0.0009) | (0.0009) |
| No WSP in effect | -0.0001 | -0.0007 | 0.0001 |
| * Profit variability | (0.0007) | (0.0009) | (0.0009) |
| Observations | 14,150 | 13,705 | 13,705 |
| $R^2$ | 0.78 | 0.78 | 0.87 |
| Financial controls | | X | X |
| Book value | | | X |

*Note:* Reported coefficients are estimated from regressions of common stock dividends divided by the market value of the firm (divided by assets in Columns 5 and 7). In Panel A, the specification is similar to those reported in Table 1.2, Column 3. The mean of the dependant variable is 0.025 in 1977, 0.017 in 1987, and 0.012 in 1999, and the standard deviation is 0.015, 0.013, and 0.012, respectively. In Panels B and C, the specifications are similar to those reported in Table 1.6. The mean of the dependant variable is 0.038 in Columns 1 and 2, 0.035 in Column 3, 0.020 in Columns 4 and 5, and 0.022 in Column 6, and the standard deviation is 0.023, 0.024, 0.015, and 0.019, respectively.

Appendix Table 1.A1: Profit variability, expected profit, and the union wage premium, 1983

| *A. Profit* | | *B. Log Hourly Wage* | |
|---|---|---|---|
| | (1) | | (2) |
| Profit variability | 0.112 | Union coverage | 0.110 |
| | (0.025) | * Profit variability | (0.032) |
| | | Union coverage | 0.137 |
| | | | (0.013) |
| Unit of analysis | Firm | Unit of analysis | Worker |
| Observations | 4,555 | Observations | 47,969 |
| $R^2$ | 0.13 | $R^2$ | 0.55 |

*Note:* In Panel A, the dependant variable is profits (operating income before interest expense, payment of current debt maturities, taxes, depreciation, and amortization) divided by assets. Profits are regressed on the (firm's historical) variability of profits and a set of controls. Profit variability is measured in units of standard deviations of sd(Δprofits/assets), calculated over the previous ten years. Controls include 4-digit SIC industry fixed effects. The standard error is reported in parentheses. In Panel B, Compustat data is merged at the 2-digit SIC level with employee wages and demographics from the Current Population Survey Outgoing Rotation Group Earnings Files. Log hourly wages are regressed on a variable indicating whether the employee is covered by a collective bargaining agreement, an interaction of union coverage with industry profit variability (the profit variability main effect is absorbed by a fixed effect), and a set of controls. The controls include age and age-squared as well as fixed effects for gender, race, years of education, occupation, and industry. Standard errors, clustered at the industry level, are reported in parentheses. Industry fixed effects and clustering are at the two-digit SIC level, and occupation fixed effects are for each of 408 census occpation codes. Compustat variables are winsorized at the 1% tails.

Appendix Table 1.A2: Effect of right-to-work laws on union organizing, 1950-1960

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| RTW law in effect | -0.382 | -0.347 | -0.423 | -0.336 |
| | (0.122) | (0.161) | (0.112) | (0.150) |
| 2 years prior to adoption | | | -0.086 | 0.011 |
| | | | (0.098) | (0.083) |
| $R^2$ | 0.83 | 0.85 | 0.83 | 0.85 |
| State and year fixed effects | X | X | X | X |
| State-specific linear trends | | X | | X |

*Note:* Reported coefficients are estimated from regressions of the log number of new members of collective bargaining units. Union organizing is regressed on a RTW law indicator variable and a set of controls. Controls in all regressions include state and year fixed effects. Where indicated, controls also include an indicator for the two years before a RTW law is adopted and state-specific linear time trends. Standard errors, clustered at state level, are reported in parentheses.

Appendix Table 1.A3: Additional robustness checks -- Alternative measures of profit variability

| | Current Debt / Market Value | | | Total Debt / Market Value | | |
|---|---|---|---|---|---|---|
| | *A. Right-To-Work Laws, 1950-1960* | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| RTW law in effect | -0.037<br>(0.035) | -0.026<br>(0.035) | -0.036<br>(0.037) | -0.099<br>(0.049) | -0.093<br>(0.049) | -0.098<br>(0.046) |
| RTW law in effect<br>* Operating income variability | -0.041<br>(0.008) | | | -0.029<br>(0.016) | | |
| RTW law in effect<br>* Operating cash flow variability | | -0.036<br>(0.006) | | | -0.024<br>(0.016) | |
| RTW law in effect<br>* Sales variability | | | -0.029<br>(0.006) | | | -0.022<br>(0.012) |
| Observations | 2,812 | 2,812 | 2,809 | 2,812 | 2,812 | 2,809 |
| $R^2$ | 0.68 | 0.68 | 0.68 | 0.82 | 0.82 | 0.82 |
| | *B. Work Stoppage Provisions, 1960-1973* | | | | | |
| | (7) | (8) | (9) | (10) | (11) | (12) |
| No WSP in effect | -0.003<br>(0.007) | -0.002<br>(0.008) | -0.003<br>(0.007) | -0.008<br>(0.007) | -0.006<br>(0.008) | -0.006<br>(0.007) |
| No WSP in effect<br>* Operating income variability | -0.010<br>(0.006) | | | -0.030<br>(0.011) | | |
| No WSP in effect<br>* Operating cash flow variability | | -0.013<br>(0.007) | | | -0.016<br>(0.010) | |
| No WSP in effect<br>* Sales variability | | | -0.018<br>(0.006) | | | -0.021<br>(0.013) |
| Observations | 13,644 | 10,517 | 13,705 | 13,644 | 10,517 | 13,705 |
| $R^2$ | 0.66 | 0.65 | 0.66 | 0.83 | 0.80 | 0.83 |

*Note:* Reported coefficients are estimated from regressions similar to those reported in Tables 1.6 and 1.7, Column 2, but the law indicator variable is interacted with alternative measures of variability. Each variability measure is in units of standard deviations of sd($\Delta x$)/assets, but for a different financial variable $x$. Operating income is EBITDA, accrual-basis operating profits before interest expense, taxes, depreciation, and amortization. Operating cash flow is a cash-basis measure of operating profits, calculated from reported accounting measures using the Sloan (1995) adjustment for accruals.

Appendix Table 1.A4: Industries included in labor law analyses (Tables 1.6 - 1.9)

| Industry | Observations: RTW analysis | Observations: WSP analysis | Union Coverage Rate, 1983 |
|---|---|---|---|
| Mineral industries | | | |
| Metal mining (10) | 33 | 150 | 0.42 |
| Coal mining (12) | 21 | 62 | 0.63 |
| Nonmetallic minerals, except fuels (14) | 33 | 91 | 0.35 |
| Construction industries | | | |
| General building contractors (15) | 11 | 122 | 0.30 |
| Heavy construction contractors (16) | 43 | 125 | 0.30 |
| Special trade contractors (17) | 0 | 29 | 0.30 |
| Manufacturing | | | |
| Food and kindred products (20) | 519 | 1,423 | 0.37 |
| Tobacco products (21) | 66 | 84 | 0.40 |
| Apparel and other textile products (23) | 78 | 551 | 0.29 |
| Paper and allied products (26) | 213 | 593 | 0.51 |
| Petroleum and coal products (29) | 191 | 365 | 0.35 |
| Rubber and miscellaneous plastics products (30) | 137 | 525 | 0.30 |
| Leather and leather products (31) | 34 | 181 | 0.26 |
| Stone, clay, glass, and concrete products (32) | 179 | 506 | 0.40 |
| Primary metal industries (33) | 303 | 899 | 0.56 |
| Fabricated metal products (34) | 181 | 979 | 0.35 |
| Electrical and electronic equipment (36) | 315 | 1,582 | 0.26 |
| Transportation equipment (37) | 403 | 1,183 | 0.50 |
| Transportation, communication, and utilities | | | |
| Railroad transportation (40) | 0 | 132 | 0.85 |
| Local and interurban highway passenger transit (41) | 0 | 12 | 0.48 |
| Motor freight transportation and warehousing (42) | 37 | 298 | 0.39 |
| Water transportation (44) | 0 | 50 | 0.42 |
| Transportation by air (45) | 153 | 337 | 0.46 |
| Communications (48) | 119 | 522 | 0.52 |
| Electric, gas, and sanitary services (49) | 28 | 371 | 0.44 |
| Retail trade | | | |
| Food stores (54) | 138 | 500 | 0.30 |
| Service industries | | | |
| Motion pictures (78) | 42 | 121 | 0.26 |
| Educational services (82) | 0 | 24 | 0.44 |

*Source:* Hirsch and Macpherson (2003), based on the Current Population Survey Outgoing Rotation Group Earnings Files, 1983; Bureau of Census Technical Paper 59 (1989)

*Note:* Industries listed are those with at least 25 percent of workforce covered by collective bargaining agreements in 1983 and at least one firm with Compustat data during the 1950s. Two-digit SIC codes are reported in parentheses. Sample includes employed wage and salary workers, ages 16 and over.

# Chapter 2

# Operating Under a Liquidity Crunch: The Impact of LBOs on Product Availability in the Supermarket Industry

## 2.1 Introduction

The financial structure of a business can impact its operations in important ways. In this chapter, I investigate how financial structure affects investments in marketing and operations management in the retail sector. In particular, I examine how leveraged buyouts (LBOs) – transactions that involve substantial increases in debt – impact retail stockouts (running out of inventory for a given product that is usually offered for sale) in the supermarket industry.

Maintaining optimal product availability is a first-order issue in the retail sector. Customer substitution upon encountering out-of-stocks is estimated to cost a typical supermarket 1.7 to 3.1 percent of sales (Andersen Consulting 1996). The costs are even greater when you include sales lost from some consumers shifting future shopping to competing stores. As an investment in future market share, product availability can be likened to R&D at an industrial firm. The results presented in this chapter suggest that leveraged buyouts increase out-of-stocks.

In theory, leveraged buyouts may either increase or decrease out-of-stocks. A liquidity-constraints hypothesis suggests that high debt levels lead firms to reduce inventories and product availability to boost liquidity (Chevalier and Scharfstein 1996). On the other hand, an agency hypothesis suggests an increased demand for liquidity may discourage shirking, stimulate productivity improvements, and lead firms to offer closer to optimal levels of product availability (Jensen 1986). Depending on the initial level, this move toward the optimum may represent either an increase or a decrease in the prevalence of out-of-stocks.

Using U.S. Consumer Price Index microdata to measure the prevalence of out-of-stocks, I examine the impact of supermarket LBOs on product availability. While there is no pre-existing trend, out-of-stocks increase by about 10 percent at supermarkets that undertake an LBO. The effect is long-lived, lasting on average 10 years or more. As a robustness test, I also show that LBOs have little impact on stockouts for product categories for which inventory is directly managed by distributors, in addition to the retailer. I also analyze price changes following an LBO, and find that LBOs lead firms to raise prices by about 1 percent – an economically significant increase for a typical retailer that earns margins of about 1 percent. These results are robust to including store and even store-item fixed effects.

This chapter illustrates an important relationship between a firm's financing and its operations in the retail sector. The findings are consistent with the empirical literature on capital-market imperfections and inventory investment, including Carpenter, Fazzari, and Petersen (1994), Kashyap, Lamont, and Stein (1994), and Calomiris, Himmelberg, and Wachtel (1995). This work also extends previous research on LBOs. Studies find that LBOs in the manufacturing sector increase operating income (Kaplan 1989; Smith 1990; Jensen 1993) and total factor productivity (Lichtenberg and Siegel 1990; Harris, Siegel, and Wright 2005), and they decrease capital expenditures (Kaplan 1989), R&D (Long and Ravenscraft 1993), and employment (Lichtenberg and Siegel 1990). Most closely related to this study, Chevalier (1995b) shows that average market price levels increase following leveraged buyouts in the supermarket industry.

The remainder of the chapter is organized as follows. Section 2.2 develops a theoretical framework for optimal retail inventory management and why it may be affected by an LBO. Section 2.3 describes the causes and consequences of LBO activity in the supermarket industry. Section 2.4 describes the data and empirical approach, and Section 2.5 presents estimates of the impact of LBOs on supermarket prices and product availability. Section 2.6 concludes.

## 2.2 Economic determinants of retail inventory management

Maintaining sufficient product availability is a major strategic issue in the retail sector. Estimates suggest that 8.2 percent of a grocery retailer's items are out-of-stock on a typical afternoon (Andersen Consulting 1996, Roland Berger Strategy Consultants 2002).[1] Accounting for consumer substitution patterns, out-of-stocks cost the average grocery retailer 1.7 to 3.1 percent of sales in the short-run.[2] Aggregating over the supermarket industry, this translates to $6 to 12 billion of lost sales per year.[3] Retailers likely also suffer longer-run reductions in sales when out-of-stocks lead previously-loyal customers to turn to other stores for future shopping needs.

Out-of-stocks primarily reflect retailer operating performance, as opposed to the performance of other parties in the vertical chain. Andersen Consulting (1996) concludes that retailers bear the responsibility for 97% of out-of-stocks of warehouse-supplied items and 76% of direct store delivery items. Similarly, Gruen et al. (2002) finds that 73% of out-of-stocks in the United States are caused in the store. Retailer operating data analyzed in these studies demonstrate that 51 to 73% of out-of-stocks are due to inaccurate forecasting (e.g., maintaining too little inventory) or ordering errors (e.g., failing to sufficiently monitor the shelf inventory and not reordering when demand exceeds forecast) and another 8 to 22% are due to failing to restock the shelf with available backroom or display inventory.

While no retailer desires out-of-stocks, reducing them is costly and maintaining 100 percent product availability at supermarkets is certainly not optimal. Optimal stocking decisions trade off expenditures on both inventory costs and monitoring the shelf for the

[1]Maintaining the right level of product availability is a long-standing issue in the industry: in 1968, Progressive Grocer reported that more than 20 percent of shoppers leave a store wanting to buy an out-of-stock item, and in their 1996 study, Andersen Consulting found that 48 percent of items they surveyed were out-of-stock at least once a month.

[2]Upon encountering an out-of-stock, consumers substitute in various ways, including purchasing an alternative item (0.4 percent decline in intended purchase expenditure), shopping at another store or cancelling the purchase (1.3 percent), and delaying purchase (1.3 percent; Andersen Consulting 1996).

[3]This figure aggregates retailers' lost intended purchase expenditure, but it does not include the sales gained by retailers because of out-of-stock at other stores. The number is meaningful from the perspective of an individual profit-maximizing retailer, but not from the perspective of the industry. The decrease in aggregate supermarket sales, which includes both terms, is likely much lower.

present value of expected lost profits from out-of-stocks. Importantly, the stores consider not only the lost margin from the current purchases of the product but also the impact on consumers' future shopping behavior. Switching supermarkets is thought to be costly for shoppers, who are accustomed to a particular store's layout and a regular food shopping routine. Given these consumer switching costs, a small but important risk of an out-of-stock is that it may trigger the "long-run" substitution of a customer's regular business to another retailer. In this sense, the provision of product availability is an investment in future market share. A similar argument can be made for price setting; firms have an incentive to attract new customers today through low prices in order to have market power over them in the future (see Chevalier 1995b).

A number of factors affect the optimal level of product availability for a given retailer and product, including the customers' substitution pattern upon encountering an out-of-stock, the price elasticity of demand, the wholesale cost, the inventory cost, and variability of demand. Customers' short- and long-run substitution behavior varies with the degree of brand loyalty and product variety in the product category as well as the degree of competition in the retailer's local market. Products facing less elastic demand earn greater markups and are more valuable to keep in stock. Other products are more costly to inventory, such as refrigerated versus shelf-stable products. And some products, such as seasonal items, have less predictable demand. There are also differences in the optimal product availability rate across retailers. Returns to scale in demand forecasting and order management may reduce the cost of providing product availability in large stores. Stores that are vertically integrated with their primary supplier may face lower wholesale and inventory costs. While most stockouts are caused in the retail store, inventory management practices at the store's supplier and the distance from that supplier may also affect the store's optimal stockout rate. Technological advances and other changes over time may also affect inventory costs. To control for these and other factors that affect the optimal rate of product availability, I control for product category, store, and year-month fixed effects as well as a number of item and time-varying store characteristics in the analysis below.

Liquidity constraints and agency problems may also lead a firm to deviate from its optimal level of product availability. The high leverage taken on by a firm during a leveraged buyout may lead cash-constrained firms to cut positive-NPV investment (Chevalier and Scharfstein 1996). By increasing the firm's cost of capital, taking on the additional debt may decrease investment in product availability, because it increases both the costs of holding inventory and the firm's discount rate. First, suppliers may reduce trade credit to LBO firms that they perceive to be a greater repayment risk. Even when trade credit terms do not change, greater cash flow demands of debt service increase the shadow cost of inventory, leading firms to reduce inventories to boost liquidity (Hubbard 1998). Second, an increased discount rate leads LBO firms to have a greater preference for current, relative to future, profits. Recall that setting product availability trades off cost today (financing inventory and monitoring shelves) and both benefits today (revenue from sale of that product) and benefits in the future (margins from incremental future sales). Thus, an increased discount rate leads LBO firms to under-value the future benefits and reduce inventory.

Agency theory, on the other hand, suggests that an increased demand for liquidity may also stimulate productivity improvements (Jensen 1986). Separating ownership and control creates an agency problem that may, among other things, depress operating performance and lead a firm to either over-invest or under-invest in product availability. On one hand, shirking managers may take insufficient precautions against stockouts. For example, retailers may fail to invest in effective demand forecasting or they may fail to effectively monitor the shelves (e.g., by "filling the holes" on the shelf with other products but failing to reorder the out-of-stock item). On the other hand, managers may over-invest in product availability. Some managers may invest in too much inventory in pursuit of "the quiet life" free of customer complaints, and other "empire-building" managers may over-invest in product availability to maximize market share rather than profits.

In a model with agency problems, a liquidity crunch may actually push the firm toward its optimal level. Undertaking an LBO increases the monitoring of managers, since non-management debt and equity are often more highly concentrated after the buyout. Debt may also mitigate the agency problem by requiring managers to disgorge "free cash flow" (Jensen 1986). Depending on whether unconstrained managers over-invest or under-invest in product availability, the agency hypothesis predicts a decrease or an increase in out-of-stocks.

The credit-constraints hypothesis and the agency hypothesis differ in their welfare implications. While the agency hypothesis contends that high levels of debt generate efficiency gains in inventory management, the liquidity-constraints hypothesis holds that debt constrains firms financially, distorting optimal decision-making. I analyze the impact of LBOs on out-of-stocks to identify the net effect. As in the previous literature that documents changes in capital expenditure, R&D, employment, and retail prices, it is difficult to determine which mechanism is affecting firm behavior.

## 2.3 Causes and consequences of LBO activity in the supermarket industry

Like many other industries, a wave of LBO activity overtook the supermarket industry in the latter half of the 1980s. In these transactions, an investment group and often senior management acquire the firm and contribute 5 to 25 percent of the total financing. The remainder of the financing is obtained through bank loans or lines of credit, usually secured by the company's assets, as well as junk bonds, debentures, or unsecured loans. After a leveraged buyout, the typical debt-to-equity ratio is greater than ten to one (Bongard and Cross 1992). By 1991, LBO firms accounted for almost a quarter of industry sales, including 19 of the largest 50 supermarket chains.

