Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 178 (2018) 153–189

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Financial contracting with enforcement externalities ☆

Lukasz A. Drozd [a], Ricardo Serrano-Padial [b,*]

[a] *Federal Reserve Bank of Philadelphia, Ten Independence Mall, Philadelphia, PA 19106, United States of America*
[b] *School of Economics, Drexel University, 3220 Market St., Philadelphia, PA 19147, United States of America*

Received 1 August 2017; final version received 31 August 2018; accepted 5 September 2018
Available online 13 September 2018

Abstract

We study the negative feedback loop between the aggregate default rate and the efficacy of enforcement in a model of debt-financed entrepreneurial activity. The novel feature of our model is that enforcement capacity is accumulated ex ante and thus subject to depletion ex post. We characterize the effect of shocks that deplete enforcement resources on the aggregate default rate and credit supply. In the model default decisions by entrepreneurs are strategic complements, leading to multiple equilibria. We propose a global game selection to overcome equilibrium indeterminacy and show how shocks that deplete enforcement capacity can lead to a spike in the aggregate default rate and trigger credit rationing.

JEL classification: D82; D84; D86; G21; O16; C72

Keywords: Contract enforcement; Default spillovers; Credit crunch; Credit cycles; Global games; Heterogeneity

☆ An earlier version was circulated under the title "Credit Enforcement Cycles." We thank Jason Donaldson, Pedro Gete, Patrick Kehoe, Marzena Rostek, Joel Sobel, Marek Weretka, and audiences at CEMFI, CIDE, Drexel, Georgetown, Indiana, NBER micro-to-macro group, UNC-Chapel Hill, Oxford, University of Pennsylvania, Pompeu Fabra, Philadelphia Fed, Queen Mary, Royal Holloway, St. Louis Fed, Surrey, Temple, Virginia, Wharton, and UW-Madison for their comments. Drozd acknowledges the financial support of the Vanguard Research Fellowship from Rodney L. White Center for Financial Research. Disclaimer: The views expressed in these papers are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.
* Corresponding author.
E-mail addresses: lukaszadrozd@gmail.com (L.A. Drozd), rspadial@gmail.com (R. Serrano-Padial).

https://doi.org/10.1016/j.jet.2018.09.002

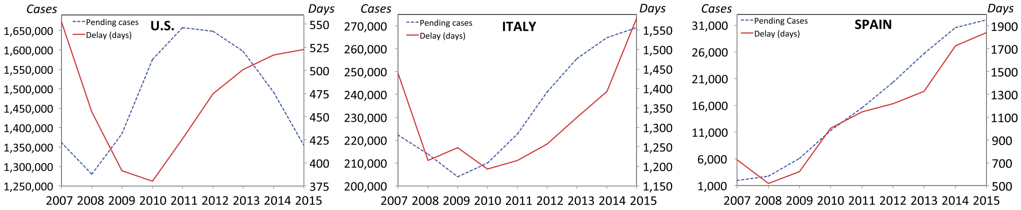

Fig. 1. Default Cases Pending in Court (left axis) and Enforcement Delay (right axis). Notes: The figure illustrates pending individual and corporate bankruptcies in the U.S., corporate property executions in Italy, and commercial bankruptcies in Spain (see data sources in the Online Appendix). The red line plots a popular aggregate measure of efficacy of court enforcement: $Delay_t = 365$*(Pending cases as of the beginning of the year + pending cases as of the end of the year)/(new cases + closed cases). In stationary equilibrium the measure returns the approximate number of days it takes to solve a case. A persistent increase in this measure indicates lower efficacy of the court system. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

1. Introduction

Financial crises are in part propagated by disrupted enforcement and liquidation mechanisms in credit markets. For example, during the 2007 financial crisis, the lack of timely enforcement has been deemed central to the understanding of the U.S. foreclosure glut as well as the depressed credit conditions in Italy and Spain.[1] In fact, in all three countries the crisis resulted in a groundswell of pending court cases that had a well-documented negative impact on the efficacy of enforcement (Fig. 1). The 2007 crisis is hardly an isolated episode. The role of enforcement has long been stressed as important in propagating financial crises by public policy practitioners.[2]

A number of empirical studies provide direct evidence in support of causal linkages between enforcement and credit market outcomes. For example, Schiantarelli et al. (2016) look at Italian firms that simultaneously owed loans both to banks in jurisdictions with weak enforcement and to banks in jurisdictions with strong enforcement to conclude that weaker enforcement induced strategic default.[3] In a similar vein, Iverson (2017) uses variation in bankruptcy court caseload to show that delayed enforcement has been associated with higher creditor losses. Based on a natural experiment, Mayer et al. (2014) argue that strategic considerations were an important factor fueling the U.S. subprime crisis. Finally, Ponticelli and Alencar (2016) and Rodano et al. (2016) relate enforcement to the availability of credit, supporting the well-documented cross-sectional correlations (Jappelli et al., 2005; La Porta et al., 1998; Djankov et al., 2007, 2008; Bae and Goyal, 2009).

Despite renewed interest in the role of enforcement mechanisms in propagating financial crises, and mounting empirical evidence, theoretical treatments of the issue remain scant. In par-

[1] See Cordell et al. (2015) for a discussion of foreclosure delays in the U.S. during the great recession. In this context, Chan et al. (2016) estimate that a foreclosure delay of nine months is associated with a 40% higher default rate on underwater mortgages while controlling for a wide array of confounding factors. Zhu and Pace (2015) find that a delay of three months increases the probability of default by 30%. Using a quantitative model, Herkenhoff and Ohanian (2015) estimate that foreclosure delays added 25% to the delinquency rate during the crisis. Carpinelli et al. (2016) document that judiciary backlogs were the primary factor behind the slower resolution of non-performing loans during this time period in Italy. The Bank of Italy (2013) documents that the average time to write off bad debt from banks' balance sheets went up from less than 4 years to over 6 years in Italy between 2007 and 2011.

[2] For example, Woo (2000), Enoch et al. (2001) and also Krueger and Tornell (1999).

[3] Ippolito et al. (2016) study repayment of loans owed to European banks differentially affected by the crisis and report similar findings. Favara et al. (2012) provide related cross-country evidence.

ticular, the workhorse financial contracting model by Gale and Hellwig (1985), building on the seminal work by Townsend (1979), and the modern applications of their framework to macroeconomic modeling (e.g., Bernanke et al. (1999) and Christiano et al., 2014), all assume that enforcement is a nondepletable resource. The few papers that do speak to the issue are too specialized to provide ground for generalizations.[4] As a result, theoretical mechanisms governing the link between enforcement and credit markets are not well understood, and empirical studies alone fall short of establishing the overall contribution of enforcement to shock propagation.

To fill this gap, we enrich the canonical model of credit-financed entrepreneurial activity à la Gale and Hellwig (1985) by incorporating enforcement as a finite resource that is accumulated ex ante by a benevolent government and hence subject to depletion ex post. We focus on debt contracts and use the model to study the equilibrium effect of the depletion of enforcement resources on credit supply and debt repayment. In our model entrepreneurs invest loans on risky projects that serve as collateral for the loans. Upon privately learning their project returns they decide whether to repay the loan or default. The enforcement (liquidation) of collateral is then determined by the amount of enforcement resources and by the aggregate default rate. The depletability of enforcement via high default rates introduces a negative feedback loop in which the expectation of weak enforcement incentivizes default among borrowers, which further depletes enforcement resources, making repayment decisions strategic complements. This amplifies shocks that initially deplete enforcement resources in the economy.

We study the equilibrium implications of our model under common knowledge of enforcement fundamentals and show that strategic complementarities can lead to multiplicity of equilibria, deeming the aggregate default rate undetermined. We overcome equilibrium indeterminacy by introducing noisy signals about enforcement as is done in the global games literature (Carlsson and van Damme, 1993; Morris and Shin, 1998, 2003). Specifically, we build on the equilibrium selection approach of Frankel et al. (2003) to show that a unique equilibrium arises as signal noise vanishes. From a technical perspective, we tackle the challenges posed by agent heterogeneity in global games by generalizing the results from Sákovics and Steiner (2012) and Frankel et al. (2003) to characterize equilibrium.

Substantively, we establish three key results. First, we pin down the necessary and sufficient conditions for equilibrium multiplicity under common knowledge and characterize the resulting equilibrium set. Second, in the absence of common knowledge, we find that equilibrium default rates are fragile in the following sense: The equilibrium default rate discontinuously jumps when enforcement resources fall below a certain threshold. This is due to a partial clustering of default decisions in equilibrium. Importantly, in the context of the equilibrium under common knowledge, the jump in default rates corresponds to a switch from the equilibrium with the lowest default rate and strongest enforcement to the equilibrium with the highest default rate and weakest enforcement. Finally, since credit is the only tool lenders have to avoid a spike in defaults caused by worsening enforcement, the credit supply tightens in response to shocks that deplete enforcement resources.

The discontinuous response of the aggregate default rate is driven by the contagion of default decisions among entrepreneurs. Specifically, agents whose payoff from defaulting is relatively low (i.e., those with high project returns), and hence have a relatively low intrinsic propensity to default, rationally anticipate that, whenever they find default optimal, so must the agents with a

[4] There are two relevant studies. Bond and Rai (2009) study a micro-finance model featuring a population of homogeneous borrowers. Arellano and Kocherlakota (2014) use a model with two firms to illustrate how coordinated defaults in the private sector can trigger sovereign default when asset liquidation is exogenously limited to one firm.

higher propensity to default (low project returns). As a result, agents who are *less* prone to default also default because they expect the default rate to be higher and enforcement to be weaker. This effect results in a formation of clusters of agents who have different intrinsic propensities to default and yet follow the exact same strategy.