While LBO activity was not concentrated in any particular geographic region, it was more prevalent among larger firms and stores. Table 2.1 reports summary statistics in 1990 for three samples of firms: firms that undertook an LBO before 1990, firms that would undertake an LBO after 1990, and non-LBO firms. LBO firms averaged 2 to 6 times as many stores as non-LBO firms, and LBO stores averaged about 50 percent larger than non-LBO stores in terms of square feet, sales, and employment.

Despite the differences in size, firms yet to LBO and firm that never LBO had nearly identical stockout rates (both with and without controls for size and other store characteristics). But the stockout rate was 4.5 percent higher at firms that had already undertaken an LBO, perhaps reflecting the impact of the buildup in debt. Differences in average price levels, on the other hand, are more difficult to interpret with firms yet to LBO charging the highest prices.[4] In light of these differences between LBO and non-LBO firms, most estimates reported in this chapter use within-firm and within-store variation to identify LBO effects. Nevertheless, these fixed-effects estimates are generally very similar to estimates

[4] On average, firms yet to LBO charge the highest (unconditional mean) prices, but these prices may actually be relatively low after controlling for store characteristics. It is possible LBO activity may have targeted low-price stores in high-price markets.

identified using cross-sectional variation with controls for size, metropolitan area, and other firm and store characteristics.

While supermarket leveraged buyouts were generally undertaken to prevent unwanted takeover attempts (Chevalier 1995b), it is unclear exactly what underlying economic factors led to the LBOs (or the takeover attempts). The restructuring premiums associated with these deals suggest the targets' assets had not been put to their highest value use, and Jensen (1989) argues this value comes from disciplining "empire-building" managers who were overinvesting the firm's resources. Post-LBO assets sales were common, and it seems many of the LBOs aimed to force the sale of unprofitable divisions (Bongard and Cross 1992). For example, Peter Magowan (1989), then-CEO of Safeway, describes this sort of strategy (in the shadow of an unwelcome takeover bid from Herbert Haft) as a primary motivation for their LBO. For LBOs targeted in this way, the ex ante operating performance of stores *retained* by firms after the buyout was not an important factor in the LBO decision. This mitigates concerns about buyout endogeneity in analyses with store fixed effects.[5] Furthermore, I show below that there do not appear to be pre-existing trends in stockouts or prices at LBO firms after controlling for store fixed effects (Table 2.2; Figure 2-2).

The leveraged buyouts had a significant impact on competition in the supermarket industry. In a series of papers, Judy Chevalier (1995a; 1995b) shows that the LBOs softened product-market competition. Following an LBO, average prices levels increase in metropolitan areas in which the LBO firm's rivals are also highly leveraged, and (in the cross-section) LBO firms charge higher prices than their rivals. The presence of LBO firms also encourages rival firms to enter and expand their operations. Chevalier also presents evidence suggesting that rivals with low leverage attempt to prey on LBO firms: following an LBO in these markets, average prices levels fall and the LBO firm is more likely to exit. Consistent with these results, LBO announcements are associated with positive stock price responses for rival firms.

[5] Beyond the issue of endogeneity, however, is one of external validity. Even if we observe the causal effect of the buyouts on the target firms, it may be that other firms with different characteristics would respond differently. In the labor economics literature, this issue is often referred to as treatment effects heterogeneity (Angrist 2004). This chapter analyzes how LBOs affected firms that undertook those transactions, but I recognize the possibility that LBOs may impact other firms differently. The takeaway from the analysis is that financial policy can affect a firm's operations, not that it always does.

## 2.4 Data and empirical approach

### 2.4.1 Supermarket LBOs

Data on supermarket LBOs come from two sources. First, Judy Chevalier provided me with a sample of supermarket LBOs consummated between 1981 and 1990. She compiled this listing from quarterly editions of *Mergers & Acquisitions* and searches of *Supermarket News*, *Supermarket Business*, and *Progressive Grocer.* While these data are of high quality, many of the LBOs in this sample precede my data on product availability, which begins in 1988. I use the Chevalier sample of LBOs for cross-sectional analysis, comparing price and product availability levels at LBO and non-LBO firms.

Second, I obtain a sample of supermarket LBOs from the Thomson Financial *Securities Data Company's Merger & Acquisitions* database, accessed through SDC Platinum. SDC claims to track more financial transactions than any other source and is used widely by researchers, investment banks, law firms, and media outlets. I include transactions explicitly coded as leveraged buyouts as well as acquisitions by buyout firms.[6] Using these data, which include supermarket LBOs from 1981 to the present, I perform panel analyses, comparing price and product availability levels at each firm before and after undertaking an LBO.

While not identical, the Chevalier and SDC samples are very similar. The samples are compared in Appendix Table 2.A1, Panel A. In 1990, 19 percent of the industry (in terms of sales) is recorded as having undertaken an LBO in both data sets.[7] Another 4 percent is included in the Chevalier LBO sample but not in the SDC sample, likely representing small buyouts picked up by Chevalier's exhaustive searches of industry publications. Two percent of the industry is recorded as having undertaken an LBO by SDC but not by Chevalier, and 75 percent is non-LBO in both samples.

Figure 2-1 shows the timing of supermarket LBOs that identify the panel analysis.[8] While LBO activity was most prevalent in the 1980s and early 1990s, it continued until the late 1990s. In all, 34 unique supermarket firms have undertaken LBOs since 1988,

[6] A small number of the acquisitions by buyout firms may not be highly leveraged. The measurement error induced by including any such transactions may attenuate (but is unlikely to bias) the estimates.

[7] Both the Chevalier sample and the SDC sample also include the Kroger leveraged recapitalization, which resulted in debt levels similar to a typical LBO.

[8] The BLS pledges confidentiality to voluntary respondents in the CPI sample; no inferences should be made from this work as to whether or not a specific firm is included in the CPI sample.

accounting for approximately 5,000 stores and $65 billion in annual sales.

### 2.4.2 Retail prices and product availability

High quality data on supermarket product availability are rare. Anecdotes suggest that most stores themselves do not systematically track availability. The most frequently cited statistics on the prevalence of out-of-stocks come from an Andersen Consulting (1996) study, sponsored by the Coca-Cola Retailing Research Council. The authors performed daily audits of 7,000 items in eight product categories in ten demographically and regionally diverse stores for one month. Such isolated (and often localized) studies do not lend themselves to either cross-sectional or longitudinal analysis at the store or firm level. Due to the cost of conducting wider-scale audits, some studies have attempted to measure out-of-stocks using purchase scanner data (e.g., Gruen et al. 2002). But such studies risk confusing low availability with low demand, which would bias estimates (Dorgan 1997).

I obtain reliable data on prices and out-of-stocks from the *CPI Commodity and Services Survey*, which is used by the Bureau of Labor Statistics (BLS) to compute the consumer price index (CPI). To calculate the CPI, the BLS collects prices on about 30,000 goods sold at grocery stores each month, where each price is specific to a particular product at a particular establishment. Generally, a product must be available for purchase at the establishment at the time of visit by the BLS surveyor in order to be included in the CPI.[9] If the product is unavailable for sale, the BLS surveyor determines whether the establishment expects to carry the item in the future. In this sense, a product may be considered out-of-stock if it is not available for sale, it is continuing to be carried by the outlet, and it is not seasonally unavailable (Bils 2005).[10]

Using these microdata, I examine observations on price and product availability at the

[9] For food items (excluding food consumed away from home), surveyors are actually instructed to record an item as available if the retailer respondent says an out-of-stock item will be restocked later that day. This complicates efforts to measure true product availability. Out-of-stocks caused by retailers failing to stock shelves with back room or display inventory will not be reflected in the BLS data. Industry studies attribute 8 to 22 percent of out-of-stocks to these sort of store shelving issues (Andersen Consulting 1996, Gruen et al. 2002).

[10] In practice, the determination of an out-of-stock is slightly more subtle. First, I condition on the outlet being available for pricing by the BLS surveyor. Second, I consider items with "different day" prices as being out-of-stock at the time of the surveyor's visit. Third, I restrict attention to observations that are at least three months prior to a product becoming permanently or seasonally unavailable. This is an attempt to address a concern that an item reported as "temporarily unavailable" that becomes "permanently unavailable" before another price quote is successfully obtained may not actually represent an out-of-stock (Bils 2005).

item-store-month level from January 1988 through June 2005. While any particular item is sampled for at most 5 years, the full data set includes about 5 million observations on price and availability for almost 220,000 unique items at 9,500 stores in more than 8,000 census tracts and 147 metropolitan areas.

I augment the CPI data with detailed store-level information from the Trade Dimensions *Retail Site Database*.[11] The *Retail Site Database* is a leading source of establishment data in the retail food industry. It includes data on ownership, sales volume, selling area, and the warehouse that primarily supplies each of the more than 33,000 supermarkets in the United States.

Using the full sample, the average out-of-stock rate among supermarket respondents is 4.4 to 5.4 percent.[12] It increases over the sample period from 3.7 to 4.5 percent around 1990 to 5.6 to 7.0 around 2004. Two factors relating to BLS data collection procedures seem to explain why these estimates are lower than the 1996 industry estimate of 8.2 percent. First, CPI data is generally collected throughout the day on weekdays, whereas out-of-stocks are most prevalent in the afternoon and on Sundays (when they reach an estimated 10.9 percent; Andersen Consulting 1996). Second, for food items consumed at home, the BLS effectively does not record out-of-stocks caused by store shelving issues (8 to 22 percent of out-of-stocks).[9,13] While these factors affect the level of out-of-stock estimates, they are

[11] I merge the data sets using the store telephone number, ZIP code, street address, and/or name. I am able to successfully match *Retail Site Database* information to 89.2 percent of the observations on product availability.

[12] The different estimates depend on whether one counts products reported to be temporarily unavailable immediately before the product is reported permanently unavailable. One concern is that the price surveyor may initially misinterpret some product cancellations as temporary out-of-stocks. A concern in the opposite direction is that some products that are repeatedly unavailable because they are out-of-stock may become classified as permanently unavailable. If a product is repeatedly unavailable, it may trigger an instruction to the surveyor to begin pricing a new item at the next visit. However, analysis conducted by Teague Ruder and cited by Bils (2005) suggests that in practice the field agents often continue to price the old version when the product becomes available for purchase. Accepting the surveyors' original classification of product unavailability yields an estimated out-of-stock rate of 5.4 percent. Another option is to only count an item as temporarily out-of-stock if it observed to be available at a later date. Under this definition, I eliminate the final observation for each product, because out-of-stock rates for those observations are zero by construction. This methodology results in an overall out-of-stock rate of 4.4 percent. The remainder of this chapter uses this more conservative algorithm for computing an out-of-stock.

[13] In some situations, it is also possible that a BLS surveyor may record an out-of-stock item as "available" if another similar item is available for purchase. However, the procedures for this sort of substitution are regulated carefully by the BLS. Key characteristics of the product must be the same for the surveyor to execute a substitution. For example, a substitute ready-to-eat cereal product must have the same brand, product name, size, sweeteners, fruit, nuts, flavorings, and more, but it does not have to have the same UPC code. To the extent that the BLS definition of a stockout more closely reflects a consumer's willingness to substitute across nearly identical items (e.g., changes in package design), it may be preferred to a simple SKU-based definition.

unlikely to bias estimates of changes in product availability caused by a leveraged buyout.

### 2.4.3 Empirical approach

I examine the impact of an LBO on supermarket operations using a difference-in-difference regression approach. The majority of supermarket LBOs were consummated in the mid-to-late 1980s. However, information on retail price and product availability from the *CPI Commodity and Services Survey* is only available from 1988. In an initial cross-sectional analysis, I examine whether LBO firms offer less product availability and charge higher prices than their non-LBO rivals in the same local market.

Using item-quote-level data from January 1988 to June 1993, I estimate a linear probability model of an item being out-of-stock.[14] Let $STOCKOUT_{isjt}$ be an indicator for whether product $i$ in store $s$, firm $j$, and month $t$ is out-of-stock, and $LBO_j$ be an indicator for whether the firm undertook a leveraged buyout before 1990.[15]

$$STOCKOUT_{isjt} = \alpha LBO_j + X_{is}^c \beta + Z_s \gamma + \psi_t + \epsilon_{isjt}$$

$X_{is}^c$ includes product and store characteristics. Product-level controls include whether the item is seasonal (i.e., not offered year round) and fixed effects for both the day of the week the item was sampled and the product category.[16] Store-level controls include whether the store is affiliated with a chain, total grocery selling space (categorized into 5 groups), the distance from the warehouse that primarily supplies the outlet (categorized into 5 groups), and whether the outlet is vertically integrated with that supplier.[17] $Z_s$ are ZIP code fixed effects, which control for differences in local markets. $\psi_t$ are year-month fixed effect, which account for seasonal, technological, and other national trends in inventory management. I use a similar framework to assess differences in price levels across LBO and non-LBO firms.

[14] Given the size of the data set and the computer resources available to me at the BLS, maximum likelihood estimation of a probit or conditional logit model is computationally infeasible.

[15] The cross-sectional analysis uses the sample of LBOs collected by Chevalier. While this sample is likely relatively comprehensive, the data provided to me do not include the date of each LBO. In the regressions reported below, the $LBO_j$ indicator equals one for all product availability observations for an LBO firm. Consequently, a small number of observations corresponding to firms that undertook LBOs between 1988 and 1990 are likely misclassified. This may slightly attenuate estimates of the impact of the LBOs.

[16] Product category fixed effects are at the level of BLS "entry level items." The sample includes items from approximately 75 grocery categories, ranging from breakfast cereal to eggs to laundry and cleaning products.

[17] Store characteristics also include metropolitan area fixed effects in specifications that exclude ZIP code fixed effects.

Estimates of $\alpha$ are identified off of a large fraction of the sample – 41 percent of the sample corresponds to LBO firms.[18] However, simple OLS significance levels may be overstated, because the observations are likely not independent. Ideally, I would cluster the standard errors in equation (1) at the firm level, but clustering is computationally infeasible (given the number of observations within each firm and the computer resources available to me at the BLS). However, I can estimate the severity of the bias.

I estimate a model similar to equation (1) but without controls. Then, I re-estimate the model at the firm level, weighting by the number of observations underlying each firm. The firm level is the natural level of variation in $LBO_j$. By construction, these two estimates of $\alpha$ are equal, but they have different standard errors. The order of magnitude difference in the standard errors provides an approximate correction factor for standard error estimates for $\alpha$ in equation (1). While this method is not ideal, it allows me to confirm with minimal computational requirements whether estimates with uncorrected $t$-statistics of 10 or 15 are indeed statistically significant.

Price setting and the degree to which a firm invests in product availability may be influenced by non-financial factors that are correlated with undertaking an LBO. For example, local market competition, which is associated with higher prices and stockout rates, may be a negative predictor of LBO activity. To control for such factors, I also analyze a second, smaller sample of supermarket LBOs for which data on prices and product availability are available both before and after the LBO took place.

Using item-quote-level data from January 1988 to June 2005, I estimate same-store changes in prices and product availability leading up to and following the transaction. I define three indicator variables: $PreLBO_{jt}$ turns on for the two years before an LBO is announced; $AnnounceLBO_{jt}$ turns on after the transaction is announced but before it takes effect; and $PostLBO_{jt}$ turns on after the LBO takes effect. While $PostLBO_{jt}$ reveals the impact of the leveraged buyout, $PreLBO_{jt}$ represents a Granger-type test for pre-existing trends.

$$STOCKOUT_{isjt} = \alpha_1 PostLBO_{jt} + \alpha_2 AnnounceLBO_{jt} + \alpha_3 PreLBO_{jt} + X_{is}^{p}\beta + \omega_s + \psi_t + \nu_{isjt}$$

[18]As compared to the industry, LBO firms are over-represented in the CPI sample, which is focused primarily in densely-populated areas.

The $X^p_{is}$ matrix of product and store characteristics is $X^c_{is}$ plus it includes primary supplier (warehouse) fixed effects. $\omega_s$ are store fixed effects.

Similar to the cross-sectional approach, estimates of $\alpha_1$ are identified off of a large fraction of the sample – 26 percent. Appendix Table 2.A1, Panel B, reports the means for the other LBO timing variables. These regressors of interest vary at the firm-month level. To assess the degree of bias in the estimated standard errors caused by correlation across products, I implement a procedure similar to what I did in the cross-sectional approach. But given the panel nature of the data, I run the collapsed regression at the firm-month level.

## 2.5 Impact of LBOs on supermarket prices and product availability

In this section, I present evidence that product availability decreases when supermarket firms undertake a leveraged buyout. The LBOs also lead supermarkets to raise retail prices. Both effects are long-lived, lasting on average 10 years or more.

Table 2.2, Panel A, presents results from the cross-sectional analysis. Controlling only for characteristics of the products sampled, LBO firms charge 5.8 percent higher prices than other supermarkets (Column 2). Including metropolitan area fixed effects and other store characteristics decreases the coefficient to 2.7 percent (Column 3). This estimate is consistent with the existing literature: using firm-level scanner price data from 1992, Chevalier (1995b) estimates an average price difference of 3.1 percent between LBO and non-LBO firms (p. 1109).[19] Controlling for local market condition with ZIP code fixed effects decreases the estimate slightly to 2.4 percent (Column 4).

These estimates are generally statistically significant. While correlation in the error term (within a firm across products and over time) likely leads the simple OLS standard errors to be understated, the estimates in Columns (2) and (3) are statistically significant even after taking this correlation into account. Column (1) presents estimates from regressions without controls at both the firm-item-month and firm levels. The standard error is 11 times greater in the collapsed regression. While this may seem high, it reflects the

[19] Chevalier finds the largest price increases in health and beauty aids, which are largely absent from my sample. Excluding these products would likely align our estimates even more.

difference between the number of observations in the each specification (1.5 million versus 251). Coming from a regression that excludes controls, 11 is a relatively conservative correction factor for the estimates of $\alpha$. Applying this correction to the standard errors reported in Columns (2) and (3), the coefficient estimates are still statistically significant at the 10 percent level. The estimate from a regression with ZIP code fixed effects is no longer statistically significant, likely due to the degrees of freedom lost by estimating so many fixed effects (Column 4).

Product availability is also lower at LBO firms. Controlling only for product characteristics, LBO firms have 27 basis points more out-of-stocks than other supermarkets (Column 6). Including metropolitan area fixed effects and other store characteristics, the coefficient increases to 36 basis points (Column 7). Controlling for local market condition with ZIP code fixed effects, the estimate increases further to 57 basis points (Column 4). Relative to the sample mean of 3.7 percent, these estimates correspond to increases in out-of-stocks of 7 to 16 percent. The regressions reported in Column (5) imply that the estimated standard errors should be multiplied by 4.7 to account for within-firm heteroscedasticity. Using this method, the coefficient estimate reported in Column (6) is statistically significant at the 6 percent level, but the estimates in Columns (7) and (8) are not statistically significant at conventional confidence intervals.

Panel analysis confirms these finding: firms increase prices and decrease availability after undertaking an LBO. These results are reported in Table 2, Panel B. While fewer LBOs occurred during the panel analysis period, pre-LBO data on prices and product availability allow me to estimate the impact of LBOs using within-firm, within-store, and even within-item variation.

Leveraged buyouts lead firms to raise prices. While regressions with firm fixed effects do not detect an effect (Column 10), same-store analysis finds that firms raise prices by 1.0 percent following an LBO (Column 11). Including item fixed effects decreases the estimate slightly to 0.9 percent (Column 12). Both of these estimates are statistically significant even after accounting for within firm-month heteroscedasticity and economically significant in an industry where margins are typically about 1 percent (Bongard and Cross 1992). Analysis of pre-existing trends yields mixed results. Estimates suggest these firms may have lowered prices immediately preceding the LBO (Column 10), but these results go away when store fixed effects are included (Column 11), and looking only at the same items before and after the transaction suggests firms may have actually slightly raised prices leading up to the LBO (Column 12).

Leveraged buyouts also lead firms to reduce product availability. Estimates of the rise in the stockout rate following an LBO are 56 basis points in regressions with firm fixed effects (Column 14) and 40 basis points in regressions with store fixed effects (Column 15). Relative to the sample mean, these correspond to a 13 percent and a 9 percent increase, respectively. These estimates are highly statistically significant even after accounting for within firm-month heteroscedasticity. Including item fixed effects decreases the estimate to 16 basis points (corrected $p < 0.08$; Column 16). Any particular item is sampled by the BLS for at most 5 years, so same-item estimates may be smaller because they reflect short-run effects. In fact, further analysis presented below finds that the LBO effects intensify over the first couple years after the buyout. An analysis of pre-existing trends shows no evidence of these firms, stores, or items having unusually high or low levels of availability in the two years before the transactions were announced. The lack of a pre-existing trend suggests both that the decrease in availability was caused by the LBOs and that unusually low levels of out-of-stocks were unlikely to have triggered the transactions.[20]

The impact of leveraged buyouts on supermarket pricing and product availability is long-lasting. Figure 2-2 depicts the timing of the LBO effects. I graph estimated regression coefficients associated with annual indicator variables for the three years prior to an LBO through the 10 years after the transaction. These estimates are from analyses that control for store and year-month fixed effects as well as product characteristics. The results for price levels are in Panel A, and the results for out-of-stocks are in Panel B. On average, prices and stockouts begin to increase within a year following an LBO, and they stay high for many years. This is consistent with the general pattern of LBOs in the supermarket industry. While LBO firms often use the proceeds from asset sales to moderately reduce leverage in the years following an LBO, IPOs and other reverse LBOs are rare. According to the *SDC Merger & Acquisitions* database, only 4 of the 34 supermarket LBOs undertaken since 1988 are associated with a subsequent IPO.

The differential impacts of LBOs on products sourced using different distribution chan-

[20]The test for pre-existing trends with fixed effects assesses whether the level of out-of-stocks differed from the historical level for that firm/store/item immediately prior to the transaction. It does not compare the historical level of out-of-stocks at LBO firms to other firms in the industry.

nels provides a useful robustness check. Most grocery products are distributed by manufacturers to retail stores through warehouses, which are owned either by the retailer or an independent operator. Other products are delivered directly to retail stores, bypassing the retailers' warehouses, and shelf inventory for these "direct store delivery" (DSD) items is usually managed jointly by the manufacturer's distributor and the retailer. Major DSD categories include carbonated drinks, bread, snacks, cookies, and crackers.

A major difference between warehouse and DSD distribution is that the manufacturer's distributor typically plays the lead role in store-level category management, merchandising, and managing the shelf inventory for DSD products.[21] The degree of the retailers' involvement in inventory management depends on the replenishment arrangement in place with the distributor. While the retailer always "checks-in" and approves deliveries when they arrive, the retailer often does not initiate orders. In some arrangements, the route driver doubles as a salesperson who monitors the shelf and replenishes as needed; in others, an account manager for the distributor forecasts sales requirements and works with the store manager to develop an order for the store. Because DSD distributors often have a good deal of autonomy over replenishment, leveraged buyouts likely have less of an impact on DSD as compared to warehouse items. On the other hand, while a DSD distributor may suggest a retail price, the retailer has ultimate control over retail price determination. Therefore, LBOs are unlikely to have a differential impact on prices for the two sets of products.

I estimate the impact of LBOs separately for product categories that are typically distributed through a warehouse versus direct store delivery. Cross-sectional regressions with ZIP code fixed effects and panel regressions with store fixed effects are reported in Table 2.3. The results for warehouse-supplied categories, presented in Panel A, are similar to the results for the full sample for both prices and out-of-stocks: focusing on the panel estimates, prices increase by 1 percent and stockouts increase by 49 basis points after an LBO (Columns 2 and 4). The results for DSD categories are reported in Panel B. While the estimate of the average price increase for items in DSD categories is imprecise (Column 6), the prevalence of out-of-stocks does not increase following an LBO. The point estimate is negative and not statistically significant (Column 8).

[21]There are also differences in the nature of products chosen to be distributed through each channel. Compared to warehouse-supplied items, DSD products tend to have shorter shelf life, higher volume, higher promotional intensity (and demand variability), lower value density, and greater merchandising difficulty due to greater weight or fragility.

I also address the impact of LBOs on supermarket product variety – an important issue related to product availability. On the supply side, SKU proliferation contributes to the growth in out-of-stocks by increasing the size and complexity of the firm's forecasting problem. And on the demand side, an out-of-stock may not be as costly for consumers when a store offers greater product variety. If LBOs lead firms to increase product variety, then the measured reductions in product availability may not actually represent a decrease in product quality. However to the contrary, evidence seems to suggest that LBOs may decrease product variety.

While I do not have direct measures of product variety, I construct a rough measure of changes in product variety from the *CPI Commodity and Services Survey.* Just as I can observe when a surveyor attempts to sample an item but encounters an out-of-stock, I can observe when the sampled item is instead discontinued by the outlet (that is, marked as "permanently" unavailable). Product cancellations are rare for food items in the CPI sample, occurring in only 0.6 to 1.4 percent of attempted price quotes.[22] While this sort of product cancellation rate is far from perfect, I analyze it as a possible proxy for changes in product variety.[23]

Average changes in the product cancellation rate before and after an LBO are presented in Figure 2-2, Panel C. There is no evidence of a decrease in product cancellations just after the LBOs. To the contrary, product cancellations increase in the years following the transaction. Same-store regression analysis similar to specifications reported in Table 2.2, Columns (11) and (15), finds that the product cancellation rate increases by 8 basis points after an LBO (corrected $p < 0.03$). At the mean, this corresponds to a 12.5 percent increase.

[22]The range on this estimate corresponds to exactly how it is constructed. As explained in footnote 12, there is a possibility that the data confuse out-of-stocks with product cancellations and vice versa. Considering all attempts to sample an item yields a product cancellation rate of 1.4 percent. As a conservative alternative, I only consider attempts for which the item was successfully sampled in the preceding month. This yields a lower estimate of 0.6 percent, and I use this algorithm to construct the variable in the analysis that follows.

[23]The product cancellation rate is an asymmetric measure of product variety. It directly measures product cancellations, but not new product introductions. When I measure an increase in product cancellation, it may be that the firms are actually changing their product mix while increasing product variety.

## 2.6 Conclusion

The effects of leveraged buyouts illustrate how debt impacts the management and operation of a firm. Once a company becomes highly leveraged, a focus on servicing debt obligations has a widespread impact on the firm. The findings presented in this chapter show that leveraged buyouts, on average, reduce a supermarket's provision of product availability – an important dimension of product quality in the retail sector. The estimates suggest that leveraged buyouts increase the incidence of retail out-of-stocks by about 10 percent. Using industry estimates for the mean out-of-stock rate, this corresponds to an increase from about 8 to 9 percent.

While leveraged buyouts seem to increase out-of-stocks at supermarket firms, it is unclear whether this increase is beneficial to the retailer. Without additional evidence, it is impossible to know whether the increase in out-of-stocks represents a move towards or away from the optimal rate. The pressures of servicing large debt obligations may lead the firms to cut profitable investments in product availability, or if agency problems had led firms to have too few out-of-stocks before the buyout, then the increase in out-of-stocks may actually represent a move towards the optimal rate. Either way, the fact that stockouts increase does not imply the buyouts were a mistake, because operating profit improvements from the asset sales, increased prices, or even improvements in other dimensions of productivity may have made the deals worthwhile. The findings in this chapter do imply, however, that it is important for firms to consider these sorts of real-side effects on the firm's operations when setting financial policy.

## References

Andersen Consulting (1996). *Where to Look for Incremental Sales Gains: The Retail Problem of Out-of-Stock Merchandise.* Coca-Cola Retailing Research Council.

Angrist, J. D. (2004). Treatment effect heterogeneity in theory and practice. *Economic Journal* 114, C52-83.

Bils, M. (2005). Deducing markup cyclicality from stockout behavior. University of Rochester mimeo.

Bongard, V. and S. Cross (1992). *The Grocery Industry: Past, Present & Future.* Murray

Hill, NJ: Duns Analytical Services.

Calomiris, C. W., C. P. Himmelberg, and P. Wachtel (1995). Commercial paper, corporate finance, and the business cycle: A microeconomic perspective. NBER working paper 4848.

Carpenter, R. E., S. M. Fazzari, and B. C. Petersen (1994). Inventory investment, internal-finance fluctuations, and the business cycle. *Brookings Papers on Economic Activity* 0, 75-122.

Chevalier, J. A. (1995a). Capital structure and product-market competition: Empirical evidence from the supermarket industry. *American Economic Review* 85 (3), 415-35.

Chevalier, J. A. (1995b). Do LBO supermarkets charge more? An empirical analysis of the effects of LBOs on supermarket pricing. *Journal of Finance* 50 (4), 1095-1112.

Chevalier, J. A. and D. S. Scharfstein (1996). Capital-market imperfections and countercyclical markups: Theory and evidence. *American Economic Review* 86 (4), 703-25.

Dorgan, Tim (1997). The awful truth. *Progressive Grocer* 76 (3), 75.

Gruen, T. W., Corsten, D. S., and Bharadwaj, S. (2002). *Retail Out-of-Stocks: A Worldwide Examination of Extent, Causes, and Consumer Responses.* Washington, DC: Grocery Manufacturers of America.

Harris, R., D. S. Siegel, and M. Wright (2005). Assessing the impact of management buyouts on economic efficiency: Plant-level evidence from the United Kingdom. *The Review of Economics and Statistics* 87, 148-53.

Hubbard, R. G. (1998). Capital-market imperfections and investment. *Journal of Economic Literature* 36, 193-225.

Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review* 76, 323-29.

Jensen, M. C. (1989). Eclipse of the public corporation. *Harvard Business Review* 67, 61-74.

Jensen, M. C. (1993). The modern industrial revolution, exit, and the failure of internal control systems. *Journal of Finance* 48, 831-80.

Kaplan, S. (1989). The effects of management buyouts on operating performance and value. *Journal of Financial Economics* 24, 217-54.

Kashyap, A. K., O. A. Lamont, and J. C. Stein (1994). Credit conditions and the cyclical behavior of inventories. *Quarterly Journal of Economics* 109, 565-92.

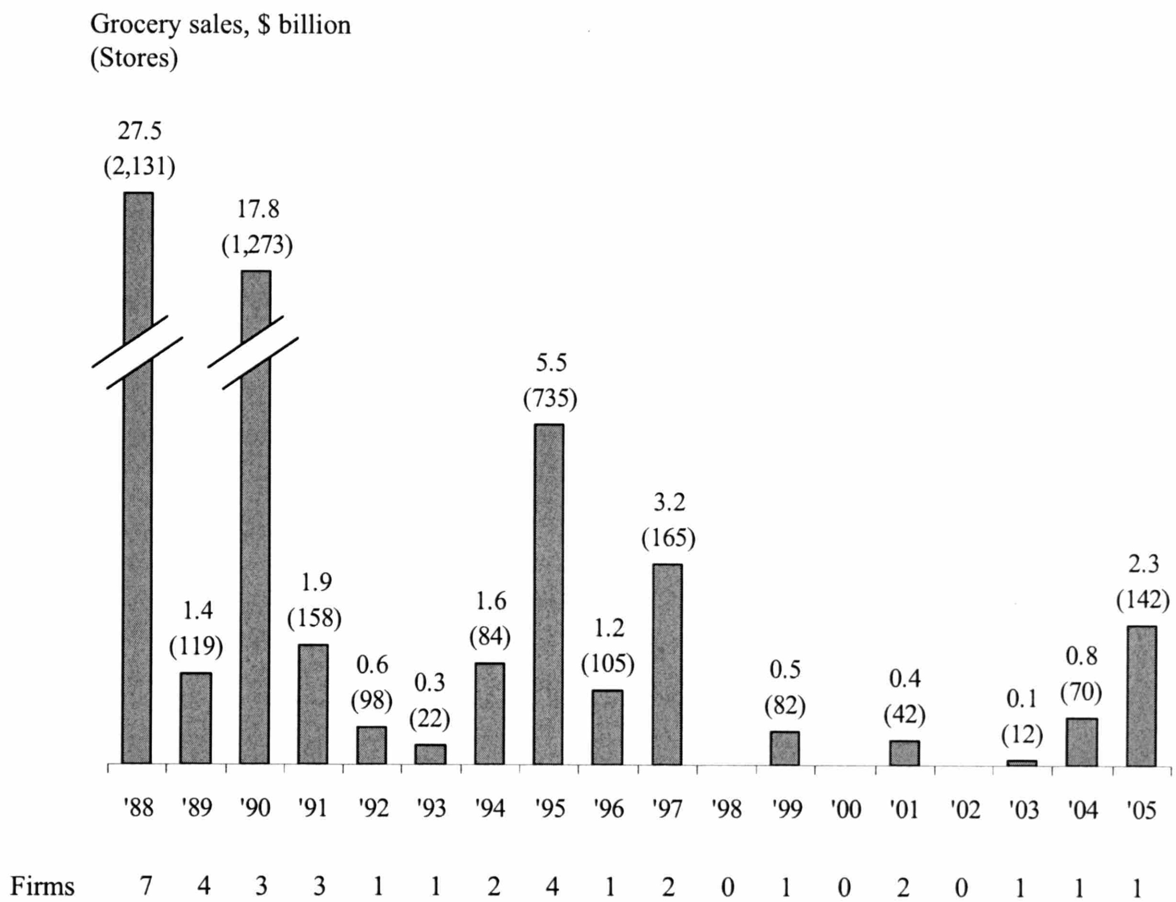
Lichtenberg, F. R. and D. Siegel (1990). The effects of leveraged buyouts on productivity and related aspects of firm behavior. *Journal of Financial Economics* 27, 165-94.

Long, W. F. and D. J. Ravenscraft (1993). LBOs, debt and R&D intensity, *Strategic Management Journal* 14 (Special issue), 119-135.

Roland Berger Strategy Consultants (2002). *Full-Shelf Satisfaction—Reducing Out-of-Stocks in the Grocery Channel: An In-Depth Look at DSD Categories.* Washington, DC: Grocery Manufacturers of America.

Smith, A. J. (1990). Corporate ownership structure and performance: The case of management buyouts. *Journal of Financial Economics* 27, 143-64.

Figure 2-1: Supermarket leveraged buyouts, 1988-2005

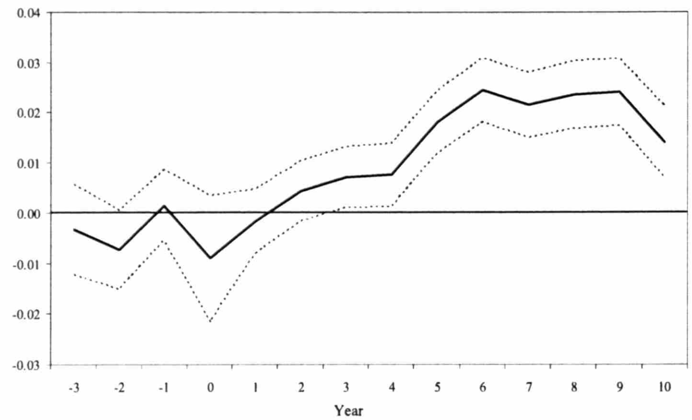



*Source*: SDC Platinum

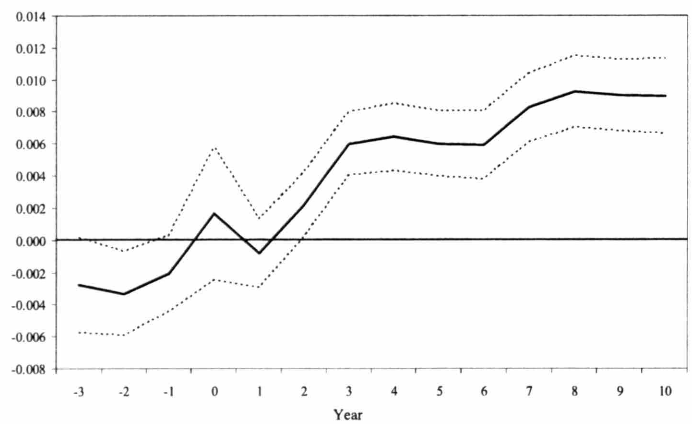
*Note*: This figure depicts the timing of supermarket LBOs undertaken between 1988 and 2005. The subset of these firms included in the CPI sample provide identification for the panel empirical approach.

Figure 2-2: Effect of LBOs on prices, out-of-stocks, and product cancellations, 1988-2005

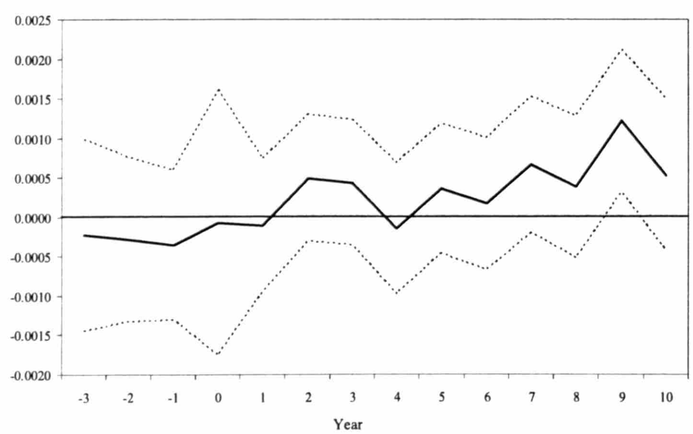
A. Log (price / ounce)




B. Out-of-stock




C. Product cancellation




*Note*: Each figure depicts regression coefficients and 95 percent confidence intervals from regressions with different dependent variables. The regressions estimate the timing of LBO effects on each dependent variable, controlling for product & time characteristics (seasonal item indicator, fixed effects for product category, day of week, year-month), store characteristics (store size, distance from primary supplier, vertically integrated with warehouse, independent indicator), and store fixed effects.

Table 2.1: Summary statistics, LBO and non-LBO firms, 1990

| | Firms that LBO before 1990 | Firms that LBO after 1990 | Firms that do not LBO |
|---|---|---|---|
| *A. Firm characteristics:* | | | |
| Stores | 171.8 | 47.2 | 26.0 |
| | (298.3) | (48.0) | (88.6) |
| States with at least one store | 4.13 | 1.95 | 1.85 |
| | (5.79) | (1.43) | (2.88) |
| *B. Store characteristics:* | | | |
| Grocery selling space (1,000 sq ft) | 29.7 | 30.4 | 20.3 |
| | (14.0) | (19.3) | (13.5) |
| Weekly grocery volume ($1,000) | 238.0 | 216.5 | 150.5 |
| | (153.3) | (169.9) | (133.1) |
| Employment (FTEs) | 70.7 | 63.1 | 43.8 |
| | (48.8) | (58.6) | (41.0) |
| Checkout counters | 8.84 | 8.79 | 6.92 |
| | (3.60) | (4.03) | (3.95) |
| *C. Performance:* | | | |
| Out-of-stock rate | 0.0371 | 0.0355 | 0.0356 |
| | (0.1889) | (0.1851) | (0.1854) |
| Average price level (¢ per ounce) | 21.32 | 23.29 | 18.71 |
| | (140.49) | (161.07) | (96.48) |

*Note:* Standard deviations reported in parentheses.