We assess the quantitative implications of our model for the propagation of shocks by calibrating it to the U.S. data. The calibrated model exhibits enforcement costs of about 0.3% of total debt outstanding and implies that, in normal times, about 83% of enforcement resources are used to liquidate projects. Nonetheless, enforcement resources are scarce in the sense that even a small shock affects credit supply due to their proximity to the threshold associated to the discontinuous jump in default rates. To demonstrate the magnitude of shock propagation in the model, we consider both a large shock on enforcement equivalent to the surge in default rates experienced by the U.S. in the great recession, and an average shock that resembles ordinary recessions. We find that credit contracts by 30% in response to the large shock and by 5% due to the average shock.

The mechanism we identify in this paper is relevant from a policy perspective. When enforcement is depleted, the model implies that the most effective way to alleviate the shock might be to prevent default by agents who are most prone to default, even if they impose the highest costs in terms of debt forgiveness. This is because such an intervention mitigates the aforementioned snowballing effect and prevents other agents from defaulting as well.[5] A recent example of a policy in this spirit is the 2009 Home Affordable Modification Program (HAMP), which targeted U.S. homeowners who could prove that they were underwater and had not yet defaulted. The success of this program has been well-documented by Agarwal et al. (2017).

Related literature The most closely related paper to ours is Bond and Rai (2009), who focus a micro-finance model featuring a population of homogeneous borrowers. Their analysis shows the possibility of a "borrower run" on a single lender driven by the link between the default rate and the continuation value that the lender may be able to deliver after the run takes place and depletes its funds. As in our model, they tackle equilibrium multiplicity by using global games. However, whereas they focus on the survival of the lending institution as a function of its financial health and the amount of credit in a homogeneous agent environment, our goal is to characterize both default and credit supply decisions in an economy with heterogeneous agents. Our model features stochastic returns to investment, which introduces private information due to nonstrategic default, whereas in their model default is always strategic. In the context of sovereign debt crises, Arellano and Kocherlakota (2014) use a related model with two firms to illustrate how coordinated defaults in the private sector can trigger self-fulfilling sovereign defaults when asset liquidation is exogenously limited to one firm.[6]

Independently, our paper complements the literature on fire sales and the propagation of shocks (Shleifer and Vishny, 2011). Enforcement capacity in this context corresponds to the ability of financial intermediaries to absorb widespread liquidation of collateral. The connection between enforcement levels and default incentives then finds its parallel in the link between the price of assets and incentives to sell collateral. Clustering loosely corresponds to fire sale of assets. The model by Dávila and Korinek (2018) provides a recent example of such a dynamic.

[5] This result echoes a related result established by Sákovics and Steiner (2012) in their setup.

[6] Coordination problems arising from limited enforcement have also been studied in the crime (Bond and Hagerty, 2010) and tax evasion literatures (Bassetto and Phelan, 2008).

Our paper makes a technical contribution to the study of global games with agent heterogeneity. In particular, we extend the characterization result established by Sákovics and Steiner (2012). They focus on a particular environment in which, by construction, all agents take the same action in equilibrium. In contrast, we characterize the endogenous emergence of decision clusters as a function of agent heterogeneity.[7]

Finally, our paper adds to the analysis of the impact of agent heterogeneity on coordination in the presence of strategic complementarities, such as the currency speculation model of Corsetti et al. (2004) and the analysis of bank bailouts by Dávila and Walther (2018).

2. Theory

This section lays out the baseline model and derives our main analytic results. We first derive the conditions for equilibrium multiplicity and introduce the global games selection that obtains uniqueness. We then characterize the equilibrium default rate and the link between enforcement and credit supply. In Section 3 we study quantitative implications of these results.

2.1. Environment

The economy is populated by a large number of deep-pocketed lenders and a continuum (of measure one) of risk-neutral and ex ante identical entrepreneurs who have an investment project. The investment project yields a random return $w \geq 0$, which is a continuously distributed i.i.d. random variable with cdf $F(w)$, with $E(w) > 1$ and density f with full support on $[0, \infty)$. (The law of large numbers is assumed.)

Entrepreneurs seek to leverage their own equity y invested in the project by taking a loan from lenders. Lenders are risk-neutral Bertrand competitors and effectively maximize the ex ante utility of entrepreneurs subject to a zero profit condition (in expectation).[8]

A loan is described by a tuple $(b, \bar{b})$, where b is principal and $\bar{b}$ is principal plus interest. Accordingly, total investment in each project is $y + b$, and the output of a project with return w is $(y + b)w$. Entrepreneur's net payoff under repayment is $(y + b)w - \bar{b}$ and hence the project is profitable if $w \geq \bar{w} := \bar{b}/(y + b)$.

Loans are defaultable and entrepreneurs use projects as a collateral to secure funding. In case of default, the project is liquidated. The liquidation value of the project is $\mu(y + b)w$, where $\mu < 1$ is a parameter that determines the deadweight loss from liquidation. Importantly, liquidation can be either *enforced* or *non-enforced*. If liquidation is enforced, the entire liquidation value is taken over by the lender. If liquidation is not enforced, the entrepreneur retains a portion $0 < \gamma < 1$ of the liquidation value at the lender's expense. The distinction between enforced and non-enforced liquidation is meant to convey the basic idea that entrepreneurs may extract rents when liquidation is delayed or disrupted.[9] The probability that a liquidation following default is

[7] In the class of games studied by Sákovics and Steiner (2012) payoff heterogeneity is restricted to guarantee that all agents cluster on the same strategy.

[8] Debt financing is assumed. In particular, we do not explore optimality of debt contracts as Gale and Hellwig (1985) do, given that enforcement is stochastic in our model, whereas they only consider deterministic monitoring. On positive grounds, debt contracts are ubiquitous, which motivates our approach. In addition, an extensive finance literature highlights reasons why debt contracts may prevail.

[9] For example, entrepreneurs may divert resources before the project is taken over by lenders. As a result, delayed enforcement may eat up the value of the project that lenders receive, meanwhile serving the entrepreneur. Alternatively,

enforced, denoted by $P(X, \psi)$, depends positively on the economywide enforcement capacity X and negatively on the aggregate default rate ψ.

Enforcement capacity X is fixed at the lending stage, and it is optimally chosen by a planner before the lending market opens. When choosing enforcement capacity, the planner maximizes the ex ante expected utility of entrepreneurs in the economy, net of capacity buildup costs. The planner's choice captures the idea that the enforcement infrastructure, e.g., the bankruptcy court system or the number of judges, is typically established by the government, fixed ex post, and costly to develop ex ante.

To analyze tractably how fluctuations in enforcement might affect credit supply, we introduce a binary aggregate shock that randomly depletes enforcement capacity before the credit market opens. This simplifies the analysis while capturing the essence of shock propagation through the depletion of enforcement resources. Accordingly, enforcement capacity available to sustain loans at the lending stage is $X = X_o - s$, where X_o is the ex-ante choice of the planner and s is a random aggregate shock. The shock captures the idea that poor performance of pre-existing loans in practice tangles enforcement infrastructure and diverts enforcement resources away from the market for new loans. (We obtain qualitatively similar results if we instead introduce a shock to the distribution of project returns.)

The timing of events is as follows: 1) the planner chooses initial capacity X_o and the shock s is realized; 2) the credit market opens and lenders compete to extend loans to entrepreneurs and investment takes place; 3) entrepreneurs privately observe their project returns and simultaneously decide whether to repay their loans or default; 4) projects are liquidated and enforced with probability P.

2.2. *Contract enforcement*

We solve the model using backward induction and hence begin our analysis from the last stage. We refer to this stage as the *enforcement game*. The solution to the enforcement game feeds into the credit market and the optimal choice of enforcement capacity, which we analyze subsequently.

2.2.1. *Enforcement technology*

Liquidated projects are assumed to be randomly selected for enforcement until enforcement capacity is fully exhausted (or there are not more projects to enforce). This is justified by the fact that w is privately observed by entrepreneurs.[10] Consequently, enforcement probability P only depends on enforcement capacity X and the measure of agents who default ψ. Without loss, we normalize the domain of X to be between 0 and 1 at the enforcement stage. Specifically, if $X_o - s < 0$, we set $X = 0$. Analogously, if $X_o - s > 1$, we set $X = 1$.

bargaining power in an out-of-court *renegotiation* settlement may depend on the threat-point of legal enforcement, which depends on its expected efficacy. Favara et al. (2012) provide evidence that weak enforcement indeed leads to a higher incidence of renegotiation between defaulting firms and creditors.

[10] Agents cannot be selected for enforcement based on payoff-irrelevant characteristics, such as the address, names, etc. Such practices would violate nondiscrimination laws and hence we assume away such a possibility (e.g., Code of Federal Regulations Title 12, Chapter V, Part 528, Section 528.9). Segmented enforcement is beneficial in our environment and is known to unravel the complementarities. This has been shown by Carrasco and Salgado (2014).

For expositional purposes, we assume that enforcement capacity X represents the measure of liquidated projects that can be enforced. In such a case, it is clear that

$$P(X, \psi) = \min\{1, X/\psi\}. \tag{1}$$

However, since our formal results do not require a specific functional form, we introduce the notion of *admissible enforcement probability* in the definition below. We use it to state them more generally.[11]

Definition 1. $P(X, \psi)$ is an admissible enforcement technology if:

(i) $P : [0, 1]^2 \to [0, 1]$ is a Lipschitz continuous function,
(ii) $P(X, \psi)$ is increasing in X, and strictly so at any (X, ψ) such that $P(X, \psi) < 1$,
(iii) $P(X, \psi)$ is decreasing in ψ, and strictly so at any (X, ψ) such that $P(X, \psi) \in (0, 1)$,
(iv) $P(0, \psi) = 0$ for all ψ, and if $\psi < 1$, there exists $X < 1$ such that $P(X', \psi) = 1$ for all $X' \geq X$.

2.2.2. Entrepreneur problem

In the enforcement game, loan terms $(b, \bar{b})$ and the enforcement capacity X are given and entrepreneurs simultaneously decide whether to pay back their loans or default after privately observing their project return w. To define the entrepreneur problem formally, let $a = 1$ denote repayment and $a = 0$ denote default. In addition, let $m = 1$ denote enforced liquidation and $m = 0$ denote non-enforced liquidation. Furthermore, recall the following. Given contract terms $(b, \bar{b})$, the repayment amount $\bar{b}$ corresponds to some minimum return $\bar{w} = \bar{b}/(y + b)$ necessary to repay the loan. Accordingly, we can equivalently represent loans by a tuple $(b, \bar{w})$, since an agent with $w < \bar{w}$ does not have enough funds to pay back the loan.

Formally, entrepreneur payoffs are

$$u(a, w, m, b, \bar{w}) := \begin{cases} (y + b)(w - \bar{w}) & a = 1 \\ (1 - m)\gamma\mu(y + b)w & a = 0. \end{cases} \tag{2}$$

Since m is a random variable determined by the enforcement probability P, entrepreneurs choose action a to maximize

$$\max_{a \in \{0,1\}} \{E(P)u(a, w, 1, b, \bar{w}) + (1 - E(P))u(a, w, 0, b, \bar{w})\}, \tag{3}$$

where $E(P)$ represents their individual expectation (belief) of enforcement probability P.

The linearity of the objective function implies that entrepreneurs' repayment decision is a cutoff rule with respect to expected enforcement probability $E(P)$. Specifically, as established in Lemma 1, entrepreneurs repay their loans if $E(P)$ exceeds a cutoff $\theta(w, \bar{w})$. We refer to $\theta(w, \bar{w})$ as the agent's *intrinsic propensity to default*. (To streamline the exposition, we relegate all proofs to the Appendix.)

[11] Allowing for different functional forms of P can help study the impact of different enforcement technologies on credit. For instance, to analyze the natural case in which the resources needed for liquidation go up with project size (i.e., with returns w) we could assume that P is strictly concave and decreasing in ψ (for the interval of ψ at which $P(\psi, X) > 0$). This is because, as we show below, the equilibrium default rate is given by $F(w)$, where w is the return associated with the largest project being liquidated, and hence a higher w is associated with a higher use of enforcement resources per project. We thank our anonymous referee for suggesting this generalization.

Lemma 1. *The debt repayment decision of an entrepreneur with a contract* $(b, \bar{w})$ *is given by*[12]

$$a = \begin{cases} 1 & \text{if } E(P) \geq \theta(w, \bar{w}) \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

where

$$\theta(w, \bar{w}) := 1 - \frac{1}{\mu\gamma}\left(1 - \frac{\bar{w}}{w}\right). \tag{5}$$

Entrepreneurs' repayment decisions determine the aggregate default rate in the economy, which is given by

$$\psi := \int_{\{w:a=0\}} dF(w). \tag{6}$$

In turn, the default rate affects enforcement probability $P(X, \psi)$. Accordingly, entrepreneurs' repayment decisions are strategic complements, since a lower expected enforcement probability incentivizes defaulting by Lemma 1.

Our next lemma establishes the key properties of the intrinsic propensity to default θ. The first part of the lemma states that agents are less inclined to default when returns are higher and when the repayment amount is lower ($\bar{w}$ is increasing in $\bar{b}$). This leads to the existence of three types of agents: *insolvent* agents with low returns who always default ($\theta(w, \bar{w}) > 1$) and act non-strategically; solvent high return agents who never default ($\theta(w, \bar{w}) \leq 0$) and also act non-strategically; and *strategic* agents with intermediate returns such that their default decision depends on their expectation of enforcement probability ($\theta(w, \bar{w}) \in (0, 1]$). The latter agents behave "strategically" in the sense that their incentives to default depend on their beliefs about the behavior of other agents in the economy (i.e., the aggregate default rate ψ). The mass of entrepreneurs who are strategic is $F(\bar{w}/(1 - \gamma\mu)) - F(\bar{w})$, and it determines the strength of strategic complementary in repayment decisions. It depends on the loan contract ($\bar{w}$), agent heterogeneity (F) and parameters determining the attractiveness of non-enforced liquidation ($\gamma\mu$). We denote the set of agent types who are strategic by $\mathcal{W} := (\bar{w}, \bar{w}/(1 - \gamma\mu))$.

Lemma 2. *Given loan contract* $(b, \bar{w})$ *with* $b, \bar{w} > 0$, $\theta(w, \bar{w})$ *is strictly decreasing in* w *and strictly increasing in* $\bar{w}$. *In addition,* $\theta(w, \bar{w}) \in (0, 1)$ *if and only if* $w \in \mathcal{W}$. *Agents with* $w < \bar{w}$ *always default (*$\theta(w, \bar{w}) > 1$*) and agents with* $w \geq \bar{w}/(1 - \gamma\mu)$ *never default (*$\theta(w, \bar{w}) < 0$*).*

2.2.3. Equilibrium enforcement under complete information

We begin our equilibrium analysis of the enforcement game by focusing on Nash equilibria under complete information about enforcement capacity X, which implies that the default rate ψ and enforcement probability $P(X, \psi)$ are common knowledge in equilibrium. We show that the game gives rise to multiple equilibria under fairly general conditions on the distribution of returns F.

We provide two different results regarding the multiplicity of equilibrium. The first one focuses on enforcement probability given by (1) and on a specific family of return distributions so that together they restrict the number of equilibria to at most three. Such a restriction will lead to

[12] Given the continuity of F, we can assume without loss that an indifferent agent always repays.

equilibrium default rates in the global game that exhibit a single discontinuity, rather than multiple jumps. Our second result lifts such restriction by identifying the necessary and sufficient condition for the existence of multiple equilibria for any admissible enforcement technology in the sense of Definition 1.

Assumption 1. The elasticity of F, $\epsilon_F(w) := \frac{wf(w)}{F(w)}$, is decreasing on $\mathcal{W}$ and $\lim_{w\downarrow 0}\epsilon_F(w) > 1$.

Intuitively, Assumption 1 requires that F be convex at zero and become gradually less elastic.[13] The family of distributions satisfying Assumption 1 includes commonly used distributions such as the log-normal or Pareto, although it does not hold for all distributions.[14]

The mechanism responsible for multiplicity of equilibria is standard. Since a higher default rate implies a lower enforcement probability, high default rate equilibria may become self-fulfilling because entrepreneurs are incentivized to default by the expectation of weak enforcement. The interesting aspect is how agent heterogeneity factors into this mechanism and which equilibria arise depending on the nature of agent heterogeneity, which we focus on next.

Nash equilibria under complete information can be summarized by a pivotal type w^* such that all agent types with $w \le w^*$ choose to default and those with $w > w^*$ choose to repay, since the intrinsic propensity to default $\theta(w, \bar{w})$ is decreasing in w (for any fixed $\bar{w}$) and the expected enforcement probability $E(P)$ is common across agents and equal to $P(X, F(w^*))$. Accordingly, any $w^* \ge \bar{w}$ is a Nash equilibrium if the pivotal type w^* is indifferent between defaulting and repaying, which by Lemma 1 implies

$$P(X, F(w^*)) = \theta(w^*, \bar{w}). \tag{7}$$

In particular, $w^* = \bar{w}$ is a Nash equilibrium if

$$P(X, F(\bar{w})) = \theta(\bar{w}, \bar{w}) = 1, \tag{8}$$

since agents with $w < \bar{w}$ are insolvent and always default.

The Nash equilibrium with pivotal type $w^* = \bar{w}$ is of particular interest. This is because this equilibrium is *efficient* in the sense of maximizing expected output, given that it minimizes the deadweight loss from liquidation. Accordingly, both the planner and lenders would like this equilibrium to be selected at the enforcement stage, since their respective problems involve maximizing entrepreneurs' payoffs. This is also the kind of equilibrium that would arise had enforcement not been fixed ex post, as is the case in Gale and Hellwig (1985). In contrast, equilibria with $w^* > \bar{w}$ are *inefficient* because also some solvent agents default.

Since we consider shocks to enforcement capacity, and X corresponds to the after-shock value of capacity, it is convenient to recast the above equilibrium condition by defining the *smallest* enforcement capacity $\chi(w, \bar{w})$ such that an agent of type w is indifferent between defaulting and repaying when she is a pivotal type; i.e., if she assumes all agents with returns higher than w repay and all agents with returns lower than w default. Using this notion, $w^* \in \mathcal{W}$ is an

13 The first condition of Assumption 1 holds when the relative (percentage) increase in the fraction of returns that are lower than some cutoff value is diminishing in that cutoff value. This condition is naturally satisfied at w where the density f is decreasing or constant. For single-peaked densities f on $[0, \infty)$, such as the Pareto or the log-normal distribution, the assumption places a restriction on the rate at which f can increase so that the distribution function is not too convex.