Table 2.2: Do LBO firms behave differently than their competitors?

*A. Cross-section of firms, January 1988 - June 1993*

| | Log (price / ounce) | | | | Out-of-stock | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| LBO firm | 0.0485<br>(0.0017)<br>[0.0188] | 0.0581<br>(0.0010) | 0.0273<br>(0.0015) | 0.0239<br>(0.0024) | 0.0022<br>(0.0003)<br>[0.0014] | 0.0027<br>(0.0003) | 0.0036<br>(0.0005) | 0.0057<br>(0.0008) |
| Observations | 1,481,008 | 1,481,008 | 1,481,008 | 1,481,008 | 1,518,906 | 1,518,906 | 1,518,906 | 1,518,906 |
| R-squared | 0.0006 | 0.6405 | 0.6464 | 0.6698 | 0.0000 | 0.0230 | 0.0263 | 0.0383 |
| Product & time characteristics | | X | X | X | | X | X | X |
| Store characteristics | | | X | X | | | X | X |
| ZIP code fixed effects | | | | X | | | | X |

*B. Panel of firms, January 1988 - June 2005*

| | Log (price / ounce) | | | | Out-of-stock | | | |
|---|---|---|---|---|---|---|---|---|
| | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) |
| Post-LBO | 0.0528<br>(0.0011)<br>[0.0023] | 0.0025<br>(0.0019) | 0.0104<br>(0.0020) | 0.0093<br>(0.0007) | 0.0014<br>(0.0002)<br>[0.0003] | 0.0056<br>(0.0006) | 0.0040<br>(0.0007) | 0.0016<br>(0.0007) |
| Announcement period | -0.0738<br>(0.0102)<br>[0.0216] | -0.0163<br>(0.0064) | -0.0040<br>(0.0062) | -0.0054<br>(0.0021) | -0.0018<br>(0.0020)<br>[0.0025] | 0.0038<br>(0.0020) | 0.0030<br>(0.0021) | 0.0007<br>(0.0020) |
| 2 years before announcement | -0.0699<br>(0.0035)<br>[0.0074] | -0.0153<br>(0.0027) | 0.0029<br>(0.0028) | 0.0045<br>(0.0010) | -0.0044<br>(0.0007)<br>[0.0009] | -0.0001<br>(0.0008) | -0.0010<br>(0.0009) | -0.0013<br>(0.0010) |
| Observations | 4,903,217 | 4,903,217 | 4,903,217 | 4,903,217 | 5,033,030 | 5,033,030 | 5,033,030 | 5,033,030 |
| R-squared | 0.0006 | 0.6557 | 0.6852 | 0.9680 | 0.0000 | 0.0362 | 0.0503 | 0.2102 |
| Controls & firm fixed effects | | X | X | X | | X | X | X |
| Store fixed effects | | | X | X | | | X | X |
| Item fixed effects | | | | X | | | | X |

*Note:* Controls (where indicated) include product & time characteristics (seasonal item indicator, fixed effects for product category, day of week, year-month) and store characteristics (store size, distance from primary supplier, vertically integrated with warehouse, independent indicator, metropolitan area fixed effects). Standard errors in brackets are from weighted regressions at the chain level in Panel A and at the chain-month level in Panel B. These regressions have 251 observations in columns (1) and (4), 56,732 observations in column (7), and 56,410 observations in column (10).

Table 2.3: How do LBOs affect warehouse-supplied versus direct-store-delivery items?

*A. Warehouse-supplied categories*

| | Log (price / ounce) | | Out-of-stock rate | |
|---|---|---|---|---|
| | (1) Cross-section 1988-1992 | (2) Panel 1988-2005 | (3) Cross-section 1988-1992 | (4) Panel 1988-2005 |
| Post-LBO | 0.0316 | 0.0106 | 0.0062 | 0.0049 |
| | (0.0025) | (0.0021) | (0.0008) | (0.0007) |
| Announcement period | | -0.0044 | | 0.0026 |
| | | (0.0065) | | (0.0021) |
| 2 years before announcement | | 0.0057 | | -0.0008 |
| | | (0.0029) | | (0.0010) |
| Observations | 1,379,289 | 4,547,574 | 1,415,203 | 4,672,528 |
| R-squared | 0.6635 | 0.6784 | 0.0402 | 0.0525 |

*B. Direct-store-delivery categories*

| | Log (price / ounce) | | Out-of-stock rate | |
|---|---|---|---|---|
| | (5) Cross-section 1988-1992 | (6) Panel 1988-2005 | (7) Cross-section 1988-1992 | (8) Panel 1988-2005 |
| Post-LBO | -0.0344 | 0.0039 | -0.0003 | -0.0041 |
| | (0.0063) | (0.0050) | (0.0031) | (0.0025) |
| Announcement period | | -0.0020 | | 0.0146 |
| | | (0.0159) | | (0.0077) |
| 2 years before announcement | | -0.0096 | | -0.0047 |
| | | (0.0072) | | (0.0035) |
| Observations | 101,719 | 355,643 | 103,703 | 360,502 |
| R-squared | 0.8475 | 0.8606 | 0.0463 | 0.0575 |

*Note:* Items in product categories typically distributed by DSD are included in Panel B; these categories are carbonated drink, bread, snacks, cookies, and crackers. Items in all other product categories are included in Panel A. The cross-section analyses include ZIP code fixed effects and correspond to Table 1.2, Columns (4) and (8). The panel analyses include store fixed effects and correspond to Table 1.2, Columns (11) and (15).

Appendix Table 2.A1: Leveraged buyout samples

*A. Comparison of Chevalier and SDC samples, matched to the 1990 RSD (stores, $ billion sales)*

| Chevalier | SDC Platinum: Non-LBO | SDC Platinum: LBO |
|---|---|---|
| Non-LBO | 26,236 (82%) | 347 (1%) |
| | 214 (75%) | 5 (2%) |
| LBO | 1,004 (3%) | 4,482 (14%) |
| | 12 (4%) | 54 (19%) |

*B. Summary statistics for CPI microdata, matched to Chevalier and SDC samples*

| | Obs | Mean |
|---|---|---|
| *Chevalier sample (matched to 1988-2002):* | | |
| LBO in 1990 | 1,532,805 | 0.410 |
| *SDC Platinum sample (matched to 1988-2005):* | | |
| Post-LBO | 5,620,827 | 0.261 |
| Announcement period | 5,620,827 | 0.002 |
| 2 years before announcement | 5,620,827 | 0.018 |

# Chapter 3

# Does Malpractice Liability Keep the Doctor Away? Evidence from Tort Reform Damage Caps

*Increasingly, Americans are at risk of not being able to find a doctor when they most need one. Doctors have given up their practices, limited their practices to patients who do not have health conditions that are more likely to lead to lawsuits, or have moved to states with a fairer legal system where insurance can be obtained at a lower price.*

U.S. Department of Health and Human Services
March 3, 2003

## 3.1 Introduction

Medical malpractice litigation has become increasingly prevalent in the United States, with liability payments totaling $4 billion in 2001, up 42 percent since 1992. In theory, the threat of liability deters iatrogenic (physician-induced) injury. While evidence on deterrence is inconclusive, the system has been shown to affect physician behavior in undesirable ways, such as inducing inefficient provision of care (Kessler and McClellan, 1996). Recently, the United States Department of Health and Human Services (HHS, quoted above) and others have claimed that malpractice liability also restricts physician supply and access to care. This study investigates the validity of such claims.

To explore this issue, I analyze reforms of the liability system. In response to growing jury awards, settlements, and insurance premiums, many states have introduced tort reforms. One popular initiative is a cap on damage awards. Thirty-three state legislatures have passed laws limiting malpractice damage awards, five of which were adopted since 2002. Proponents argue that such limits boost physician supply — a claim that was featured prominently in the 2004 presidential campaign.[1] Research on the effects of such reforms on physician supply is crucial for state and federal policymakers considering similar initiatives and provides evidence on how malpractice litigation affects health care markets more broadly.

In theory, the effects of caps depend on market-specific factors. An effective damage cap reduces a physician's costs, lowering both insurance premiums and uninsurable costs. If prices of medical services do not fully adjust or if demand is sufficiently elastic, local physician net incomes will increase, attracting additional entry. However, if price changes offset cost changes but demand is inelastic, damage caps will not affect physician supply. Although the health care market demand elasticity is generally believed to be low in most of the United States, it may be greater in the most rural areas, where potential customers (i.e., patients) are more likely to be uninsured (Ormond et al. 2000). Furthermore, malpractice liability may engender greater pressure on physicians in rural areas where I demonstrate that non-liability costs of practice are lower, physician-to-population ratios are smaller, and rates of malpractice claims per doctor are greater than in metropolitan areas.

I use state adoption of damage caps to identify changes in local malpractice climates. Based on data representing the universe of physician malpractice claims in the 1990s, my results suggest that damage caps reduce damage payments by 24 percent. These estimates align with existing evidence from previous decades (Danzon 1986; Zuckerman et al. 1990; Sloan et al. 1989). Then, using three decades of data on physician-to-population ratios, I estimate the effect of these reforms on physician supply. I find that malpractice caps do not increase physician supply for the average American, but they do increase total physician supply in the least densely-populated areas by 3 to 5 percent. This effect appears to be driven by a relative increase in the supply of specialists by 10 to 12 percent, with no effect

[1] For example, in remarks delivered at a Little Rock, Arkansas, hospital on January 26, 2004, President George W. Bush maintained: "One of the reasons...it's hard to find a doc these days, is because frivolous and junk lawsuits are threatening medicine across the country....For the sake of making sure health care is accessible and affordable, we need a $250,000 cap on non-economic damages."

on the supply of general practice physicians.

The remainder of the chapter is organized as follows. Section 3.2 describes state tort reforms and presents evidence on their effects on damage payments and liability premiums. Section 3.3 develops a theoretical framework for the effects on physician supply, and Section 3.4 describes the data, empirical approach, and primary results. Possible explanations of the urban-rural differential effect are explored theoretically and empirically in Section 3.5. Section 3.6 discusses the welfare implications, and Section 3.7 concludes.

## 3.2 Tort reform: Background and proximate impact

Beginning during the liability insurance "crisis" of the mid-1970s, state governments have enacted a variety of tort reform measures. Motivated by a popular perception that juries were overly generous to plaintiffs,[2] reforms were aimed at (1) reducing the frequency of malpractice claims, (2) reducing the amounts recoverable, and/or (3) curbing the costs of the legal process. Examples include altering the statute of limitations, placing limits on attorney contingency fees, instituting mandatory medical review or pretrial screening panels, and enacting caps on damage awards.

Since 1975, thirty-three states have passed laws limiting medical malpractice liability damage awards.[3] Eight states have added or stiffened caps on noneconomic damages since 2002, and the President has proposed a nationwide cap on noneconomic damages. This analysis focuses on this sort of cap on noneconomic damages. More restrictive caps are also included in my analysis—specifically the caps on total damages in Indiana, Louisiana, New Mexico, and Virginia—and less restrictive caps are excluded.[4] Some of these laws are specific to iatrogenic injury cases, whereas others apply to general liability. States cap noneconomic damages at different levels, ranging from $250,000 to $875,000, with some indexed to inflation and others not. As there is substantial legislative heterogeneity across states, my analysis measures the average effect.[5]

[2] Although popular perception is that jury awards, particularly for pain and suffering, are random and capricious, Viscusi and Born (2004) claim "there is no general empirical evidence to that effect." In fact, analysis in Viscusi (1996) suggests they are not as random as is often thought.

[3] Seven of which were overturned in state courts.

[4] Minnesota law limits the award of damages for intangible losses (loss of consortium, emotional distress, or embarrassment) but not other noneconomic damages, such as pain and suffering. Nebraska law limits total damages recoverable against a limited subset of health care providers.

[5] State laws vary in other dimensions as well. For example, the Michigan cap does not apply to cases

Existing evidence suggests that liability caps reduce damage awards by 23 to 48 percent. Danzon (1986) compares states with caps to those without and finds that damage caps are associated with 23 percent lower damage awards but no change in the frequency of malpractice claims. Using state-year panel data based on a survey of insurers, Zuckerman, Bovbjerg, and Sloan (1990) find that noneconomic damage caps reduce malpractice payments by 48 percent and do not affect frequency.[6] Based on closed claims data, Sloan, Mergenhagen, and Bovbjerg (1989) conclude that, on average, noneconomic (total) damage caps reduce malpractice payments by 31 (38) percent.[7]

I employ data derived from the National Practitioner Data Bank (NPDB) to verify this relationship using more recent closed claims data. The NPDB is a comprehensive clearinghouse of information that records, among other things, all malpractice liability payments made on behalf of a health care provider in the United States. Congress created the data bank as part of the Health Care Quality Improvement Act of 1986 to retard a physician's ability to move from state to state without disclosing his malpractice record. Payment reporting began in September 1990.

The NPDB is a census of medical malpractice payments made on behalf of any health care provider in the United States.[8] It includes all payments—whether made as a result of a judgment or a settlement. Any medical malpractice payer that fails to report medical malpractice payments is subject to an $11,000 fine for each payment not reported. Although the database covers all types of licensed health care practitioners, I restrict my sample to

where there is loss of life.

[6] The authors cannot precisely estimate the effect of caps on total damages on malpractice payment severity; their point estimates of 20-30 percent cannot statistically be distinguished from either 0 or 48 percent.

[7] A recent study of jury verdicts in California find that that state's cap on noneconomic damage awards was imposed in 45 percent of trials resulting in a plaintiff verdict, reducing defendants' liabilities by 30 percent (Pace, Golinelli, and Zakaras 2004).

[8] The NPDB is the most comprehensive database of malpractice payments available. It includes all malpractice payments on behalf of licensed health care practitioners except those made by practitioners themselves from their own personal funds. Thus, if a malpractice insurer or a self-insured entity makes a malpractice payment on behalf of a practitioner, the law requires that it be reported. A self-insured entity includes incorporated practices. Data is not available on how many physicians self-insure; the number is believed to be small but growing (GAO 2003a). Another potential deficit in the database is that, in the case of settlement, it only requires the reporting of health care practitioners named in the final agreement. Therefore, payments may not be recorded in the NPDB when a doctor's name is removed from the claim before (or allegedly as part of) the final settlement. Industry participants reportedly call it the "corporate shield." After the doctor's name is removed from the claim, only the hospital or another corporate entity is identified as the responsible party. Speculation suggests use of the corporate shield may be widespread (Hallinan 2004). However, there is no obvious reason why use of the corporate shield would vary systematically with the adoption of damage caps or across areas with different population density.

licensed non-Federal physicians, including allopathic physicians (MDs), osteopathic physicians (DOs), interns, and residents.[9]

Analysis of NPDB data is consistent with results reported in the literature for earlier decades; damage awards are lower in states with laws limiting liability than in other states. Despite their accuracy and completeness, NPDB data have a major drawback in the context of this study. There are too few changes in state law since the initiation of the NPDB to identify a within-state effect.[10] Results from cross-sectional regressions are in Table 3.1, where each column represents a regression with a different dependent variable, measured in logs. Each regression includes year fixed effects and demographic controls, but without law changes, identification is achieved from cross-sectional variation. States with caps have 24 percent lower damage payments per capita than other states, driven by differences in severity but not frequency.[11] Damage awards are 28 percent lower on average and 24 percent lower at the median in states with laws limiting awards, but these states have a similar number of claim payments per capita.

Existing evidence suggests that liability caps have a more modest effect on liability premiums. Zuckerman, Bovbjerg, and Sloan (1990) find that caps on total damages decrease insurance premiums for all specialties by 13 to 16 percent in the year following enactment and significantly more in the long-run, but that caps on only noneconomic damages have limited effect. In contrast, administrative data from the National Association of Insurance Commissioners suggest that premiums in states with noneconomic damage reform are 6 to 17 percent lower than in other states (Viscusi and Born 2004; Thorpe 2004). Recent rate surveys also find an effect in the cross-section. The Medical Liability Monitor (2002) reports that of the nineteen states with caps on noneconomic damages in 2002, premiums were below the national average in all but one state for internists (Michigan) and all but three states for general surgeons (Michigan, Missouri, Utah) and obstetricians/gynecologists (Michigan, Massachusetts, Maryland).

Damage caps may also affect physician supply. To my knowledge, Danzon, Pauly,

[9] Nationally, DOs represent 5.6 percent of physician malpractice payments reported in the NPDB. Residents and interns represent only 0.9 percent.

[10] Since the NPDB began collection in 1990, only Montana and North Dakota adopted damage caps before 2002. For the five states adopting caps since 2002, there is not yet a sufficient post period to conduct a complete analysis.

[11] The cross-sectional design of the analysis of course does not resolve the potential concern that omitted variables drive these differences and not the damage cap itself.

and Kington (1990) were the first to consider the effect of malpractice liability on cross-market adjustments in physician stocks. They find no significant effects of changes in the malpractice climate—measured by claim frequency, claim severity, and liability premiums—on changes in physician stocks. Rather, they present evidence that physician fees reflect more than full pass-through of insurance costs and that net income is not affected by the malpractice environment. However, the source of identification in their analysis is unclear. An omitted factor, such as the prevalence of invasive medical procedures, may be driving both malpractice claims and the price of health care.

Hellinger and Encinosa (2003) revisit the question, focusing specifically on the effect of laws limiting malpractice liability awards. Using predominantly unweighted county-level analysis to compare growth in physician supply from 1970 to 2000 in states that adopt caps to states that do not, they find that damage caps increase physician supply by 12 percent. In a related study, Klick and Stratmann (2003) measure the effect of various state tort reform measures on physician supply. Their population-weighted results suggest damage caps have at most a limited effect on physician supply.[12]

These divergent results can be reconciled by considering geographic heterogeneity in the physician supply response to malpractice liability. Whereas the unweighted county-level research design employed by Hellinger and Encinosa (2003) measures the effect for the average *county*, the population-weighted analysis conducted by Klick and Stratmann (2003) measures the effect for the representative *individual.* Although neither study focuses on this explanation, both results are consistent with a supply response concentrated in the least densely-populated counties.[13] The magnitude of the physician supply response is particularly important in the most rural areas where baseline physician-to-population ratios are lowest, and policymakers from both sides of the aisle have expressed particular concern for access to care in rural areas.[14] Other policymakers and numerous press articles (e.g.,

[12]Table 3.5, columns 4-6, and Table 3.6, columns 3-4, present specifications that weight all physicians equally regardless of specialty. When the standard errors are clustered to allow for within state correlation over time, virtually all estimates associated with the adoption of a damage cap are likely not statistically different than zero. Furthermore, the point estimates are small in magnitude, suggesting that noneconomic (total) damage caps have a +0.6 (-0.8) percent impact on physician supply.