14 E.g., the exponential distribution violates the second condition.

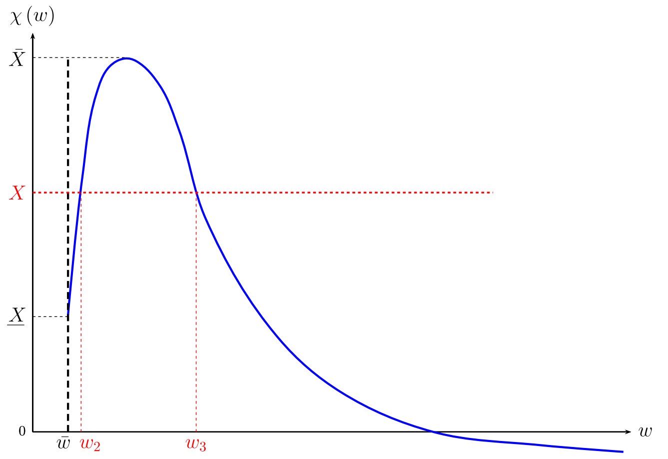

Fig. 2. Multiplicity of Equilibria. Notes: The figure plots $\chi(w, \bar{w}) = F(w)\theta(w, \bar{w})$ and illustrates all Nash equilibria that arise under complete information for $\chi(w, \bar{w})$ strictly quasi-concave and increasing at $\bar{w}$. There is an efficient equilibrium at $\bar{w}$ and two inefficient equilibria w_2 and w_3. Their respective default rates are $F(\bar{w})$, $F(w_2)$ and $F(w_3)$.

equilibrium if $X = \chi(w^*, \bar{w})$ and $w^* = \bar{w}$ is an equilibrium if $X \geq \chi(w^*, \bar{w})$, since agents with $w < \bar{w}$ always default.

Definition 2. Let $\chi(w, \bar{w})$ denote the smallest enforcement capacity X satisfying $P(X, F(w)) = \theta(w, \bar{w})$ for all $w \in \mathcal{W}$.

Lemma 3. *$\chi(w, \bar{w})$ is a real-valued continuous function on $\mathcal{W}$ for any admissible P and any fixed $\bar{w}$. Moreover, if P is given by (1) then $\chi(w, \bar{w}) = F(w)\theta(w, \bar{w})$, and, if F additionally satisfies Assumption 1, then $\chi(w, \bar{w})$ is strictly quasi-concave on $\mathcal{W}$.*

As Lemma 3 shows, if P is given by (1) and F satisfies Assumption 1, $\chi(w, \bar{w})$ is given by $F(w)\theta(w, \bar{w})$, which is a continuous single-peaked function, as illustrated in Fig. 2. Accordingly, if $\bar{w}$ falls to the left of the peak of required capacity $\chi(\cdot, \bar{w})$, which implies that $\chi(\cdot, \bar{w})$ is increasing at $\bar{w}$, there are exactly three equilibria for intermediate values of enforcement capacity between $\chi(\bar{w}, \bar{w})$ and the peak of $\chi(w, \bar{w})$, as illustrated in the figure. The first equilibrium is efficient and occurs at $w^* = \bar{w}$, while the other two equilibria, w_2 and w_3, exhibit higher default rates.[15] Proposition 1 summarizes this result formally.

Proposition 1. *If P is admissible and $\chi(w, \bar{w})$ is increasing $w = \bar{w}$ and strictly quasi-concave, then the complete information enforcement game has: 1) a unique equilibrium at $w^* = \bar{w}$ for all $X > \overline{X} := \max_{w \in \mathcal{W}} \chi(w, \bar{w})$; 2) a unique equilibrium at some $w^* > \bar{w}$ for all $X < \underline{X} := \chi(\bar{w}, \bar{w})$; 3) three equilibria w_i^* satisfying $\bar{w} = w_1^* < w_2^* < w_3^*$ for all X in the non-degenerate set $(\underline{X}, \overline{X})$[16]; and 4) two equilibria at $w_1^* = \bar{w}$ and at some $w_2^* > \bar{w}$ when $X = \overline{X}$ or $X = \underline{X}$.*

[15] It is worth noting that the equilibrium with pivotal type w_2 is unstable in the sense that a small deviation toward $\bar{w}$ (resp. w_3) would trigger all agents in $(\bar{w}, w_2)$ to repay (resp. all agents in (w_2, w_3) to default).

[16] By non-degenerate set we mean a set of positive Lebesgue measure.

Multiplicity of equilibria plagues the enforcement game precisely when it matters. Note that the planner would like to sustain the efficient equilibrium at the lowest cost possible, which implies that in practice enforcement capacity is likely to be close to the multiplicity region $[\underline{X}, \overline{X}]$, especially when the probability of capacity being depleted by the shock s is low. This is because, on the one hand, uniquely sustaining the efficient equilibrium requires $X > \overline{X}$, but on the other hand any excess above $\overline{X}$ that would not be utilized in normal times ($s = 0$) is costly.

We conclude this section by establishing the necessary and sufficient condition for equilibrium multiplicity under any admissible enforcement technology. This result shows that indeterminacy is a prevalent feature of the model and obtains whenever $\chi(w, \bar{w})$ is increasing around some $w \in \mathcal{W}$. $\chi(w, \bar{w})$ being increasing at w requires that F exhibits some concentration of mass around w. Intuitively, an increase in w has two effects on indifference condition $P(X, F(w)) = \theta(w, \bar{w})$: a drop in P due to a higher default rate and a drop in θ. Hence, the required X would go up when the first effect dominates, i.e., when $F(w)$ grows fast enough with w. This can be the case, for example, when the density function f is high at w.

Proposition 2. *If P is admissible then the complete information enforcement game has multiple equilibria for a non-degenerate set of values of X if and only if $\chi(w, \bar{w})$ is strictly increasing at some $w \in \mathcal{W}$.*

2.2.4. Global games formulation of the enforcement game

Equilibrium indeterminacy motivates us to relax the assumption of common knowledge of enforcement capacity X by considering an economy in which agents instead receive a noisy signal about X, which implies that neither ψ nor $P(X, \psi)$ are common knowledge. From an economic point of view the global games formulation of the enforcement game is of interest because indeterminacy of equilibrium under complete information relies on the strong assumption that entrepreneurs can perfectly observe X, and thus coordinate their beliefs about ψ and P. In practice, borrowers may have a hard time knowing enforcement fundamentals with exact precision or, more generally, they may not be able to perfectly infer the equilibrium default rate or the underlying enforcement probability.

Formally, we follow Frankel et al. (2003) and assume that each agent receives a signal $x = X + \nu\eta$, where $\nu > 0$ is a scaling factor and η is an i.i.d. random variable with continuous distribution H that has full support on $[-1/2, 1/2]$. That is, agents' signals are correlated through X but their signal noise is i.i.d. To simplify the proofs, we assume agents' prior about X is uniformly distributed on [0, 1], albeit our results do not hinge on this assumption. As Frankel et al. (2003) show, equilibrium selection arguments work in the limit as signal error goes to zero, since any well-behaved prior will be approximately uniform over the small range of X that are possible given an agent's signal.

Importantly, we assume that the signal summarizes all available information to entrepreneurs. In particular, we do not allow entrepreneurs to learn new information from observing loan terms $(b, \bar{w})$, even though they are contingent on enforcement capacity X. As we further discuss in Section 4, this assumption is not essential for the results but it significantly simplifies the proofs. In that section, we outline a generalization of the setup in which lenders do not directly observe X, as they do here, but instead receive a public signal at the lending stage that is noisier than the private signals borrowers receive when loans are due. Such a generalization is realistic, since the available enforcement capacity evolves over time, that is, from the time the lenders issue loans to the time repayment is due. As we argue in Section 4, the presence of a public signal at the time of obtaining the loans merely changes the initial prior about X but does not undermine the selection mechanism or change our characterization of the resulting equilibrium. As already mentioned, Frankel et al. (2003) show that the selection does not hinge on the choice of the prior.

On a technical note, while our results from now on are stated for a continuous distribution of returns for ease of exposition, our proof technique relies on the use of discrete return distributions. Specifically, we characterize the equilibrium associated to a sequence of discrete distributions that in the limit converge to the continuous distribution F.[17]

2.2.5. Equilibrium enforcement in the global game

Our first result shows that the game has a unique equilibrium as signal noise vanishes in the limit. In this equilibrium agents' strategies are characterized by a type-specific cutoff $k(w, \bar{w})$ with respect to signal x such that an agent with w who receives a signal above the cutoff repays and otherwise defaults.

Proposition 3. *For any admissible P, the enforcement game has a unique equilibrium as $\nu \to 0$. Equilibrium strategies are characterized by a signal cutoff $k(w, \bar{w})$ such that*

- *If $x \geq k(w, \bar{w})$, agents choose to repay ($a = 1$)*
- *If $x < k(w, \bar{w})$, agents choose to default ($a = 0$), where $k(w, \bar{w})$ satisfies the system of limit indifference conditions*

$$\lim_{\nu \to 0} \mathbb{E}(P|x = k(w, \bar{w})) = \theta(w, \bar{w}) \text{ for all } w \in \mathcal{W}. \tag{9}$$

The intuition for why equilibrium is unique follows from standard arguments in the global games literature. Even under infinitesimal uncertainty about enforcement capacity X, the lack of common knowledge exposes agents to large strategic uncertainty about the actions of others, hindering their ability to coordinate on a high default equilibria when X is low based solely on a self-fulfilling belief. In other words, entrepreneurs use signals to learn about X rather than to coordinate their choices.[18,19]

Our next results provide a complete characterization of the unique equilibrium cutoff $k(w, \bar{w})$ underlying Proposition 3. To ease notation, we now omit the explicit dependence of $k(w, \bar{w})$, $\chi(w, \bar{w})$ and $\theta(w, \bar{w})$ on $\bar{w}$.

We begin with the most general characterization that applies to any admissible P in the sense of Definition 1. Albeit technical in nature, this result shows that $k(w)$ generally "irons" out any peaks of the function $\chi(w)$, resulting in clustering of types on the same equilibrium strategy around its peaks.

[17] A discrete distribution is technically convenient to prove uniqueness in the global game version of the model.

[18] Intuitively, common knowledge no longer applies because, when an agent observes her idiosyncratic signal x, she considers it possible that the true value X is $\nu/2$ away from the signal x. As a result, she considers it possible that other agents may get a signal as far as ν away from her own x, thus admitting a possibility that other agents' signals are as far as 2ν away from x and so on and so forth.

[19] Our formal proof uses the fact that games with strategic complementarities feature a smallest and largest Nash equilibrium, both in cutoff strategies (Van Zandt and Vives, 2007), and shows that both must coincide. Specifically, it shows that, given the noise structure, $E(P|x = k(w, \bar{w}))$ goes up when $k(\cdot)$ is raised for all w, implying that there can be only one profile of cutoffs at which $E(P|x = k(w, \bar{w})) = \theta(w, \bar{w})$ for all $\theta(w, \bar{w}) \in (0, 1)$, i.e., at which each strategic agent is indifferent between defaulting or not when she receives her cutoff signal.

Proposition 4. *If P is admissible, there exists a unique partition of types with $\theta(w) \in (0, 1]$ into intervals $\{(\underline{w}_j, \bar{w}_j]\}_{j=1}^J$ s.t.:*

(a) *if $k(w)$ is strictly decreasing in interval j then it is constant in intervals $j-1$ and $j+1$ and vice versa;*
(b) *if $\chi(w)$ is strictly decreasing in interval j then $k(w) = \chi(w)$ for all $w \in (\underline{w}_j, \bar{w}_j]$;*
(c) *if $\chi(w)$ is not strictly decreasing in $(\underline{w}_j, \bar{w}_j]$ then $k(w) = k_j$ for all $w \in (\underline{w}_j, \bar{w}_j]$ with k_j satisfying $k_j = \chi(\bar{w}_j) \geq \chi(\underline{w}_j)$ (with equality if $\underline{w}_j > \bar{w}$) and*

$$\int_{F(\underline{w}_j)}^{F(\bar{w}_j)} P(k_j, z)dz = \int_{\underline{w}_j}^{\bar{w}_j} \theta(w)dF(w). \tag{10}$$

To illustrate the implications of this clustering, we next consider $\chi(w)$ that is increasing at $w = \bar{w}$ and single-peaked, as is the case when P is given by (1) and Assumption 1 is satisfied (see Lemma 3). Recall that under these assumptions the game of complete information exhibits multiple equilibria for $X \in [\underline{X}, \overline{X}]$, as illustrated in Fig. 2 and established in Proposition 1. In the context of Proposition 4 above, the single-peakedness of $\chi(w)$ leads to a single cluster of types using the same (limit) signal cutoff in equilibrium.

Proposition 5. *If P is admissible and $\chi(w)$ is increasing at $w = \bar{w}$ and strictly quasi-concave, then there exists $w^* > \bar{w}$ such that:*

$$k(w) = \begin{cases} \chi(w^*) & \text{for all } w \in [\bar{w}, w^*] \\ \chi(w) & \text{for all } w > w^*, \end{cases}$$

where w^* is the unique solution in $\mathcal{W}$ to

$$\int_{F(\bar{w})}^{F(w^*)} P(\chi(w^*), z)dz = \int_{\bar{w}}^{w^*} \theta(w)dF(w). \tag{11}$$

Corollary 1. *If, additionally, P is given by (1), then there exists $w^* > \bar{w}$ such that:*

$$k(w) = \begin{cases} \theta(w^*)F(w^*) & \text{for all } w \in [\bar{w}, w^*] \\ \theta(w)F(w) & \text{for all } w > w^*, \end{cases}$$

where $w^* = \bar{w}$ if $\bar{w} \geq w_{\max} := \text{argmax}_{w \in \mathcal{W}} \chi(w)$ and, if $\bar{w} < w_{\max}$, w^* is the unique solution in $(w_{\max}, \infty)$ to

$$\theta(w^*)F(w^*)\big(1 - \log\theta(w^*)\big) - F(\bar{w}) = \int_{\bar{w}}^{w^*} \theta(w)dF(w). \tag{12}$$

Proposition 5 and Corollary 1 show that under these assumptions the equilibrium cutoff $k(w)$ is as illustrated in Fig. 3. That is, $k(w) = k(w^*)$ is constant up to some $w^* > \bar{w}$, which is the solution to (12), while $k(w)$ tracks $\chi(w)$ for $w > w^*$. In particular, the flat region of $k(w)$ – over

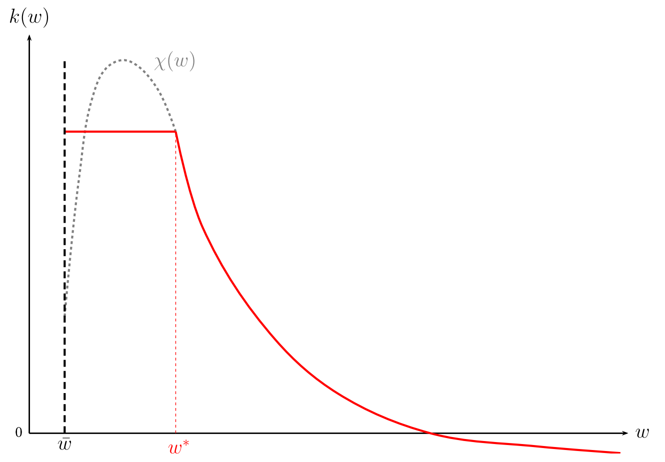

Fig. 3. Equilibrium Default Cutoffs. Notes: The figure plots strategy profile $k(w)$ as implied by Corollary 1 for single-peaked $\chi(w)$ (red solid line). $k(w)$ features a flat region for the cluster of returns $[\bar{w}, w^*]$. When X is above $k(w^*)$ the efficient equilibrium with default rate $\psi = F(\bar{w})$ is selected. The strategy profile follows $\chi(w)$ for $w > w^*$. The presence of the cluster leads to *fragility* of the efficient equilibrium, as for any X lower than $k(w^*)$ all entrepreneur types between $\bar{w}$ and w^* default, resulting in a discontinuous jump of the default rate from $F(\bar{w})$ to $F(w^*)$.

which different types of agents choose the same strategy $k(w^*)$ – occurs between the lowest-default equilibrium that maximizes output and the highest-default equilibria among the multiple equilibria under complete information. Accordingly, the efficient equilibrium in which only insolvent agents default is *fragile* in the following sense: If X is depleted below $k(w^*)$, a cluster of agents with returns between $\bar{w}$ and w^* defaults. This implies that, if there is enough enforcement capacity in the economy, the default rate is at the lowest possible level. However, if enforcement capacity is depleted below the level associated with the flat region, the default rate discontinuously jumps by the measure of agents with returns in the cluster.

To illustrate the key intuition behind this result, consider a simplified numerical example with only three types of agents $i = 1, 2, 3$ of mass $1/3$ each and default propensities θ_i satisfying $\theta_1 > 1$ and $1/2 > \theta_2 > \theta_3 > 0$. That is, the first type always defaults, and types 2 and 3 default if they expect enforcement probability to be lower than θ_2 and $\theta_3 < \theta_2$, respectively. P is given by (1). To ease notation, k_i denotes the signal cutoffs of each type $i = 1, 2, 3$.[20]

To see why clustering of types 2 and 3 forms in this case, observe first that k_i must be (weakly) decreasing in type, since propensity to default θ_i is decreasing in type. Hence, the following two cases are possible: $k_2 > k_3$ or $k_2 = k_3$. Moreover, by Proposition 3 and the fact that $x = X$ in the limit, strategy cutoffs k_2 and k_3 must satisfy:

$$\lim_{\nu \to 0} \mathbb{E}(P|x = k_i) = \lim_{\nu \to 0} \mathbb{E}(\min\{1, X/\psi\}|x = k_i)$$
$$= \mathbb{E}(\min\{1, k_i/\psi\}|x = k_i) = \theta_i, \quad i = 2, 3. \tag{13}$$

[20] Note that indeterminacy obtains in this example under complete information. For example, if $X = 1/6$, there exists an equilibrium in which only agents of type-1 default. In this equilibrium $P = (1/6)/(1/3) = 1/2$, and hence $P \geq \theta_2 > \theta_3$. However, for $\theta_2 > 1/4 > \theta_3$, there is also an equilibrium in which agents of type-2 also default, since $P = 1/4$ in such a case. Analogously, for $\theta_3 > 1/6$, there is an equilibrium in which all agents default, since $P = 1/6$.

That is, an agent of type 2 or 3 who receives a signal *equal* to signal cutoff k_i must be *indifferent* between defaulting and repaying. We now show how to solve for k_2 and k_3 and that it is possible to have $k_2 = k_3$ despite $\theta_2 > \theta_3$.

To that end, consider what happens when $k_2 > k_3$. In such a case, when signal noise is sufficiently small, we must have $k_2 - \nu > k_3$, since all signals are within ν of each other by the law of large numbers. This implies that a type-2 agent knows that, when her signal x is equal to her strategy cutoff k_2, all type-3 agents are receiving signals above k_3, and hence they all repay. Consequently, the only unknown from her perspective is the default rate among her own type, since she does not know how many agents of type-2 might be receiving signals below k_2. Let ψ_2 be this random variable.