[13]In fact in a recent paper, Baicker and Chandra (2004) find that malpractice liability premiums are negatively correlated with rural physician supply but not with the size of the overall physician workforce. Given the potential endogeneity of liability premiums, it is difficult to assign their results a casual interpretation. Although their analysis is not dispositive of the effects of malpractice liability, it nicely complements the findings in this chapter.

[14]For example, on July 17, 2002, Congressman Ted Strickland (D-OH) stated in a hearing before the

Malcolm 1985) have focused on the supply response of specialist physicians, particularly obstetricians.[15,16] Using state adoption of damage caps to identify changes in local malpractice climates, I measure the response in physician supply and assess the extent to which this effect differs both across broad categories of physician specialties and across urban and rural areas.

## 3.3 Theoretical framework

Although the theoretical impact of damage caps on malpractice payments and liability premiums seems straightforward, the effect of caps on physician supply is less clear. It depends on the incidence of malpractice costs, which is a function of (1) the extent to which doctors pass on the costs of malpractice litigation to their patients in the form of higher fees and (2) the market elasticity of demand. Differences in these factors across geographies will drive positive effects of caps on physician supply in some areas and no effects in others.[17]

The following theoretical treatment extends Danzon, Pauly, and Kington (1990) to study the impact of damage caps. Let $M$ index the probability and expected size of a malpractice recovery, conditional on the occurrence of an injury, under a given legal regime. Then, as illustrated in Table 3.1, the adoption of limits on damage awards corresponds to a decrease

---

Subcommittee on Health:

> As a Representative of a rural area, I am particularly concerned about this issue. My District already suffers from chronic access problems and I am very worried that a malpractice crisis in which doctors simply cannot buy insurance would exacerbate this problem to the point of emergency.

And in a floor statement on March 18, 2003, Senator Orrin Hatch (R-UT) pronounced, "The [access to health care] crisis is particularly acute in the farming and ranching communities of rural America."

[15]For example, on April 6, 2004, Senate Majority Leader Bill Frist, M.D. (R-TN) stated on the Senate floor:

> While the crisis affects all people seeking access to quality care, it affects those who are seeking help from high risk specialist physicians the most.... Our litigation system is increasingly forcing needed medical specialty doctors like neurosurgeons and obstetricians to drop or limit their services, to move to states not in crisis or to simply retire early from the practice of medicine.

[16]While popular press reports motivate further analysis, they are far from dispositive. In fact, GAO (2003b) reviewed reports of physicians reducing certain high litigation risk services, such as spinal surgeries and mammograms. Analysis of Medicare data and contacts with physicians who were reportedly affected found that access to these services was not widely affected.

[17]In theory, a negative effect is also possible. Damage caps increase the costs borne by patients in the face of an adverse event, and thereby increase the full implicit price of medical care. To the extent that liability leads physicians to change their medical practice, damage caps may lower the quality of medical care as well. Depending on the demand elasticity and the magnitude of the response in physician fees and quality, patients may demand less medical care, attracting fewer physicians.

in $M$.

Malpractice litigation creates both insurable and uninsurable costs for physicians. Virtually all physicians hold insurance policies covering the direct monetary costs of malpractice litigation, including settlement and judgment payments and the cost of legal defense. Lawthers et al. (1992) report that, of New York State physicians sued for malpractice, 97.8 percent made no out-of-pocket damage payments and 93.6 percent paid no out-of-pocket attorney expenses.[18] The cost of insurance typically depends on the limits of coverage as well as basic, fixed characteristics of a physician's practice, such as specialty, whether high risk procedures (such as surgery or obstetrics) are performed at all, and practice location.[19] The liability premium, thus, represents a fixed cost for physicians and is increasing in $M$.[20]

Physicians also bear uninsurable costs from malpractice litigation, such as the time costs, forgone revenue, reputation damage, and mental anguish associated with defending a malpractice claim. For example, Lawthers et al. (1992) find that the median physician accused of malpractice loses 3 to 5 practice days to depositions, attorney's meetings, other defense preparation activities, and court appearances. With each patient served, a physician incurs a cost equal to the expected value of these uninsurable costs. This marginal per patient cost is increasing in $M$.[21] In this sense, malpractice litigation increases a physician's marginal cost.

How does a decrease in $M$ in one market affect physician location decisions? Consider the simple case where physicians are differentiated Bertrand-Nash competitors, and assume that long-run equilibrium requires the equalization of (real) net incomes across markets.[22]

[18] Based on a survey of 739 physicians, conducted in 1989. Anecdotal evidence suggests that doctors are increasingly not carrying excess-layer liability coverage, making them vulnerable to judgments beyond the limits of their standard liability insurance.

[19] Although individual experience rating is not prohibited, class rating is most prevalent in the industry. For descriptive evidence and a discussion of experience rating for medical malpractice, see Sloan (1990).

[20] See GAO (2003a) for evidence that insurance premiums are largely driven by the magnitude of losses incurred by insurance companies.

[21] Even though damage caps do not reduce the probability of facing a payment (see Table 3.1), they likely decrease a physician's marginal cost. For example, less time is required to prepare a defense when less money is at stake. Even if they do not, changes in $M$ may still affect prices. Since the physician have market power, changes in fixed costs may serve as a signal to raise prices. Alternatively, insurers may rapidly and automatically incorporate premium cost increases into reimbursement levels, insulating physician incomes from changes in $M$ (Danzon, Pauly, and Kington, 1990).

[22] Existing empirical evidence demonstrates that physician location decisions respond to financial incentives, such as reimbursement levels, income, and a town's predicted economic potential. See a survey by Fruen, Hadley, and Korper (1980) and more recent work by Wright et al. (2001).

The price change implied by a change in $M$ is

$$\frac{dP}{dM} = \frac{dc}{dM}\left(1 + \frac{1}{\varepsilon_i}\right)^{-1}$$

where $c$ is marginal cost and $\varepsilon_i$ is the physician-specific price elasticity of demand. As demand is likely relatively elastic at the physician level, prices will rise by more than the increase in marginal cost.[23] Finally, the market demand elasticity will determine whether the total demand for healthcare increases when $M$ decreases. If the market demand elasticity is low (as it is generally believed to be in most of the United States), a decrease in $M$ causes a decrease in price, a minimal increase in quantity, roughly no change in net income, and almost no change in physician supply. Whereas if market demand is elastic, a decrease in $M$ is associated with an increase in quantity, net income, and physician supply.

This theoretical structure suggests that differences in market characteristics could explain a geographically heterogeneous physician supply response to malpractice tort reform. Specifically, if the market demand elasticity or the magnitude of a physician's uninsured malpractice litigation cost vary with population density, we would expect to find a differential effect across urban and rural areas.

## 3.4 Evidence on the supply consequences of reform

### 3.4.1 Data

My research design uses physician geographic counts derived from American Medical Association (AMA) administrative records and reported in the Area Resource File (ARF), published by HHS. The data in my sample span 1970 to 2000, with slightly less coverage for specialty specific counts. The AMA, the leading national professional organization for physicians, tracks the universe of medical doctors using administrative records from U.S.medical schools, licensing exams, state licensing boards, residency programs, certification boards, and the U.S.Surgeon General. When a physician moves, his location is updated using the U.S.Postal Service Address Correction System even if the doctor does not directly inform the AMA. Cherkin and Lawrence (1977) and Williams, Whitcomb, and

[23] For example, McCarthy (1985) estimates $\varepsilon_i \approx -3$, implying an increase in price of 1.5 times the change in marginal cost.

Kessler (1996) study the data and find it to be generally accurate.

Doctor's locations come from mailing addresses associated with either the doctor's home or office. This lack of uniformity is likely of little importance at the county level.[24] The American Medical Association estimates are better than census estimates both because they are available at greater frequency (every year versus every 10 years) and because census estimates occasionally include interns and residents as well as chiropractors, dentists, and veterinarians in physician counts when enumerators fail to properly differentiate the response of "doctor" (ARF User Documentation, 2002).

Information about state laws limiting malpractice liability and other tort reforms were collected by Hellinger and Encinosa (2003) and Klick and Stratmann (2003) from summaries prepared by the National Conference of State Legislatures, the American Tort Reform Association, and the law firm McCullough, Campbell & Lane. I updated the authors' listings using similar legislative summaries as well as the original text of various statutes.

Other variables are obtained from the ARF and from the Regional Economic Information System, published by the Bureau of Economic Analysis. Table 3.2, Panel B, shows county summary statistics, separating counties into quartiles based on population density in 1970. The least densely-populated ("frontier") counties have less than fourteen residents per square mile, whereas the most densely-populated ("metropolitan") counties have more than 72 (and up to almost 55,000). Physicians are concentrated in more densely-populated counties as is population. Metropolitan counties have higher income and employment, on average, than other counties. Rural counties have a higher proportion of both farm employment and retired persons (as proxied by per capita social security income).

### 3.4.2 Econometric framework

The initial analysis uses state-year aggregates. Let $y_{it}$ be the number of doctors in state $i$ at time $t$, $pop_{it}$ represent the population, and $C_{it}$ indicate a state law limiting malpractice

[24] Of the least densely-populated quartile of counties, 63 percent are more than 40 miles from the nearest urban area (area with population greater than 50,000).

damage awards. $C_{it}$ excludes laws overturned by state courts.[25]

$$\ln\left(\frac{y_{it}}{pop_{it}}\right) = \alpha C_{it} + \eta_i + \phi_t + u_{it} \tag{3.1}$$

Specification (3.1) includes state fixed effects, $\eta_i$, and year fixed effects, $\phi_t$. Standard errors are clustered at the state level, allowing for unspecified within-state correlation over time.

The number of doctors per capita in a given market results from a number of demand and supply factors, many of which are not observable. Doctors are probably forward looking, and their location decisions driven by expectations of the demand for healthcare in future periods in addition to contemporaneous market conditions. As long as these and any other unobservables which comprise $u_{it}$ are not correlated with the legislative adoption of caps, the estimate of $\alpha$ in equation (3.1) has a causal interpretation. This assumption of uncorrelation is the principal identifying assumption of this study.

Liability reform almost certainly responds to economic and social conditions that may be correlated with regional physician supply. Some of these factors may induce omitted variables bias in estimates of the effect of damage caps. At the same time, a number of factors support a causal interpretation of these estimates. Many damage caps came out of a broader crisis in commercial casualty insurance not limited to medical lines (Priest 1987, Rabin 1988). Moreover, the legislative process that generates reforms has a substantial random component (e.g., logrolling), which is likely uncorrelated with the determinants of physician supply. Figure 3-1 illustrates the geographic diversity of states adopting damage caps reforms.

Existing empirical evidence supports the notion that malpractice damage caps can be taken as essentially exogenous in equations that control for state and year effects. Danzon (1984) shows that the concentration of physicians and their degree of organization, as measured by membership in state and local medical societies, had little impact on the adoption of tort reforms in the 1970s. In more recent work, Campbell, Kessler, and Shepherd (1995) show that while the concentrations of physicians and lawyers in a state are correlated with many liability reforms, neither is correlated with caps on malpractice damages. And, while

[25]Liability caps which face a high probability of being stricken down by state courts are unlikely to affect long-term variables such as physician location decisions. Specifications which include overturned laws in the period before they are overturned produce attenuated and less precise results. Whether the effect of caps that are not overturned intensifies over time is addressed below.

estimating physician supply, Klick and Stratmann (2003) use political variables to perform an overidentification test and cannot reject the exogeneity of legal malpractice reforms.[26] My own analysis suggests the adoption of caps is essentially unrelated to identifiable drivers of physician supply. Table 3.2, Panel C, shows that the population affected and the population unaffected by the policy changes have similar per capita income, employment, elderly populations, and farm employment shares. States that adopted caps may be, on average, slightly less densely populated. Additionally, an analysis of pre-existing trends, presented below, shows that changes in physician supply do not precede the passage of laws limiting malpractice awards.

More precise estimates of $\alpha$ may be obtained from a second specification, which employs county-level data on physician-to-population ratios and other demographic variables. Although $C_{it}$ only varies at the state level, the county-level analysis allows for the inclusion of county-level covariates which improve precision by absorbing variation in doctors per capita that is not caused by tort reform. The most important factors affecting demand for physicians relate to residents' healthcare needs and preferences, including age, income, employment status, and health insurance coverage. Although not all of the relevant variables are available, such as health insurance coverage or managed care penetration on the county-level back to 1970, county fixed effects and several controls are included in specification (3.2):

$$\ln\left(\frac{y_{ijt}}{pop_{ijt}}\right) = \alpha C_{it} + X_{ijt}\beta + \eta_j + \phi_t + \omega_i t + u_{ijt} \tag{3.2}$$

$X_{ijt}$ represents observed determinants of physicians per capita in county $j$ – specifically income, employment, and social security payments per capita (a proxy for age), and farm employment share. State-specific linear time trends, $\omega_i t$, improve the precision of the estimates by controlling for the long-run trends in the unobserved determinants of physician supply. Again, standard errors are clustered at the state level.

Another important feature of physician relocation is that it is often costly. As the opportunity cost of not utilizing the installed base of customers is sizeable and state licens-

[26]Klick and Stratmann (2003) use political variables, such as indicators of whether the state legislatures were controlled by the Democratic party and whether corporations were prohibited from making political contributions, to test the exogeneity of tort reforms. Although the authors report that the first stage had a "high" $F$ statistic, the results should be interpreted with caution as large standard errors in the 2SLS second stage may hinder the rejection of exogeneity.

ing requirements are nontrivial, long-term or persistent factors overwhelmingly influence a physician's location decision. Existing evidence suggests that for a majority of doctors, the original practice location selected upon completing residency training becomes permanent, reinforcing the importance of long-term factors in the location decision. West et al. (1996) find that, in 1991, 58.7 percent of University of Washington family practice residency graduates were still practicing in the location where they had started their careers as many as 18 years earlier, and only 16.7 had changed practice location more than once. Very few physicians moved between rural and urban settings after the first 4 years of practice, and physicians were more likely to switch from rural to urban environments (19.3 percent of those initially in rural practices) than the reverse (5.7 percent of those initially in urban practices).

Despite the significant relocation cost, I assume there is relatively rapid adjustment to equilibrium levels in the market for physicians.[27] The adjustment is achieved by the sizeable annual inflow of new physicians completing residency programs, including both U.S.medical school graduates and immigrant physicians. For example, in 1996, almost 1 in 20 physicians completed residency programs (American Medical Association 1997).[28]

Relatively rapid adjustment to equilibrium levels and the importance of long-term factors in a residency graduate's location decision implies that the unobservable time-varying determinates of physician supply are correlated over time. Econometrically, this causes the error term in my regression equation to be serially correlated.[29] Although estimates of $\widehat{\alpha}$ derived from specifications (3.1) and (3.2) are consistent, modeling the error term can help reduce the standard errors of the estimates. Suppose the unobservables, $u_{ijt}$, are first-order autoregressive with a state-specific coefficient, $\rho_i$, so the error model is

$$u_{ijt} = \rho_i u_{ijt-1} + \epsilon_{ijt} \tag{3.3}$$

This specification is estimated using a two-step procedure, where $\widehat{\rho}_i$ is estimated in the first step and $\widehat{\alpha}$ is estimated in the second after quasi-differencing. This GLS estimator

[27]Partial adjustment models were rejected by the data.

[28]In 1996, 27,739 physicians (4.4 percent of all physicians) completed residency programs. This rate is likely representative of the sample period. Since 1970, the stock of physicians in the U.S.has been growing rapidly at 3.2 percent on average per year. As this growth rate is net of retirees and others who exit the market, an even larger number share of doctors enters the market each year.

[29]The average estimated autocorrelation coefficient, $\widehat{\rho}_i$ in equation (3.3), is 0.729.

explicitly models the correlation of $u_{ijt}$ over time. To allow for possible cross-sectional correlation in $\epsilon_{ijt}$ (across counties within a state at a point in time), standard errors in the second step are clustered at the state-year level.[30,31]

As an alternative to the fixed effects specifications, I also estimate a regression in first differences.

$$\ln\left(\frac{y_{ijt}}{pop_{ijt}}\right) - \ln\left(\frac{y_{ijt-1}}{pop_{ijt-1}}\right) = \alpha\left(C_{it} - C_{it-1}\right) + \left(X_{ijt} - X_{ijt-1}\right)\beta + \left(\phi_t - \phi_{t-1}\right) + \omega_i + \nu_{ijt} \tag{3.4}$$

Since the estimated autocorrelation coefficient in specification (3.3) is large (on average, $\widehat{\rho}_i = 0.729$), the first difference specification should yield similar results and may be easier to interpret. If the FGLS model for the error term in specification (3.3) is not correct, then the first difference estimator in specification (3.4) may be more precise.

### 3.4.3 Results

Difference-in-difference comparisons provide a simple measure of the effect of a cap. Table 3.3 presents population-weighted average physician-to-population ratios in 1970 and 2000. As depicted in Figure 3-1, state laws limiting damages in malpractice liability cases were adopted throughout this period, beginning in 1975. The difference-in-difference comparison, reported in the final column, measures the increase in physician supply in adopting states minus the increase experienced in other states. The analysis is also performed on subpopulations of physicians in two dimensions—physician specialty and population density. The analysis is performed for active physicians overall as well as for physicians engaged in office-based patient care in the following specialties: general or family practice; medical specialties, including allergy, cardiovascular diseases, dermatology, gastroenterology, internal medicine, and pediatrics; surgical specialties, including obstetrics and gynecology, ophthalmology, and urology; and support specialties, including anesthesiology, neurology, pathology, psychiatry, and radiology. The analysis is performed for entire states as well as by population density quartile. The difference-in-difference comparisons in Panel I suggest

[30] There may be some concern that the error process is misspecified. For example, the true process may be AR(2). Clustering standard errors at the state (rather than state-year) level is robust to such misspecification and produces virtually identical results.

[31] Ideally, I would be able to implement a GLS procedure such as Hansen (2004) to parametrically model the cross-sectional correlation in $u_{ijt}$. However, with as few as 3 and on average only 63 counties per state, the large sample assumptions required are not justifiable in my application.

that although damage caps do not have a sizeable effect on physician-to-population ratios for the average individual in affected states, they do increase physician supply in the most rural areas. In these areas, caps on liability increased the growth in physician supply by 33 percent over the three decades.[32] Panel II shows that broadly similar difference-in-difference comparisons hold for each specialty group.

Results from population-weighted regression analysis are reported in Table 3.4, where each coefficient is from a separate regression. Column (1) contains state-level OLS analysis with no covariates—specification (3.1). Column (2) includes state-specific linear trends, and Column (3) adds county-level covariates and fixed effects—specification (3.2). Column (4) presents GLS results assuming a first-order, state-specific autoregressive process for the error term—specification (3.3). And, Column (5) reports regressions of first differences—specification (3.4). The analysis reported in Panel A includes each entire state, whereas Panel B presents separate estimates by population density. For example, the first row of Panel B includes only counties with population density less than 14 residents per square mile in 1970. In this way, the most rural areas in states that do not adopt caps provide a counterfactual for similarly densely populated areas in states that do.