Determining the distribution of ψ_2 is key to solving the indifference condition (13). In this case the distribution is uniform, which is a standard result known as the *Laplacian property* (Morris and Shin, 2003). To see why, note that, since all type-2 agents use the same cutoff k_2, when a type 2 receives signal $x = k_2$, she understands that the rank of her signal among signals of type-2 agents determines the default rate among type-2 agents. For example, a median signal would imply default rate $\psi = 1/2$, a first quartile signal would imply default rate $\psi = 1/4$, and so on and so forth. The key observation here is that the rank of *any* random variable is uniform, and hence the default rate ψ_2 is also uniform. Accordingly, in the limit we have

$$\mathbb{E}(\min\{1, k_2/\psi\}|x = k_2) = \int_0^1 \frac{k_2}{\frac{1}{3} + \frac{1}{3}\psi_2} d\psi_2 = 3k_2 \log(1/3 + \psi_2/3)\Big|_0^1 = \theta_2,$$

and hence $k_2 = \dfrac{\theta_2}{3\log 2} \approx 0.48\theta_2$.[21]

We can analogously calculate k_3 by using the fact that a type-3 agent receiving $x = k_3$ must believe that the default rate is $\psi = 2/3 + (1/3)\psi_3$, where ψ_3 pertains to the default rate among her own type, which type-3 deems uniformly distributed. Note that the belief about ψ of a type-3 agent with $x = k_3$ entails a distribution that is shifted upward relative to type-2 agent receiving $x = k_2$, which was $1/3 + (1/3)\psi_2$. This will become important for the emergence of a cluster, as we shall see. Solving the indifference condition, we obtain $k_3 = \dfrac{\theta_3}{3\log(3/2)} \approx 0.82\theta_3$.

When θ_2 and θ_3 are not too far apart, it is clear that $k_2 > k_3$ is not possible. Accordingly, the two types must cluster on the same strategy profile despite having different intrinsic propensities to default, which implies $k_2 = k_3$. The above disparity in beliefs about ψ is the key driver behind clustering: a type-3 agent receiving $x = k_3$ expects a *higher* default rate and hence a lower enforcement probability than a type-2 agent, which incentivizes a type-3 agent receiving $x = k_3$ to default.

Proposition 4 generalizes this result and provides a formula to solve for the common limit threshold $k^* = k_2 = k_3$. To derive this result, we build on the key property of beliefs in global games with heterogeneous agents established by Sákovics and Steiner (2012) and known as the *belief constraint*. They show that, even though the Laplacian property no longer holds when signal cutoffs of different types are within ν of each other, it holds *on average* in the following

[21] The min operator in P does not bind in this example, since $k_2 < 1/3$ when $\theta_2 < 0.5$ and $\psi = 1/3 + \psi_2/3 \geq 1/3$ given that all type-1 agents default.

sense:[22] If we randomly draw an agent of type 2 or 3, each receiving her respective cutoff signal, the average belief about the distribution of the *joint* default rate of both types, denoted by ψ_{23}, will still be uniform. This is sufficient to solve for the limit equilibrium cutoff because $k_2 = k_3$ as signal noise vanishes in the limit, and hence we do not need to make a distinction between individual beliefs across the two types. Specifically, instead of solving (13), we can use the average indifference condition of types 2 and 3 and substitute out the average belief of the two types by the uniform distribution as we integrate over ψ_{23}:

$$\frac{1}{2}\left(\mathbb{E}(\min\{1, k^*/\psi\}|x = k_2) + \mathbb{E}(\min\{1, k^*/\psi\}|x = k_3)\right) = \int_0^1 \frac{k^*}{\frac{1}{3} + \frac{2}{3}\psi_{23}} d\psi_{23}$$
$$= \frac{1}{2}(\theta_2 + \theta_3).$$

Applying this approach to our example, we obtain that the common limit threshold for both types 2 and 3 is $k^* = \dfrac{\theta_2 + \theta_3}{3 \log 3}$.[23]

Importantly, the two types have different signal thresholds away from the limit, which leads to differences in beliefs about P to offset differences in propensities to default θ. It is only in the limit that the two cutoffs coincide. The key observation behind the belief constraint, which is instrumental to derive the above result, is that when signal noise is small, the probability that agents of type-2 receiving $x = k_2$ assign to the event that signals generated for type-3 agents fall in the interval between the two cutoffs exactly offsets the probability that agents of type-3 receiving $x = k_3$ assign to the event that signals of type-2 agents fall into the same interval.

2.2.6. Policy implications of equilibrium clustering

Clustering turns out relevant from a policy perspective as it is driven by an externality, which we refer to as the *enforcement externality*. This can be readily gleaned from our numerical example. Note that if there is a cluster of agents who may default, the most effective policy may be to prevent type-2 agents from defaulting by reducing their repayment amount, as this automatically prevents agents of type-3 from defaulting. As long as θ_2 and θ_3 are not too far apart, which is precisely when the cluster forms, it will be better than preventing type-3 from defaulting. A recent example of a policy in this spirit is the 2009 Home Affordable Modification Program (HAMP), which targeted homeowners in the U.S. who could prove that they were underwater and had not yet defaulted.[24]

2.3. Provision of credit

Having characterized equilibrium in the enforcement game, we next turn to the analysis of the model implications for the provision of credit. To this end, we lay out the lender's problem

[22] Although we rely on Sákovics and Steiner (2012), our characterization technique in the proof goes beyond their specific framework and it determines when an endogenous cluster forms. While they restrict agent preferences to guarantee the emergence of a single limit cutoff for all agents, we characterize the endogenous emergence of decision clusters as a function of agent heterogeneity, generalizing the applicability of the belief constraint to games with asymmetric equilibria.

[23] The min operator in the average indifference condition does not bind, since $\theta_2 + \theta_3 < 1$ and thus $k^* < 1/3$.

[24] Sákovics and Steiner (2012) derive a related result in the context of their model.

and analyze the equilibrium of the lending market, after X_o has been chosen and the shock s has been realized. We focus on the case of a single default cluster to ease the exposition.

2.3.1. Lender problem

Lenders are risk neutral and rationally anticipate the outcome of the enforcement game. They deal with a mass of entrepreneurs and must break even in expectation by the law of large numbers. They compete in a Bertrand fashion. For convenience, we normalize their cost of funds to zero.

Bertrand competition implies that lenders offer an equilibrium contract $(b, \bar{w})$ that maximizes entrepreneurs' expected payoffs subject to a non-negative profit condition and to the repayment behavior determined by capacity X and thresholds $k(w)$.[25] Profits from a loan $(b, \bar{w})$, given project return w, are

$$\pi(w, m, b, \bar{w}) := \begin{cases} (y+b)\bar{w} - b & k(w) \leq X \\ \big(m + (1-m)(1-\gamma)\big)\mu(y+b)w - b & k(w) > X. \end{cases} \tag{14}$$

Accordingly, the lender problem is given by:

$$\max_{b,\bar{w}} \left[\int_{\{w:k(w)\leq X\}} (y+b)(w-\bar{w})dF + (1-P)\gamma \int_{\{w:k(w)>X\}} \mu(y+b)wdF \right], \tag{15}$$

subject to

$$b \leq \int_{\{w:k(w)\leq X\}} (y+b)\bar{w}dF + (P + (1-P)(1-\gamma)) \int_{\{w:k(w)>X\}} \mu(y+b)wdF, \tag{16}$$

and the enforcement capacity constraint $P = P(X, \psi)$.

To ensure an interior solution, we impose an assumption that the dead-weight loss from defaulting $(1-\mu)$ is high enough so that arbitrarily large loans cannot be recouped through liquidation.

Assumption 2. $\int_{\bar{w}}^{\infty} \bar{w}dF(w) + \mu \int_0^{\bar{w}} wdF(w) < 1$ for all $\bar{w} \geq 0$.[26]

This assumption is needed because production technology is linear in our model and hence it may pay to scale up projects ad infinitum and simply liquidate all of them ex post.

2.3.2. Equilibrium provision of credit

Proposition 6 states the main result, which applies to an initial situation when lenders choose loan size b such that the default rate is minimal, i.e., equal to $\psi = F(\bar{w})$. It shows that, if enforcement capacity X falls to some level $X' < k(w^*)$, there are only two possibilities: the default rate

[25] Payoff maximization subject to zero profits follows from the usual arguments. Since agents are ex-ante identical, there is no profitable deviating contract for a lender: either such contract is less attractive to consumers or yields negative profits. In addition, if the equilibrium contract does not maximize agent payoffs, a lender can make positive profits by offering a contract arbitrarily close to the payoff-maximizing one and attract all agents.

[26] The assumption guarantees that, as b goes to infinity and hence $b/(y+b)$ goes to one, there is no repayment cutoff $\bar{w}$ that would make lenders earn non-negative profits, even when $P = 1$.

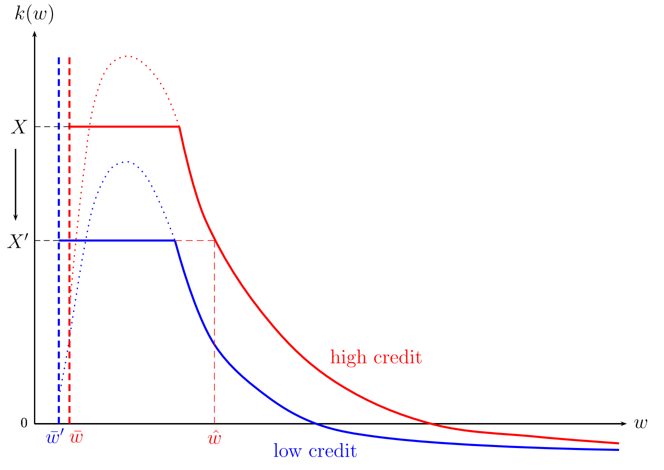

Fig. 4. Equilibrium Strategy in the Global Game. Notes: The figure plots equilibrium cutoff $k(w)$ as implied by Corollary 1 for quasi-concave $\chi(w)$ and two different values of b: high value (red line, labeled high credit) and a low value (blue line, labeled low credit). The plot illustrates how a drop in enforcement capacity from X to X' leads to a contraction of credit.

jumps or b contracts. The immediate corollary of this result is that the price schedule for loans as a function of their size is discontinuous. While larger loans may be offered, they are offered at a disproportionately higher price.

Lemma 4. *$k(w^*)$ in Proposition 5 is strictly increasing with respect to $\bar{w}$.*

Proposition 6. *Let $(b_X, \bar{w}_X)$ be the equilibrium loan size associated to X under the assumptions in Proposition 5 and let ψ_X be the corresponding equilibrium default rate. If $X = k(w^*)$ the equilibrium contract $(b_{X'}, \bar{w}_{X'})$ and the default rate $\psi_{X'}$ associated to any $X' < X$ satisfy either $b_{X'} < b_X$ or $\psi_{X'} > F(\bar{w}_{X'})$.*

The intuition behind the above result is that larger loan sizes b are necessarily associated with a higher equilibrium repayment cutoff $\bar{w}$. This follows from the zero profit condition. A higher repayment cutoff, in turn, increases default propensities of all agents by equation (5), and raises cluster's strategy threshold $k(w^*) = \chi(w^*)$ (as shown in Lemma 4).

Fig. 4 provides the impulse-response analysis of a shock s that reduces enforcement capacity. Initially, enforcement capacity X is high and sustains the lowest default rate $\psi = F(\bar{w})$ for some value of b, which determines the initial position of the strategy profile $k(w)$, which corresponds to the line labeled "high credit." After a shock lowers enforcement capacity to some X', this equilibrium is no longer optimal because agents with returns between $\bar{w}$ and $\hat{w}$ would choose to default – as implied by the original position of $k(w)$ ("high credit"). As a result, only two scenarios are possible. The first scenario involves a jump in the default rate to $F(\hat{w})$, in effect raising the cost of credit and repayment cutoff $\bar{w}$, and a downward shift of the "high credit" line resulting in even more defaults. In the second scenario b falls so as to prevent the jump in the default rate by shifting down the strategy profile $k(w)$ to the position labeled as "low credit", which sustains the low default rate $F(\bar{w}')$. As we show in the next section, since defaults and liquidations are socially wasteful, preventing default is generally optimal for reasonable parameterizations of the model.[27] (In the Online Appendix, we provide a numerical example illustrating how credit reacts to a depletion of enforcement based on the previous three-type setup.)

2.4. Accumulation of enforcement capacity

We close the model by assuming that before the lending market opens a benevolent planner chooses the enforcement capacity to maximize the expected entrepreneur payoffs net of capacity costs $c(X_o)$:

$$\max_{X_o}[E(V(X_o - s)) - c(X_o)], \tag{17}$$

where $V(X)$ denotes the expected payoff of agents from the equilibrium contract associated with X, i.e., the loan that solves (15), and the expectation is taken over the distribution of shocks s.

3. Quantitative analysis

Here we calibrate our model to study its quantitative implications. To this end, we consider a binary shock that hits with probability $p \in [0, 1]$ and consider two shock sizes. In the first case, we assume the shock is large, simulating a severe financial crisis. In the second case, we focus on a smaller shock, simulating regular recessions.

For demonstrational purposes, we set the size of the stylized *crisis* shock to $s = 0.018$, which corresponds to the difference in delinquency rates on commercial loans in the U.S. between the pre-crisis period 2000–07 (2.3% delinquency rate) and the crisis period 2008–09 (4.1%).[28] We set the stylized *recessionary* shock to 0.4% ($s = 0.004$), which is half the average absolute deviation of the default rate on commercial loans between 2000 and 2007. Note that, when X corresponds to the mass of liquidations that can be enforced, the size of the hike in default rate on existing loans coincides with the magnitude of the decline in enforcement resources from the perspective of the lending market.[29] We are motivated by this basic logic in considering a direct shock to capacity and using the spike in default rates on loans in the data as a measure of the size of the shock to enforcement.

We choose the remaining parameters to ensure that, when the probability of the shock is nil ($p = 0$), the model's static implications are consistent with the U.S. data.

3.1. Parameterization

We independently select the value of average returns Ew, liquidation value μ, and the shape of return distribution F following the quantitative business-cycle literature that focuses on enforce-

[27] If X is very low it is possible that the loan size b at which $X = \chi(w^*)$ is so small that the cluster contains only a tiny mass of entrepreneurs. In such a case, it may be optimal to increase b and let the cluster default.

[28] Source: Federal Reserve Bank of St. Louis.

[29] For example, if each default proceeding takes 100 hours of a judge's time per year for a period of n years, and there is a spike of 1000 new default cases in the economy from the usual level of 3000 cases, this event subtracts 100,000 judge hours from the available court resources per year for the next n years. Hence, if initial enforcement capacity was, say, 1,000,000 judge hours per year, the residual capacity from the perspective of new lending was 700,000 judge hours per year before the shock and 500,000 after the shock.

Table 1
Parameters and Statistics.

Parameter	Value	Source
Independently selected parameters		
y	1	
Ew	1.02	Bernanke et al. (1999)
μ	0.88	Bernanke et al. (1999)
F	Lognormal	Bernanke et al. (1999), Christiano et al. (2014)
Jointly selected parameters		
σ	$3/8$	
$c(X)$	$0.088X$	
γ	0.33	

Statistic	Model	Data
Targeted moments		
Debt-to-equity ($\frac{b}{y}$)	0.8	0.8[a]
Default rate (ψ)	2.2%	2.3%[b]
Enforcement cost/project value	12.7%	12.5%[c]
Other relevant moments		
Capacity costs/debt ($c(X_o)/b$)	0.3%	
Capacity utilization ($\frac{\psi}{X}$)	83.4%	
Cluster size ($F(w^*) - F(\bar{w})$)	4.5%	
% strategic agents ($\theta(w) \in (0, 1)$)	11.2%	

[a] Gomes et al. (2016) report the average leverage ratio $\frac{b}{k}$ in the U.S. between 1971 and 2012 of 0.42, which corresponds in our model to a debt-to-equity ratio of 0.84 (if $k = y + b$). The debt-to-equity ratio is 0.52 in Christiano et al. (2014) and 1 in Bernanke et al. (1999).
[b] 2000–2007 average delinquency rate on commercial and industrial loans. Source: Federal Reserve Bank of St. Louis.
[c] Average bankruptcy costs, relative to pre-bankruptcy firm value, of chapter 7 (8.1%) and chapter 11 (16.9%) bankruptcies in the U.S. (Bris et al., 2006).

ment frictions (we also normalize equity y to one). In particular, the values of these parameters are drawn from Bernanke et al. (1999) and Christiano et al. (2014), two of the workhorse models of the business cycle literature on financial frictions. We then jointly choose the volatility of returns σ, linear capacity costs $c(X_o) = c_o X_o$, and private benefit from lack of enforcement γ to match the average debt-to-equity ratio in the U.S., the average delinquency rate on commercial loans, and the average bankruptcy costs relative to pre-bankruptcy firm value in the U.S. Table 1 lists the value of the parameters and the corresponding targets in the data and their values in the calibrated model.

As Table 1 shows, the calibrated model not only matches data targets but has sensible predictions in several respects that we did not target. For example, capacity buildup costs relative to debt are 0.3 percent. The cluster that arises in the calibrated model includes only 4.5% of the agents, and the pool of strategic agents includes 11.2% of the population. This is consistent with the notion that most agents in the economy are nonstrategic and hence not sensitive to enforcement probability.

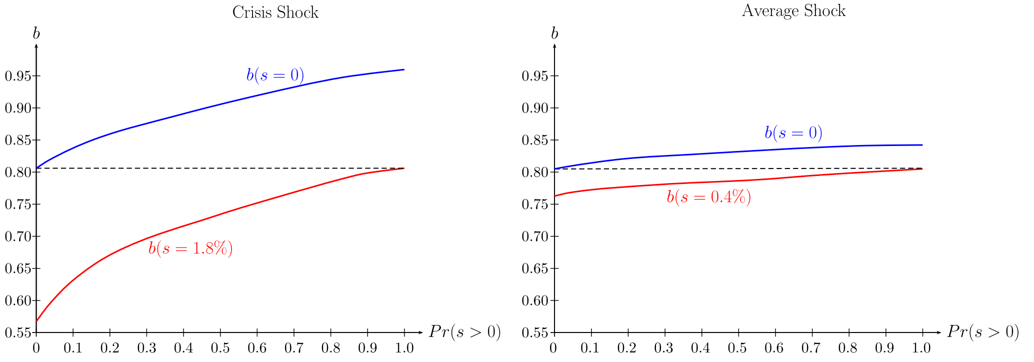

Fig. 5. Enforcement Capacity and Shock Propagation. Notes: The figure shows the impact of the shock on equilibrium loan size b. It plots loan levels for each realization of the shock s (two levels) as a function of shock probability p.

3.2. *Quantitative findings*

3.2.1. *Static implications*

Our model implies that in the steady state the utilization rate of enforcement capacity is 83 percent; that is, the steady state default rate $\psi = F(\bar{w})$ uses up 83 percent of enforcement resources. Yet, enforcement capacity is binding in the sense that any shock that depletes enforcement resources leads to a contraction of credit. This is because $X = k(w^*)$, implying that credit supply is the highest compatible with the low-default equilibrium for the given level of enforcement resources. Furthermore, we note that credit supply would be 20 percent higher had enforcement been freely available in this economy.