Two important facts are evident in Table 3.4. First, malpractice liability caps do not increase physician supply for the average resident of adopting states. Panel A presents results from population-weighted regressions on statewide data. From all of the specifications, the greatest point estimate is 1.4 percent (not significant) and none of the estimates are positive and statistically significant. Second, caps do, however, increase physician supply in the most rural areas. The first row of Panel B presents estimates for the least densely populated quartile of counties. In these areas, damage caps increase physician supply for the average resident by 3 to 5 percent.[33] Estimates in the remaining rows suggest that physician supply does not increase in other parts of states that adopt the caps.[34]

[32] Caps are associated with an increase in total physician supply of 14 doctors per 100,000 residents in the most rural quartile of counties, and physician supply increased by 42 (97 minus 55) doctors per 100,000 residents in similar counties that were unaffected by the policy change.

[33] In principle, splitting the sample based on population density while estimating an equation that includes population in the dependent variable may bias the regression estimates. However, estimating the equation with log population on the right-hand side with an unrestricted coefficient is robust to such concerns and produces similar results.

[34] Amy Finkelstein thoughtfully suggested an alternative hypothesis. Suppose that a principal margin on which physician supply adjusts is shifting retirement and that doctors in rural areas tend to be older. Then, a damage cap may affect all physicians near retirement, but the measured effect would be strongest in rural areas. Rudimentary empirical analysis does not support this hypothesis. In 2000 (data for the

The effects of damage caps on physician supply may intensify with time after the reform is adopted. Such a phenomenon could explain observed differences between estimates from the various specifications. If the policy has dynamic effects, OLS estimates in Columns (1)-(3) measure the average effect, first difference estimates in Column (5) represent the effect in the year of adoption, and the GLS estimates in Column (4) present a weighted average favoring initial effects. The effect of laws limiting damage awards may increase over time for three primary reasons. First, as time passes and the law survives legal challenges, there is less uncertainty about whether the cap will be overturned in court.[35] Second, it may take time for the expectation of both liability insurers and physicians to adjust to a new equilibrium.[36] Third, many liability caps have not been adjusted regularly with inflation, and thereby, have become effectively more binding over time.[37]

This dynamic effects hypothesis is examined in Table 3.5. In these regression specifications, the liability cap variable is interacted with variables indicating the number of years since the law was passed. Each column represents a separate regression on different samples of counties. The first column includes the full sample and each subsequent column includes only counties in each population density quartile. The regressions reported in Table 3.5 include county-level covariates and fixed effects, and the results are robust to using other specifications from Table 3.4. Although a constant policy effect cannot be rejected at conventional confidence intervals, point estimates suggest that the effects may be weaker in the initial period following adoption. This pattern is most consistent with both legal uncertainty and slow adjustment to the new equilibrium.

The analysis in Table 3.5 also provides a Granger-type test for pre-existing trends. An indicator variable for the 1 to 5 years before adoption of the cap is also included. The

early 1970s is not readily available), the age distribution of physicians was similar across rural and urban areas. Furthermore, cross-sectional regressions for the year 2000 suggest that, whereas physicians in rural areas of state with damage caps retire later than rural physicians in other states, the evidence for physicians in more urban areas is inconclusive. Although the cross-sectional analysis is not dispositive, malpractice liability likely influences a physician's retirement decision and professional activity. As liability premiums are generally not volume-related, an increase in premiums makes part-time medical practice an economically less viable option.

[35] As caps have not survived judicial review in over a fifth of adopting states, the probability of being overturned is not trivial.

[36] When it is difficult to predict losses on claims, insurers will generally adopt relatively conservative expectations of losses (GAO 2004a). As the accuracy of doctors' perceptions of uninsurable risk is also likely to improve with time, tort reform liability premiums and physician supply may respond slowly to tort reform.

[37] For example, the California damage cap was fixed at $250,000 in 1975 and has not been adjusted for inflation. By 2003, the cap corresponded to a $70,000 limit in 1975 dollars.

coefficient corresponding to this variable is not statistically significantly different from zero at the 5 percent level. Increases in physician supply do not seem to precede the laws, and the laws do not seem to be generated during periods of relatively low physician supply.

Limits on malpractice damage awards are sometimes adopted in conjunction with other tort reforms. Even when not adopted simultaneously, the timing of various reforms is often correlated. In principle, disentangling the effects of the various tort reforms may be difficult with only a limited number of policy experiments. To confirm that damage caps in particular affect physician supply in the most rural areas, I add a control for the adoption of other tort reforms. The results reported above are indeed robust to such a specification.

A control for the presence of other tort reforms is included in Table 3.6. Each column represents a separate regression on different samples of counties. In addition to the damage cap variable, each regression includes an indicator variable for whether any of the following five other popular tort reforms was in effect: collateral-source rule, joint-and-several liability, attorney fee restrictions, pretrial screening, and periodic payments.[38,39] These other reforms do not seem to affect physician supply for the average resident of adopting states. In contrast, even after controlling for the adoption pattern of other reforms, medical malpractice damage caps are associated with a 5 percent increase in physician supply in the most rural areas. The impact of damage caps on physician supply appears to be restricted to these areas, and none of the other reforms seem to have a comparable effect.

Dividing physician supply into major specialty groups enables me to further isolate the effects of the caps. Table 3.7 reports regressions similar to those in Table 3.6, but where the dependent variables are specialty-specific physician-to-population ratios. Only the damage cap coefficients are reported. Each estimate is from a separate regression; each row corresponds to a different dependent variable, and each column corresponds to a different sample of counties. Similar to physicians overall, damage caps do not increase the supply of any particular type of physician for the average state resident. The effect of damage caps on physician supply in the most rural areas is limited to specialist physicians.

[38] A collateral-source rule allows damages to be reduced by the value of compensatory payments already made to the plaintiff. Joint-and-several liability limits damages recoverable from parties only partially responsible for the plaintiff's harm. Attorney fee restrictions subject contingency fee payments (to the plaintiff's attorney) to statutory caps or court approval. Pretrial screening reform is included when claims must be submitted to a hearing panel, whether or not the results are admissible at trial. Periodic payments reform requires part or all of damages be paid in the form of an annuity.

[39] The damage cap estimates are also robust to including a separate indicator variable for each reform.

The effect is driven by a 10 percent increase in surgical specialists and a 12 percent increase in support specialists.

There are several reasons why laws limiting malpractice damage awards would not affect generalist physicians, such as internists, pediatricians, and those in general or family practice. First, the direct financial effect of damage caps may differ for generalists and specialists (see Section 3.5).[40] Second, longer-lived and perhaps closer patient relationships among generalists may reduce their sensitivity to the malpractice environment.[41] Third, over the past several decades, primary care has been increasingly delivered by internists and pediatricians, driving nationwide decreases in the relative number of general and family practice physicians (see Table 3.3). The extent to which damage caps have contributed to this shift cannot be discerned from Table 3.7 or the data available to this study.

The asymmetric physician supply response across rural and metropolitan areas is more of a mystery and, to my knowledge, has not been systematically explored by researchers. I turn to considering potential explanations of the urban-rural effects of malpractice liability in the next section.

## 3.5 Explanations of the urban-rural heterogeneity in supply response

*It is not that they can't practice in those counties; it is that they cannot afford to practice in those counties. Why can't they? This problem is a uniquely rural problem in some ways.*

Senator Judd Gregg (R-NH)
February 23, 2004

One of the few previous attempts to focus on urban-rural differences in the effects of malpractice liability is a study by Ventä et al. (1998), who examine Finnish malpractice claims against dentists and oral surgeons for permanent nerve injuries related to third molar extractions. They find that claims originated more often from rural areas (3.8 claims per 100,000 residents) than urban areas (2.4), and claims were brought against 2 percent of total dental surgeons (generalists) and 26 percent of all oral and maxillofacial surgeons

[40] Although they are difficult to quantify, nonfinancial effects may differ as well. For example, the reputation costs of facing a malpractice claim may not be symmetric across fields.

[41] May and Stengel (1990) finds that patients are less likely to sue if their doctors show concern for them personally.

(specialists).[42] Fondren and Ricketts (1993) look at the limited availability of high-risk obstetric care in rural North Carolina in 1989. They find that rural obstetricians were twice as likely to stop or reduce providing care to medically high-risk patients as metropolitan obstetricians (25 versus 13 percent), and almost three times as many rural (46 percent) as urban (16.7 percent) physicians whose obstetric patient volume had decreased reported fear of malpractice lawsuit as an important factor. Rural physicians were also more sensitivity to liability premiums. More rural (65 percent) than urban physicians (54 percent) indicated that a premium of $50,000 would be too high to continue practicing obstetrics.[43]

What drives these observed heterogeneous effects of malpractice liability across urban and rural areas? Theoretical consideration in Section 3.3 points to both differences in how caps affect a physician's costs as well as differences in the market elasticity of demand for health care.

### 3.5.1 Malpractice litigation costs

Malpractice liability may more severely affect the cost structure of rural physicians. Returning to the model presented in Section 3.3, geographic differences in the physician cost structure may impact the supply response. If price changes following a change in $M$ do not fully offset changes in costs, physician supply will respond. The magnitude of the response is increasing in the percentage change in costs. Empirical evidence suggests that malpractice litigation costs compose a disproportionate share of both fixed and marginal costs in rural areas.

Rural medical practice differs from urban practice along several dimensions. Table 3.8 presents some statistics on the economics of medical practice, broken down by specialty and practice location. The data are derived from physician surveys conducted as a part of the AMA Socioeconomic Monitoring System. Fees charged for office visits in rural areas are substantially below those in metropolitan areas (Panel B), and rural doctors work more hours per week and perform more patient visits (Panel A). In net, consistent with the theoretical assumptions in Section 3.3, they earn similar net incomes to their metropolitan

[42] The authors recognize that the concentration of claims among specialists may reflect the relative difficulty of cases referred to them. They attribute the relatively large number of claims in rural areas to the scarcity of specialists in those areas.

[43] Thirteen percent of rural physicians and 20 percent of urban physicians paid obstetrics premiums in the $40,000-$50,000 range.

counterparts (Panel E).

As a share of fixed costs, liability premiums are greatest for rural specialists. Average selected fixed costs by geography and specialty are listed in Table 3.8, Panel C.[44] Even though premiums are generally not experience-rated on the individual level, physician surveys, such as the one which generated this data, are likely more representative of actual premiums paid than insurer reported rack rates.[45] The liability premium as a share of fixed costs ("relative premium") is highest for rural physicians with a surgical specialty. Among surgical specialists, relative premiums are greater in rural areas than in metropolitan areas, but the difference is only marginally statistically significant.[46] Physicians in general or family practice have lower relative premiums than specialists, and among primary care physicians, rural doctors do not have the lowest relative premiums. These facts are broadly consistent with the effect of damage caps measured in Section 3.4.

Nevertheless, changes in liability premiums likely explain relatively little of the effect of caps on rural physician supply. Although the changes in premiums may be sizeable, they compose a relatively small share of a typical physician's total costs (approximately 12 percent, excluding the shadow cost of the physician's time and uninsured malpractice litigation cost). Evaluated at the means reported in Table 3.8, a ten percent decrease in malpractice premiums is economically equivalent to 1.1 (1 percent) more visits with established patients per week or a $0.37 (1 percent) increase in the fee charged for such a visit. Although the effect of caps on liability premiums may be economically significant for physicians in the tail of the distribution of premiums, it is unlikely that premiums are driving the bulk of the observed rural physician supply response.

Geographic differences in the physician supply response to malpractice liability are more plausibly explained by differences in uninsured malpractice litigation costs. NPDB data

[44]To avoid issues of cost allocation, data is for self-employed physicians in solo medical practice only. Office expense includes rent, mortgage interest, depreciation on medical buildings used in the physician's practice, utilities, and telephone. Medical equipment expense is imprecisely measured and excluded from the analysis.

[45]Industry surveys suggest rack rates are uniform within many states (e.g., Medical Liability Monitor 2002).

[46]The standard error of the relative premium is calculated using the delta method and assuming that liability premiums and office expenses are uncorrelated. If the costs are positively correlated, then the true standard errors are even lower. A one-tailed t-test rejects the hypothesis that relative premiums are lower in rural areas than in metropolitan areas with less than one million residents ($p < 0.10$); however the difference between rural areas and metropolitan areas with more than one million residents is not statistically significant at standard confidence levels.

demonstrate that malpractice pressure is particularly intense on rural doctors, suggesting uninsured malpractice litigation costs are significantly greater in rural areas. The Division of Practitioner Data Banks at HHS has provided me with annual summary data for groupings of counties, classified by their population density.[47] I regress each of several dependent variables, relating to the frequency and severity of malpractice awards on variables indicating population density quartiles, these indicators interacted with a damage cap dummy, and state and year fixed effects. Metropolitan areas form the omitted category, and states without such areas (Montana, North Dakota, and Wyoming) are excluded from the analysis. The results are reported in Table 3.9, where each column corresponds to a regression with a different dependent variable. Per physician, there are the most malpractice payments in the most rural areas in terms of both dollars (23 percent greater than metropolitan areas, Column 1) and numbers (67 percent, Column 2). Although severity is lower on average in rural areas (Column 7), no appreciable difference in medians (Column 8) suggests this may be attributable to a few extremely large metropolitan area awards. In fact, the per doctor number of claim payments of at least $250,000 is also greatest in the most rural areas—52 percent greater than in metropolitan areas (Column 3).

The skewed distribution of physicians drives a wedge between the per doctor and per capita measures of malpractice payments. In fact, Columns (4)-(6) show that per capita measures of malpractice payments are lower in frontier rural areas than in metropolitan areas by 54 to 65 percent. At least two factors may contribute to these differences. First, many residents of frontier rural communities likely travel to metropolitan areas for high-risk medical care. Based on hospital records in New York State, Weiler et al. (1993) find that in 1984 rural hospitals had lower iatrogenic injury rates than New York City hospitals, but similar rates of negligence conditional upon such an injury. Second, rural patients may be less likely to sue their physicians for a variety of reasons.[48] For example, rural physicians

[47] The Division of Practitioner Data Banks provided aggregated data in order to protect identifying information, which is strictly limited to hospitals, HMOs, the individual physician, and in certain circumstances a plaintiff. Physicians are matched with counties based on their work zip code. If work zip code is unavailable, home zip code is used. Some records contain invalid zip codes and could not be matched. Frequent changes to zip codes by the postal service also interfered with matching some records. In all, 98 percent of records were successfully mapped to counties.

[48] Considering that rural residents have lower incomes and less frequent employment (see Table 3.2), one might expect them to be more likely to sue. Donohue and Siegelman (1993) show that downturns in the business cycle increase the propensity of plaintiffs to file federal employment discrimination cases. However, Burstin et al. (1993) actually show that the poor, elderly, and uninsured (all of whom are overrepresented in rural communities) are less likely to sue for medical malpractice, even after controlling for the presence

may maintain closer doctor-patient relationships, and survey evidence presented by May and Stengel (1990) finds that patients are less likely to sue if their doctors show concern for them personally. Even though there may be a lower chance that any particular rural patient will sue his doctor, rural physicians face greater uninsured malpractice costs after aggregating risk over their larger caseloads. Being less numerous per capita, they work more hours and conduct more patient visits than their metropolitan colleagues (see Table 3.8, Panel A). Although each rural patient visit may bear lower expected malpractice costs, the relative intensity of a rural physician's workload leads rural physicians to face greater uninsured litigation costs over the course of a year.[49]

Are uninsured litigation costs in rural areas disproportionately affected by damage caps? Unfortunately, the timing of legal changes and the collection of NPDB data do not allow a proper analysis of this question. Due to extreme heterogeneity in the mix of population density across states, a cross-sectional analysis would compare different states for each population density quartile.[50] I adopt an alternate approach, exploiting within state rather than between state comparisons. I compare the rural-metropolitan contrast across states. If the ratio of rural to metropolitan malpractice payments in states that do not adopt damage caps form the appropriate counterfactual for the similar ratio in adopting states, then adding interaction variables to the regressions estimated in Table 3.9 will represent the effect of the law. None of these interaction terms are statistically different than zero at the five percent level for any of the dependent variables. The estimates are simply not precise enough to determine the relative effect of the law.[51] The effect of damage caps on uninsured litigation costs may also vary by specialty. Unfortunately, NPDB data cannot be used to test this hypothesis, because information on physician specialty is not collected.[52]

and degree of iatrogenic injury.

[49] These results complement Danzon (1984) who concludes that "Urbanization is the single most powerful predictor of both frequency and severity [of malpractice claims]." She uses two surveys of claims closed by insurance companies in 1970 and 1975-78. Her regressions control for the physician-to-population ratio rather than focus on per physician measures of the malpractice environment. In this sense, Danzon's results correspond more to Columns (4)-(8) than to Columns (1)-(3) in Table 3.9.

[50] For example, Wyoming does not contain a metropolitan county, and Alabama does not contain a frontier rural county. This lack of correspondence severely complicates the interpretation of cross-sectional analyses disaggregated below the state level.

[51] For example, a 95 percent confidence interval for the effect of a damage cap in the least densely populated areas (relative to the most densely populated areas) on malpractice payments per capita includes both a 16 percent decrease and an 84 percent increase.

[52] A limited description of the malpractice act or omission (e.g., diagnosis, anesthesia, treatment) is recorded in the NPDB. With respect to these classifications, malpractice payments follow a similar pattern across population density quartiles.

Even if damage caps have a constant effect on uninsured malpractice costs across geographic areas, they may have a greater effect on physician marginal costs in rural areas if other (non-litigation related) variable costs are lower in rural areas. Table 3.8, Panel D, suggests this may be the case. Total physician expenditure on medical materials and supplies, such as drugs, x-ray films, and disposable medical products, is lower in rural than urban areas, although the difference is not statistically significant. Factoring in the greater caseload among rural physicians, this difference is potentially sizeable. In sum, empirical evidence suggests that malpractice liability may have a more severe effect on the cost structure of rural versus urban physicians.

### 3.5.2 Market elasticity of demand

Even if physician costs can be fully passed through to consumers, unless the market demand for health care is sufficiently inelastic, malpractice liability will affect health care utilization and physician net income and supply. Differences in health insurance coverage rates across urban and rural areas may drive differences in the market elasticity of demand. Full health insurance coverage effectively reduces a patient's demand elasticity to near zero. The severe effect of health insurance coverage on the demand for health care is well known.[53] For example, based on a randomized experiment, Manning et al. (1987) find that a catastrophic insurance plan reduces expenditures 31 percent relative to a zero out-of-pocket price. Rates of health insurance coverage are lowest in rural areas, where full-time employment is less prevalent and agricultural work is more prevalent.[54] Ormond, Zuckerman, and Lhila (2000) find that 21.9 percent of residents of rural counties removed from urban areas are uninsured, compared to 17.5 percent for rural adjacent counties and 14.3 percent in urban counties. Further, in a study of Nebraska households, Mueller, Patil, and Ullrich (1997) find that the median spell of uninsurance is approximately 22 months in the most rural areas (less than 6 persons per square mile), 16 months in other rural communities, and 6 months in urban areas (MSAs).

Greater rates of Medicaid enrollment may also increase the market demand elasticity in the most rural areas. Danzon, Pauly, and Kington (1990) find that physicians are unable

[53] Studies from the medical literature include Newacheck (1989), Freeman et al. (1990), Cornelius, Beauregard, and Cohen (1991), and Hafner-Eaton (1993).

[54] As county-level data on health insurance rates are not publicly available, I rely on evidence presented by other authors. For other county descriptive statistics, see Table 3.2.

to pass-through malpractice costs to Medicaid reimbursement as they do with Medicare and private insurance.[55] Without the ability to adjust price, decreases in $M$ may increase physician net income and supply. Therefore, greater rates of Medicaid coverage would be associated with greater affects of $M$ on physician supply.[56] Estimates from the Current Population Survey suggest that Medicaid enrollment rates are highest in rural areas.[57] In 2000, 12 percent of non-metropolitan residents were on Medicaid, compared to 10 percent of those in metropolitan areas ($p < 0.001$). Enrollment rates are likely significantly even greater in the most rural areas (Ormond et al. 2000). Fondren and Ricketts (1993) find that among obstetricians in North Carolina, twice as many urban as rural physicians have minimal to no Medicaid caseload (64 versus 33 percent). Thus, not only do rural physicians face greater uninsured malpractice litigation costs, but they are likely less able to pass these costs onto their patients. These factors likely drive the observed geographic heterogeneity in the effect of damage caps on physician supply.

## 3.6 Welfare implications

Although increases in the physician-to-population ratio are generally considered to be desirable outcomes, evidence of a resulting increase in physician supply is not sufficient to imply that imposing a damage cap is welfare enhancing or even beneficial to public health. Limits on malpractice damage awards interfere with the two key functions of the tort liability system: deterrence and compensation.

A damage cap may undermine the deterrence incentive provided by medical malpractice liability. If quality of care and medical errors are elastic with respect to the degree of liability, patients may be harmed by the introduction of a liability cap. Weiler et al (1993) provide a study of deterrence of medical negligence based on 49 hospitals in New York State. Although they are not statistically significant, the point estimates imply that, on the

[55] Medicare practice costs adjustments ostensibly account for malpractice premium differences and changes over time, but they are unlikely to reflect uninsured malpractice litigation costs. The responsiveness of reimbursement from private insurance presumably reflects the degree of insurer market power.

[56] As compared to 1978-83, Danzon, Pauly, and Kington (1990) report finding less evidence for the relationship between malpractice variables and fees during 1976-78, when malpractice premiums generally fell. While fees may respond asymmetrically to increases and decreases in $M$, greater price stickiness in areas with more Medicaid patients will lead to greater increases in physician net income when $M$ decreases.

[57] The Centers for Medicare & Medicaid Services at HHS do not collect county-level Medicaid enrollment statistics.

margin, tort liability reduces the rate of negligent injuries per admission by 29 percent and the overall rate of medical injuries by 11 percent.[58] Klick and Stratmann (2003) provide evidence that some malpractice tort reforms, including noneconomic damage caps, may increase the 6-day infant mortality rate.

Damage caps also limit the compensation paid to victims of iatrogenic injuries. Although caps may not affect the number of patients compensated, they have been shown to reduce the size of damage awards by 23 to 48 percent on average (Table 3.1; Danzon 1986; Zuckerman et al. 1990; Sloan et al. 1989) and to most severely affect those who suffered grave injury, pain, and disfigurement (Studdert et al. 2004; Pace et al. 2004). As tort compensation is in effect compulsory insurance tied to the purchase of medical care, reducing damage awards may be suboptimal if adverse selection or another market failure prevents patients from purchasing supplemental insurance.

However, reforms may also increase the efficiency of the tort system. Evidence suggests the system may provide deleterious incentives while failing to take action following the majority of negligent iatrogenic injuries. Kessler and McClellan (1996) present evidence that the fear of malpractice drives some physicians to practice defensive medicine, administering precautionary treatments with minimal expected benefit out of fear of legal liability. And, the Harvard Medical Practice Study (1990) finds that only 1 in 17 victims of negligent iatrogenic injury in New York State in 1984 received compensation. Reforms that curtail such inefficiencies may be optimal.

Danzon (2000) advocates a schedule of damage award limits that vary with the severity of the injury. Danzon argues that caps on noneconomic damages are unlikely to undermine deterrence, because very large awards are typically not used for rating individual (as opposed to class) liability premiums. She implicitly assumes that caps do not affect the magnitude a physician's uninsured liability costs. Recognizing the insurance component of tort compensation, Danzon posits that optimal compensation is the amount that patients would choose to purchase voluntarily. She argues that the lack of voluntary insurance coverage for noneconomic loss in other private or social insurance programs suggests that such coverage is not optimal.

In net, assessment of the welfare implications of the adopted caps relies on the optimality

[58] Interestingly in the context of this study, the authors instrument for claims per negligent event using urbanization and population density.

of the original tort system and whether the caps address any inevitable inefficiencies. Without further evidence, the welfare implications of limits on liability awards are impossible to determine.

## 3.7 Conclusion

Over the last four decades, rural areas in the United States have experienced a steady relative increase in the number of specialist physicians. Williams et al. (1983) find that the distance that residents of outlying areas (towns of less than 25,000, outside metropolitan areas) must travel to reach medical and surgical specialists was substantially reduced in the 1970s. Schwartz et al. (1980) conclude that this trend is driven by an overall increase in the supply of specialists, potentially aided by a shift in preferences toward small-town living. This chapter adds to the understanding of one of the drivers of this change. Back-of-the-envelope calculation using estimates presented in this chapter implies that the enactments of damage caps are responsible for approximately 17 percent of the increase in frontier rural specialists in these states since 1970.

The estimates in this chapter also suggest that, should the United States adopt a national cap, the typical resident of a frontier rural area of a state that has not otherwise adopted such a law would see an increase of almost 7 specialist physicians in his county – an increase of 13 percent. These gains come in places that arguably have the greatest need; both overall- and specialist-physician-to-population ratios in the least densely populated quartile of counties are less than forty percent of the national average. Although the effects of damage caps are sizeable in these communities, the associated potential social costs should be compared with those of a more targeted policy.

The particularly acute effects of malpractice liability in rural areas have been largely unexplored in the economics or policy literature. Analysis of NPDB data finds that the frequency of malpractice claims per doctor is greatest in the most rural areas, and liability costs seem to constitute a larger share of both fixed and marginal costs for rural physicians. Theoretical and empirical evidence also suggests that elastic market demand for health care in the most rural areas of the United States exacerbates the effects of medical malpractice liability by preventing the complete pass-through of malpractice litigation costs to patient fees.

While malpractice liability has particularly strong effects on physician supply in the most rural areas, laws limiting damage awards for medical malpractice have no significant effect on physician supply for most Americans. Key fundamentals underlying the lion's share of the U.S. healthcare market explain why adopting reforms such as liability damage caps does not produce all of their intended effects. As the market demand for health care is generally considered to be highly inelastic, physician fees likely adjust to fully reflect changes in malpractice litigation costs. As a result, most physicians' net incomes are effectively insulated from the malpractice environment, and tort reform damage caps have no significant effect on most physicians' location decisions.

## References


American Medical Association (1997). Graduate medical education. *Journal of the American Medical Association* 278 (9), 775-84.

ARF User Documentation (2002). National Center for Health Workforce Analysis, Bureau of Health Professions, Health Resources and Services Administration, Department of Health and Human Services.

Baicker, K. and A. Chandra. (2004). The effect of malpractice liability on the delivery of health care. NBER working paper 10709.

Burstin, H. R., W. G. Johnson, S. R. Lipsitz, and T. A. Brennan (1993). Do the poor sue more? A case-control study of malpractice claims and socioeconomic status. *Journal of the American Medical Association* 270 (14), 1697-701.

Campbell, T. J., D. P. Kessler, and G. B. Shepherd (1995). The causes and effects of liability reform: Some empirical evidence, NBER working paper 4989.

Cherkin, D. and D. Lawrence (1977). An evaluation of the American Medical Association's physician masterfile as a data source – One state's experience. *Medical Care* 15 (9), 767-9.

Cornelius, L., K. Beauregard, and J. Cohen (1991). Usual sources of medical care and their characteristics (AHCPR Pub. No. 91-0042). *National Medical Expenditure Survey Research Findings* 11, Rockville, MD: Agency for Health Care Policy and Research.

Danzon, P. M. (1984). The frequency and severity of medical malpractice claims. *Journal of Law and Economics* 27 (1), 115-48.

Danzon, P. M. (1986). The frequency and severity of medical malpractice claims: New evidence. *Law and Contemporary Problems* 49 (2), 57-84.

Danzon, P. M. (2000). Liability for medical malpractice. *Handbook of Health Economics* 1B, 1339-404.

Danzon, P. M., M. V. Pauly, and R. S. Kington (1990). The effects of malpractice litigation on physicians' fees and incomes. *American Economic Review* 80 (2), 122-27.

Donohue, J. J., and P. Siegelman (1993). Law and macroeconomics: Employment discrimination litigation over the business cycle. *Southern California Law Review* 66: 709-765.

Fondren, L. K. and T. C. Ricketts, (1993). The North Carolina obstetrics access and professional liability study: A rural-urban analysis. *The Journal of Rural Health* 9 (2), 129-37.

Freeman, H. E., L. H. Aiken, R. J. Blendon, and C. R. Corey (1990). Uninsured working-age adults: Characteristics and consequences. *Health Services Research* 24 (6), 811-23.

Fruen, M. A., Hadley, J., and Korper, S. P. (1980). Effects of financial incentives on physicians' specialty and location decisions. *Health Policy and Education* 1 (2), 143-59.

GAO (2003a, June 27). *Medical malpractice insurance: Multiple factors have contributed to increased premium rates,* GAO-03-702. Washington, DC.

GAO (2003b, August 8). *Implications of rising premiums on access to health care,* GAO-03-836. Washington, DC.

Hafner-Eaton, C. (1993). Physician utilization disparities between the uninsured and insured: Comparisons of the chronically ill, acutely ill, and well nonelderly populations. *Journal of the American Medical Association* 269 (6), 787-92.

Hallinan, J. T. (2004, August 27). Doctor is out: Attempt to track malpractice cases is often thwarted. *The Wall Street Journal,* A1.

Hansen, C. B. (2004). Generalized least squares estimation in differences-in-difference and other panel models, MIT mimeo.

Harvard Medical Practice Study (1990). *Patients, doctors, and lawyers: Medical injury, malpractice litigation, and patient compensation in New York*, the report of the Harvard Medical Practice Study to the state of New York, Cambridge, MA: President and Fellows of Harvard College.

Hellinger, F. J. and W. E. Encinosa (2003). The impact of state laws limiting malpractice awards on the geographic distribution of physicians, U. S. Department of Health and Human Services, Agency for Healthcare Research and Quality mimeo.

Kessler, D. and M. McClellan (1996). Do doctors practice defensive medicine? *Quarterly Journal of Economics* 111 (2), 353-90.

Klick, J. and T. Stratmann (2003). Does medical malpractice reform help states retain physicians and does it matter? Florida State University College of Law mimeo.

Lawthers, A. G., A. R. Localio, N. M. Laird, S. Lipsitz, L. Hebert, and T. A. Brennan (1992). Physicians' perceptions of the risk of being sued. *Journal of Health Politics, Policy and Law* 17 (3), 463-82.

Malcolm, A. H. (1985, February 12). Fear of malpractice suits spurring some doctors to leave obstetrics. *The New York Times*, A1.

Manning, W. G., J. P. Newhouse, N. Duan, E. B. Keeler, A. Leibowitz, and M. S. Marquis (1987). Health insurance and the demand for medical care: Evidence from a randomized experiment. *American Economic Review* V77, N3: 251-77.

May, M. L. and D. B. Stengel (1990). Who sues their doctors? How patients handle medical grievances. *Law and Society Review* 24: 105-120.

Mccarthy, T. R. (1985). The competitive nature of the primary-care physician services market. *Journal of Health Economics* 4 (2), 93-117.

Medical Liability Monitor (2002, October). Trends in 2002 rates for physicians' medical professional liability insurance.

Mueller, K. J., K. Patil, and F. Ullrich (1997). Lengthening Spells of uninsurance and their consequences, *Journal of Rural Health* 13 (1), 29-37.

Newacheck, P. W. (1989). Improving access to health services for adolescents from economically disadvantaged families. *Pediatrics* 84 (6), 1056-63.

Ormond, B., S. Zuckerman, and A. Lhila (2000). Rural/urban differences in health care are not uniform across states. *Assessing The New Federalism*, Series B, No. B-11, Washington, DC: The Urban Institute.

Pace, N. M., D. Golinelli, and L. Zakaras (2004). *Capping non-economic awards in medical malpractice trials: California jury verdicts under MICRA,* MG-234-ICJ. Santa Monica, CA: Rand Corporation.

Priest, G. L. (1987). The current insurance crisis and modern tort law, *Yale Law Journal* 96: 1521-89.

Rabin, R. L. (1988). Some reflections on the process of tort reform, *San Diego Law Review* 25: 13-48.

Schwartz, W. B., J. P. Newhouse, B. W. Bennett, and A. P. Williams (1980). The changing geographic distribution of board-certified physicians, *The New England Journal of Medicine* 303 (18), 1032-8.

Sloan, F. A., P. M. Mergenhagen, and, R. R. Bovbjerg (1989). Effects of tort reforms on the value of closed medical malpractice claims: A microanalysis. *Journal of Health Politics, Policy and Law* 14 (4), 663-89.

Sloan, F. A. (1990). Experience rating: Does it make sense for medical malpractice insurance. *American Economic Review* 80 (2), 128-33.

Studdert, D. M., Y. T. Yang, and M. M. Mello (2004). Are damages caps regressive? A study of malpractice jury verdicts in california. *Health Affairs* 23 (4), 54-67.

Thorpe, K. E. (2004). The medical malpractice "crisis": Recent trends and the impact of state tort reforms. *Health Affairs* Web Exclusive, January 21.

U. S. Department of Health and Human Services (2003). *Addressing the new health care crisis: Reforming the medical litigation system to improve the quality of health care.* Washington, DC.

Ventä, I., C. Lindqvist, and P. Ylipaavalniemi (1998). Malpractice claims for permanent nerve injuries related to third molar removals. *Acta Odontologica Scandinavica* 56 (4), 193-6.

Viscusi, W. K. (1996). Pain and suffering: Damages in search of a sounder rationale. *Michigan Law and Policy Review* 1: 141-178.

Viscusi, W. K. and P. Born (2004). Damages caps, insurability, and the performance of medical malpractice insurance. Harvard John M. Olin Discussion Paper No. 467.

Weiler, P. C. et al. (1993). *A measure of malpractice: Medical injury, malpractice litigation, and patient compensation.* Cambridge, MA: Harvard University Press.

Weiss, B. D. (1986). The effect of malpractice insurance costs on family physicians' hospital practices. *The Journal of Family Practice* 23 (1), 55-8

West, P. A., T. E. Norris, E. J. Gore, L. M. Baldwin, and L. G. Hart (1996). The geographic and temporal patterns of residency-trained family physicians: University of Washington family practice residency network. *The Journal of the American Board of Family Practice* 9 (2), 100-8.

Williams, A. P., W. B. Schwartz, J. P. Newhouse, and B. W. Bennett (1983). How many miles to the doctor? *The New England Journal of Medicine* 309 (16), 958-63.

Williams, P. T., M. Whitcomb, and J. Kessler (1996). Quality of the family physician component of AMA masterfile. *The Journal of the American Board of Family Practice* 9 (2), 94-9.

Wright, G. E., C. H.A. Anrilla, and L. G. Hart (2001). How many physicians can a rural community support? A practice income potential model for Washington state. WWAMI Rural Health Research Center Working Paper No. 45.

Zuckerman, S., R. R. Bovbjerg, and F. Sloan (1990). Effects of tort reforms and other factors on medical malpractice insurance premiums. *Inquiry* 27 (2), 167-82.

Figure I: State adoption of tort reform damage caps

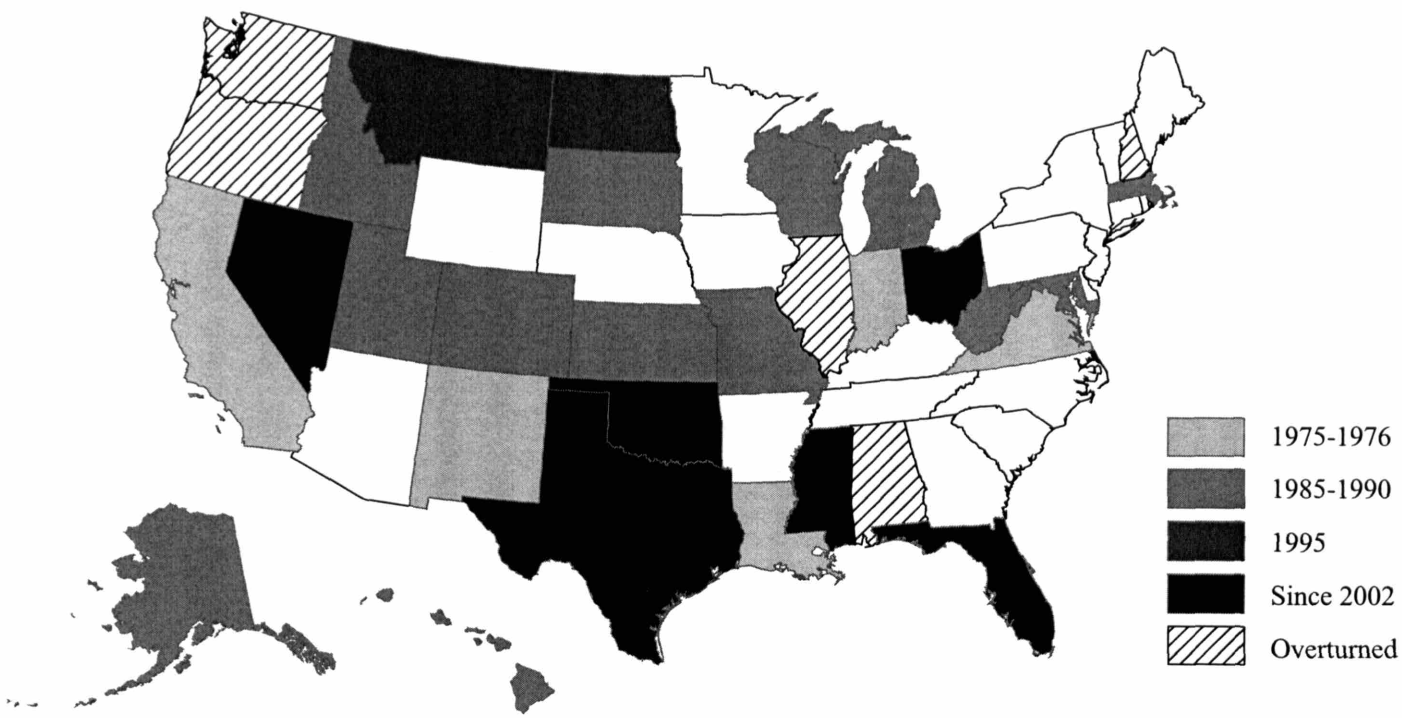


Table 3.1: Effect of damage cap on closed malpractice claims, 1991-2000

| | Payments Per Capita | Claims Per Capita | $250K+ Claims Per Capita | Mean Payment | Median Payment |
|---|---|---|---|---|---|
| Damage cap | -0.27* | 0.00 | -0.39** | -0.33** | -0.28** |
| | (0.10) | (0.13) | (0.14) | (0.11) | (0.10) |
| $R^2$ | 0.49 | 0.35 | 0.46 | 0.50 | 0.46 |
| N | 1556 | 1556 | 1294 | 1556 | 1556 |

* significant at 5%; ** significant at 1%

*Notes:* Each column represents a separate regression. Reported coefficients are estimated from state-level regressions of various log dependent variables on a damage cap dummy variable and a set of controls. Controls include year fixed effects, income, employment, and social security payments per capita, farm employment share, and population density in 1991. Regressions in columns (1)-(3) are weighted by population and regressions in columns (4)-(5) are weighted by the number of claims. Standard errors corrected for within-state correlation in the error term are reported in parentheses.

Table 3.2: County summary statistics

| | Population Per Square Mile, 1970 | | 1970 | | 2000 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Min | Max | Total Active MDs | Population (000) | Total Active MDs | Population (000) | Income ($000) | Percent Employed | Social Security Income ($) | Farm Employment Share |
| *A. Statewide:* | | | | | | | | | | |
| All counties | 0 | 54,759 | 1644 | 878 | 3170 | 1057 | 15.7 | 0.52 | 0.86 | 0.12 |
| | | | (3048) | (1548) | (5477) | (1889) | (9.0) | (0.16) | (0.65) | (0.12) |
| *B. Population density:* | | | | | | | | | | |
| Quartile 1 | 0 | 14 | 11 | 18 | 54 | 44 | 12.2 | 0.47 | 1.43 | 0.21 |
| | | | (14) | (18) | (68) | (48) | (6.2) | (0.13) | (2.19) | (0.14) |
| Quartile 2 | 14 | 32 | 19 | 25 | 83 | 54 | 11.9 | 0.44 | 1.04 | 0.15 |
| | | | (29) | (21) | (165) | (65) | (6.2) | (0.10) | (0.72) | (0.09) |
| Quartile 3 | 32 | 72 | 91 | 87 | 496 | 301 | 12.9 | 0.44 | 0.89 | 0.10 |
| | | | (189) | (140) | (844) | (500) | (6.6) | (0.10) | (0.53) | (0.07) |
| Quartile 4 | 72 | 54,759 | 2084 | 1105 | 4083 | 1331 | 16.7 | 0.54 | 0.81 | 0.03 |
| | | | (3314) | (1680) | (6005) | (2085) | (9.5) | (0.17) | (0.46) | (0.03) |
| | | | | | Population Density (residents per square mile) | | | | | |
| *C. Damage cap in 2000:* | | | | | | | | | | |
| Cap | | | | | 51 | | 16.0 | 0.53 | 0.81 | 0.13 |
| | | | | | (250) | | (8.8) | (0.13) | (0.67) | (0.12) |
| No Cap | | | | | 90 | | 15.6 | 0.52 | 0.88 | 0.12 |
| | | | | | (472) | | (9.2) | (0.18) | (0.63) | (0.11) |

*Notes:* Population-weighted, county-level means are reported. Standard deviations are reported in parentheses. Income meaures are in per capita terms. Social security income includes old age, survivors, and disability insurance payments as well as railroad retirement and disability payments, and serves as a proxy for the percentage of county residents that are retired.

Table 3.3: Effect of damage cap on doctors per 100,000 residents

| | Percent Represented | | With Cap | | Without Cap | | |
|---|---|---|---|---|---|---|---|
| | Counties | Population | 1970 | 2000 | 1970 | 2000 | Diff-in-Diff |
| | | | *I. Total Active MDs* | | | | |
| A. Statewide | 1.00 | 1.00 | 143 | 259 | 134 | 249 | 2 |
| B. County population density: | | | | | | | |
| Quartile 1 | 0.25 | 0.03 | 54 | 110 | 55 | 97 | 14 |
| Quartile 2 | 0.25 | 0.06 | 66 | 118 | 53 | 88 | 17 |
| Quartile 3 | 0.25 | 0.12 | 82 | 139 | 66 | 134 | -11 |
| Quartile 4 | 0.25 | 0.79 | 159 | 294 | 154 | 290 | -1 |
| | | | *II. Office-Based Patient Care MDs* *(a) General/Family Practice Physicians* | | | | |
| A. Statewide | 1.00 | 1.00 | 26 | 25 | 24 | 23 | 0.2 |
| B. County population density: | | | | | | | |
| Quartile 1 | 0.25 | 0.03 | 35 | 32 | 34 | 28 | 3.2 |
| Quartile 2 | 0.25 | 0.06 | 30 | 27 | 33 | 26 | 3.0 |
| Quartile 3 | 0.25 | 0.12 | 30 | 25 | 28 | 25 | -1.7 |
| Quartile 4 | 0.25 | 0.79 | 24 | 23 | 23 | 22 | -0.1 |
| | | | *(b) Medical Specialists, Including Internists and Pediatricians* | | | | |
| A. Statewide | 1.00 | 1.00 | 22 | 63 | 21 | 61 | 0.4 |
| B. County population density: | | | | | | | |
| Quartile 1 | 0.25 | 0.03 | 3 | 19 | 4 | 20 | 0.4 |
| Quartile 2 | 0.25 | 0.06 | 8 | 25 | 4 | 18 | 2.2 |
| Quartile 3 | 0.25 | 0.12 | 11 | 33 | 8 | 32 | -1.9 |
| Quartile 4 | 0.25 | 0.79 | 25 | 72 | 25 | 71 | -0.1 |
| | | | *(c) Surgical Specialists* | | | | |
| A. Statewide | 1.00 | 1.00 | 31 | 43 | 28 | 42 | -2.0 |
| B. County population density: | | | | | | | |
| Quartile 1 | 0.25 | 0.03 | 8 | 19 | 9 | 18 | 2.3 |
| Quartile 2 | 0.25 | 0.06 | 15 | 23 | 8 | 16 | 0.6 |
| Quartile 3 | 0.25 | 0.12 | 19 | 27 | 15 | 26 | -3.2 |
| Quartile 4 | 0.25 | 0.79 | 34 | 47 | 32 | 48 | -2.7 |
| | | | *(d) Support Specialists* | | | | |
| A. Statewide | 1.00 | 1.00 | 19 | 48 | 16 | 44 | 1.2 |
| B. County population density: | | | | | | | |
| Quartile 1 | 0.25 | 0.03 | 3 | 20 | 4 | 16 | 4.8 |
| Quartile 2 | 0.25 | 0.06 | 7 | 23 | 4 | 14 | 6.2 |
| Quartile 3 | 0.25 | 0.12 | 9 | 25 | 7 | 24 | -2.1 |
| Quartile 4 | 0.25 | 0.79 | 22 | 55 | 19 | 52 | 0.4 |

*Notes:* Population-weighted, county-level mean physician-to-population ratios are reported in the third through sixth columns. The final column reports a simple difference-in-difference comparison.

Table 3.4: Effect of damage cap on log doctors per resident

| | Percent Represented | | | | | | |
|---|---|---|---|---|---|---|---|
| | Counties | Population | (1) | (2) | (3) | (4) | (5) |
| *A. Statewide* | 1.00 | 1.00 | -0.014 | 0.014 | -0.014 | -0.007+ | -0.003 |
| | | | (0.023) | (0.013) | (0.011) | (0.004) | (0.003) |
| *B. County population density:* | | | | | | | |
| Quartile 1 | 0.25 | 0.03 | 0.071+ | 0.045* | 0.044** | 0.031* | 0.040** |
| | | | (0.042) | (0.023) | (0.015) | (0.015) | (0.012) |
| Quartile 2 | 0.25 | 0.06 | 0.060 | -0.005 | 0.008 | -0.005 | -0.011 |
| | | | (0.037) | (0.018) | (0.015) | (0.009) | (0.011) |
| Quartile 3 | 0.25 | 0.12 | -0.041 | 0.036 | 0.004 | 0.006 | 0.011 |
| | | | (0.033) | (0.031) | (0.019) | (0.006) | (0.007) |
| Quartile 4 | 0.25 | 0.79 | -0.009 | 0.010 | -0.020+ | -0.010* | -0.008** |
| | | | (0.026) | (0.013) | (0.012) | (0.004) | (0.003) |
| Level of observations | | | State | State | County | County | County |
| Fixed effects | | | State | State | County | County | County |
| State-specific linear time trend | | | | X | X | X | X |
| Covariates | | | | | X | X | X |
| AR(1) error term, state-specific $\rho_i$ | | | | | | X | |
| First difference | | | | | | | X |

+ significant at 10%; * significant at 5%; ** significant at 1%

*Notes:* Each coefficient represents a separate regression, where each row corresponds to a different sample of counties, and each column corresponds to a different specification. Reported coefficients are estimated from regressions of log physician-to-population ratios on a damage cap dummy variable and a set of controls. Only the coefficient corresponding to the damage cap dummy variable is reported. Controls in all regressions include state/county (as indicated) and year fixed effects. Where indicated, controls also include a state-specific linear trend and a set of covariates: income, employment, and social security payments per capita, and farm employment share. All regressions are weighted by population. Standard errors corrected for within-state correlation in the error term are reported in parentheses, except in Column (4) where reported standard errors corrected for state-by-year correlation in the error term . All regressions are for 1970-2000 (except 84, 90, 94).

Table 3.5: Effect of damage cap on log doctors per resident -- Dynamics

| | | County Population Density | | | |
|---|---|---|---|---|---|
| | Statewide | Quartile 1 | Quartile 2 | Quartile 3 | Quartile 4 |
| Years since cap adoption: | | | | | |
| -5 to -1 | 0.016 | 0.025 | 0.012 | 0.019 | 0.013 |
| | (0.010) | (0.031) | (0.014) | (0.018) | (0.011) |
| 0 to 5 | 0.002 | 0.061+ | 0.025 | 0.023 | -0.006 |
| | (0.014) | (0.032) | (0.018) | (0.024) | (0.018) |
| 6 to 10 | 0.014 | 0.127** | 0.029 | 0.039 | 0.007 |
| | (0.017) | (0.030) | (0.023) | (0.029) | (0.026) |
| 11+ | 0.021 | 0.118** | 0.048 | 0.045 | 0.019 |
| | (0.018) | (0.043) | (0.037) | (0.031) | (0.031) |
| N | 80,572 | 17,722 | 20,743 | 21,358 | 20,713 |

+ significant at 10%; ** significant at 1%

*Notes:* Each column represents a separate regression on different samples of counties. Reported coefficients are estimated from regressions of log physician-to-population ratios on a set of damage cap dummy variables, interacted with the number of years since the cap was adopted, and a set of controls. Controls include county and year fixed effects, a state-specific linear trend, and a set of covariates: income, employment, and social security payments per capita, and farm employment share. All regressions are weighted by population. Standard errors corrected for within-state correlation in the error term are reported in parentheses. All regressions are for 1970-2000 (except 84, 90, 94).

Table 3.6: Effect of damage cap and other tort reforms on log doctors per resident

| | | County Population Density | | | |
|---|---|---|---|---|---|
| | Statewide | Quartile 1 | Quartile 2 | Quartile 3 | Quartile 4 |
| Damage cap | -0.014 | 0.045** | 0.008 | 0.005 | -0.019 |
| | (0.012) | (0.016) | (0.016) | (0.019) | (0.013) |
| Other tort reforms | 0.000 | -0.008 | 0.001 | -0.002 | -0.005 |
| | (0.007) | (0.018) | (0.018) | (0.009) | (0.009) |
| N | 80,572 | 17,722 | 20,743 | 21,358 | 20,713 |

** significant at 1%

*Notes:* Each column represents a separate regressionon different samples of counties. Reported coefficients are estimated from couty-level regressions of log physician-to-population ratios on tort reform indicator variables and a set of controls. The other tort reform indicator includes collateral-source reform, joint-and-several liability, attorney fee restrictions, pretrial screening, and periodic payments. Controls include county and year fixed effects, a state-specific linear trend, and a set of covariates: income, employment, and social security payments per capita, and farm employment share. All regressions are weighted by population. Standard errors corrected for within-state correlation in the error term are reported in parentheses. All regressions are for 1970-2000 (except 84, 90, 94).

Table 3.7: Effect of damage cap on log doctors per resident -- By specialty

| | | County Population Density | | | |
|---|---|---|---|---|---|
| | Statewide | Quartile 1 | Quartile 2 | Quartile 3 | Quartile 4 |
| *Dependent variable:* | | | | | |
| General/family practice physicians | -0.029 | 0.008 | -0.017 | 0.021 | -0.029 |
| | (0.021) | (0.019) | (0.028) | (0.029) | (0.026) |
| Medical specialists, including internists and pediatricians | -0.015 | 0.058 | -0.059 | 0.013 | -0.018 |
| | (0.017) | (0.043) | (0.053) | (0.029) | (0.018) |
| Surgical specialists | -0.020 | 0.098** | 0.032 | -0.031 | -0.030* |
| | (0.013) | (0.021) | (0.030) | (0.022) | (0.014) |
| Support specialists | 0.005 | 0.111* | 0.027 | 0.019 | -0.004 |
| | (0.018) | (0.056) | (0.045) | (0.032) | (0.021) |

* significant at 5%; ** significant at 1%

*Notes:* Each coefficient represents a separate regression, where each row corresponds to a different dependent variable, and each column corresponds to a different sample of counties. Reported coefficients are estimated from couty-level regressions of log physician-to-population ratios on damage cap and other tort reform indicator variables and a set of controls. Only the coefficient corresponding to the damage cap indicator variable is reported. The other tort reform indicator includes collateral-source reform, joint-and-several liability, attorney fee restrictions, pretrial screening, and periodic payments. Controls include county and year fixed effects, a state-specific linear trend, and a set of covariates: income, employment, and social security payments per capita, and farm employment share. All regressions are weighted by population. Standard errors corrected for within-state correlation in the error term are reported in parentheses. All regressions are for 1970, 1975, 1980, 1985-2000 (except 90, 94).

Table 3.8: Economics of medical practice

| | General/Family Practice | | | Surgical Specialties (except Ob/gyn) | | |
|---|---|---|---|---|---|---|
| | Rural | <1M | 1M+ | Rural | <1M | 1M+ |
| **A. Quantity** | | | | | | |
| Hours per week | 62.2 | 56.4 | 55.9 | 63.0 | 60.2 | 56.9 |
| | (1.5) | (1.0) | (1.2) | (2.1) | (0.8) | (0.8) |
| Patient visits per week | 163.4 | 134.3 | 116.3 | 113.0 | 100.7 | 91.8 |
| | (6.4) | (4.1) | (4.2) | (8.0) | (2.2) | (2.2) |
| **B. Revenue** | | | | | | |
| Fee for office visit with established patient ($) | 38.8 | 44.6 | 53.3 | 38.4 | 57.9 | 66.7 |
| | (1.2) | (1.1) | (2.2) | (2.0) | (2.0) | (1.9) |
| **C. Fixed costs** | | | | | | |
| Liability premium ($000) | 8.7 | 11.4 | 9.5 | 20.3 | 21.7 | 23.2 |
| | (0.7) | (2.5) | (1.4) | (2.2) | (1.0) | (1.1) |
| Office expense ($000) | 39.9 | 47.8 | 44.9 | 53.1 | 68.6 | 69.5 |
| | (8.0) | (4.5) | (5.1) | (13.6) | (5.6) | (5.5) |
| $\frac{\text{Premium*}}{\text{Premium + Office}}$ | 0.179 | 0.193 | 0.175 | 0.277 | 0.240 | 0.250 |
| | (0.016) | (0.034) | (0.022) | (0.026) | (0.011) | (0.011) |
| $H_0$: $X_r < X_m$, p-value | | 0.645 | 0.442 | | 0.096 | 0.170 |
| **D. Variable cost** | | | | | | |
| Medical supplies ($000) | 20.6 | 21.5 | 22.5 | 19.5 | 23.0 | 23.5 |
| | (2.7) | (2.3) | (4.7) | (4.8) | (3.1) | (3.0) |
| **E. Profit** | | | | | | |
| Net income before taxes ($000) | 127.1 | 119.6 | 119.9 | 254.1 | 274.8 | 239.6 |
| | (5.8) | (4.2) | (5.8) | (29.1) | (10.0) | (8.1) |

*Assumes premium and office expense are uncorrelated.

*Notes:* Physician means are reported, with standard errors in parentheses. Hours, fees, visits are for 1995; net income and costs for 1994. Cost data is for self-employed physicians in a solo medical practice. Office expense includes rent, mortgage interest, depreciation on medical buildings used in the physician's practice, utilities, and telephone. Medical materials and supplies includes drugs, x-ray films, and other disposable medical products.

*Source:* Socioeconomic Characteristics of Medical Practice, American Medical Association, 1996.

Table 3.9: Geographic distribution of closed malpractice claims

| | (1) Payments Per Doctor | (2) Claims Per Doctors | (3) $250K+ Claims Per Doctor | (4) Payments Per Capita | (5) Claims Per Capita | (6) $250K+ Claims Per Capita | (7) Mean Payment | (8) Median Payment |
|---|---|---|---|---|---|---|---|---|
| Population density Quartile 1 | 0.21* | 0.51** | 0.38** | -1.11** | -0.78** | -0.90** | -0.31** | 0.06 |
| | (0.09) | (0.07) | (0.08) | (0.14) | (0.11) | (0.22) | (0.06) | (0.06) |
| Population density Quartile 2 | 0.06 | 0.42** | 0.34* | -1.24** | -0.83** | -0.84** | -0.31** | 0.03 |
| | (0.12) | (0.11) | (0.14) | (0.13) | (0.08) | (0.09) | (0.06) | (0.04) |
| Population density Quartile 3 | 0.12 | 0.31** | 0.13 | -0.73** | -0.53** | -0.74** | -0.19** | -0.12* |
| | (0.08) | (0.09) | (0.08) | (0.08) | (0.06) | (0.08) | (0.05) | (0.05) |
| Damage cap * Quartile 1 | 0.17 | 0.01 | 0.04 | 0.22 | 0.07 | 0.31 | 0.17+ | -0.05 |
| | (0.15) | (0.13) | (0.14) | (0.20) | (0.15) | (0.25) | (0.09) | (0.12) |
| Damage cap * Quartile 2 | 0.11 | -0.01 | -0.06 | 0.35 | 0.22 | 0.10 | 0.09 | -0.19+ |
| | (0.14) | (0.14) | (0.16) | (0.23) | (0.15) | (0.26) | (0.08) | (0.10) |
| Damage cap * Quartile 3 | 0.19+ | 0.11 | 0.18+ | 0.26 | 0.17 | 0.32* | 0.12+ | 0.06 |
| | (0.10) | (0.10) | (0.10) | (0.18) | (0.12) | (0.15) | (0.07) | (0.06) |
| $R^2$ | 0.70 | 0.82 | 0.74 | 0.72 | 0.80 | 0.73 | 0.72 | 0.69 |
| N | 1323 | 1323 | 1115 | 1476 | 1476 | 1227 | 1476 | 1476 |

+ significant at 10%; * significant at 5%; ** significant at 1%

*Notes:* Each column represents a separate regression. Reported coefficients are estimated from state-population-density-quartile-level regressions of various log dependent variables on population-density-quartile dummy variables, interacted with a damage cap dummy variable, and a set of controls. Population density quartile 4 areas form the omitted category. States that do not contain a population density quartile 4 area are excluded from the analysis. Controls include state and year fixed effects. Regressions in columns (1)-(3) are weighted by the number of doctors, regressions in columns (4)-(6) are weighted by population, and regressions in columns (7)-(8) are weighted by the number of claims. Standard errors corrected for within-state correlation in the error term are reported in parentheses. Regressions in columns (1)-(3) are for 1992-2000, and regressions in columns (4)-(8) are for 1991-2003.