3.2.2. *Dynamic implications*

Fig. 5 illustrates the impact of a sudden depletion of enforcement resources on the equilibrium loan size b for each type of the shock and as a function of its ex ante probability p (rationally anticipated by the planner). When p is very low, the crisis shock implies a contraction of credit by as much as 30%. Under the average shock, credit goes down by 5%. Importantly, in neither of the two cases is it optimal to let the cluster default and, hence, the enforcement constraint $X = \chi(w^*)$ is always binding. This result underscores the strength of the endogenous link between enforcement and credit supply predicted by our model.

Interestingly, although the gap $b(s = 0) - b(s > 0)$ slightly shrinks as the shock becomes more likely, it is still substantial for any p. The key reason is that, while higher shock probabilities induce the planner to accumulate precautionary capacity X_o to dampen the effect of the shock on credit provision, this extra capacity also leads lenders to extend more credit in normal times by relaxing the borrowing constraint imposed by $X = \chi(w^*)$. Accordingly, our model is capable of generating large credit fluctuations in response to fluctuations in enforcement resources.

In Appendix A, we further discuss our choice of some key parameters of the model, such as the variance of returns (σ) and the rents from lack of enforcement (γ), and relate them to existing micro-level and cross-country evidence. There we show, for example, that substantial heterogeneity is needed to match observed default rates and to obtain meaningful enforcement utilization rates. We also discuss evidence that suggests our choice of the value of the critical parameter γ is reasonable in light of the data. In addition, we illustrate in the Online Appendix the role of strategic complementarities in the provision of credit by comparing the credit supply under the assumption that the efficient equilibrium is always played to the amount of credit under the assumption that the highest default-equilibrium is always selected.

4. Robustness to information aggregation

We conclude the analysis by discussing how to extend the baseline setup to prevent borrowers from perfectly inferring X from loan terms, which we ruled out by assumption.

The possibility of multiple equilibria in the presence of *endogenous public signals* in a global games setup, such as prices ($b, \bar{b}$ in our case), has been originally raised by Angeletos and Werning (2006) and Angeletos et al. (2006) as a criticism of the global games methodology when applied to environments in which markets aggregate information through publicly observed prices. The basic idea is that market prices generate a public signal that is naturally more precise than the idiosyncratic private signals they aggregate, serving as a coordination device that brings back equilibrium multiplicity.

This criticism does not readily apply to our analysis for the following reason: In credit markets, lenders extend credit *before* borrowers decide to default and hence their information, however precise it may be at that time, becomes outdated when the loans are due and agents decide whether to default or repay. This delay makes the *earlier* public signal potentially *less* precise than *later* private signals, even if the public signal aggregates some *earlier* private signals. According to the result established by Angeletos and Werning (2006) and Angeletos et al. (2006), the presence of a *less* precise public signal is *not* enough to resuscitate indeterminacy.

To formalize this argument, we now sketch an extension of our model that allows agents to infer information about X along the lines above and argue that it boils down to our setup in the limit. We do not incorporate this extended setup in the paper because it would significantly complicate our proofs. This alternative specification is based on the idea that lenders do not perfectly observe X at the time the credit market opens but instead observe a public signal of its true value. However, between the time credit is extended and loans are due, entrepreneurs receive more precise information regarding enforcement resources.

Specifically, consider the following change to the baseline model. After the planner chooses X_o and the shock s is realized, lenders and entrepreneurs observe a public signal $\tilde{x}$ regarding enforcement capacity $X = X_o - s$. Assume that $\tilde{x}$ is distributed according to continuous cdf $\tilde{H}(\tilde{x}|X)$ with full support on $[0, 1]$ for all values of X. After observing this public signal, lenders issue loan contracts to entrepreneurs, who then invest these funds in their projects. Next, project returns are realized and agents receive, as in the benchmark model, idiosyncratic private signals $x = X + \nu\eta$ and simultaneously decide whether to repay or default.

The introduction of a public signal observed by lenders requires two modifications. First, entrepreneur beliefs about X conditional on the public signal $\tilde{x}$ before receiving their idiosyncratic signals are no longer uniform (recall that agents have uniform priors in the benchmark model). However, given that $\tilde{H}(\tilde{x}|X)$ is continuous and has full support, agents' conditional beliefs will also be continuous and have full support on $[0, 1]$. As Frankel et al. (2003) show, the global games selection works whenever the prior belief satisfies these conditions. Intuitively the reason is that, as $\nu \to 0$ any continuous prior with full support becomes roughly uniform in the interval $[X - \nu, X + \nu]$. Second, the entrepreneurs problem must be changed to incorporate the fact that lenders do not know X. Accordingly, lenders in this case solve

$$\max_{b,\bar{w}} E\left[\int_{\{w:k(w)\leq X\}} (y+b)(w-\bar{w})dF + (1-P)\gamma \int_{\{w:k(w)>X\}} \mu(y+b)wdF \middle| \tilde{x}\right], \quad (18)$$

subject to

$$b \leq E\left[\int_{\{w:k(w)\leq X\}} (y+b)\bar{w}dF + (P+(1-P)(1-\gamma)) \int_{\{w:k(w)>X\}} \mu(y+b)wdF \middle| \tilde{x}\right], \quad (19)$$

and the enforcement capacity constraint $P = P(X, \psi)$.

It is clear that, if $\tilde{x}$ is sufficiently precise, the solution to (18) will be responsive to changes in X, and the relationship between credit and enforcement capacity will closely approximate the one in the benchmark model where X is observed by lenders without noise.

5. Conclusions

We have analyzed capacity-constrained enforcement in a canonical model of debt-financed entrepreneurial activity. We have shown that shocks that deplete enforcement resources may lead to a spike in the aggregate default rate and trigger credit rationing. We have identified policies that can mitigate this outcome. We have also illustrated the quantitative implications of our model when calibrated to the data. Although stylized in nature, this analysis shows that the propagation mechanism we focus on is quantitatively relevant.

Appendix A. Discussion of key parameters and sensitivity analysis

Here we relate more closely the key parameters of the model to the data and discuss how they affect our results. For additional analysis see the Online Appendix.

Heterogeneity of returns (σ) Heterogeneity of returns affects equilibrium outcomes in our model by modulating the degree of strategic complementarity. As returns become more concentrated, so do agents' propensities to default, leading to a larger default cluster. This leads to unrealistically large default waves if capacity falls below the cluster threshold. For instance, when $X = 0.05$, the cluster size is 4.5% for $\sigma = \frac{3}{8}$, while it is 75% for $\sigma = \frac{1}{16}$.

It is worth mentioning that our selection of $\sigma = 3/8$ represents the midpoint between the value of $\sigma = 1/4$ in Christiano et al. (2014) and $\sigma = 1/2$ in Bernanke et al. (1999).

Heterogeneity is essential to deliver the targeted default rate and leverage. To illustrate the necessity of having heterogeneous returns, Fig. 6 depicts the relationship between X and both b (left) and ψ (right), assuming different variance of returns. As is clear from the figure, the default rate hardly rises above zero for $\sigma = \frac{1}{16}$ in normal times. This is because more concentrated return distributions have a thinner lower tail; hence, as b and $\bar{w}$ go up with capacity, the default rate barely changes. Having a concentrated distribution of returns ($\sigma = \frac{1}{16}$ or $\frac{1}{8}$) also leads to counterfactually high debt-to-equity ratios (b/y) and to low enforcement capacity utilization (ψ/X).

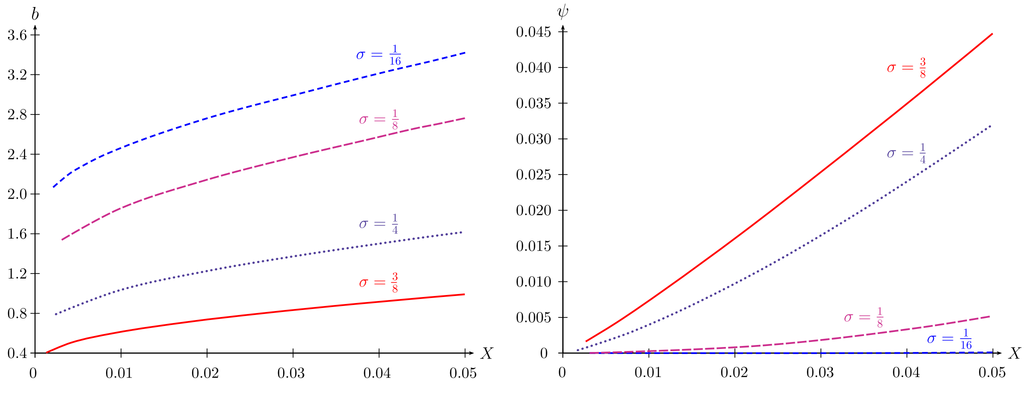

Fig. 6. Model-implied Relationship between Enforcement, Credit Supply and Default Rate. Notes: The figure shows the relationship between enforcement capacity X and loan size b (left) and default rates ψ (right), for log-normal distributions of returns $F(w)$ with different variances for the calibrated parameters.

Capacity costs (c_o) Our choice of unit capacity costs $c_o = 0.088$ represents 12.7% of the pre-liquidation value of the projects subject to liquidation in equilibrium, which we consider reasonable. This is roughly the midpoint of the two estimates reported by Bris et al. (2006), who use bankruptcy data from Arizona and New York State and estimate that direct bankruptcy costs represent 8.1% of the firm's pre-bankruptcy value in Chapter 7 bankruptcies (liquidation) and 16.9% in Chapter 11 (restructuring). Djankov et al. (2008) estimate that legal enforcement costs in the U.S. are 7% of the pre-bankruptcy value of a generic firm filing for bankruptcy. These include court and attorney fees, as well as other fees such as bankruptcy administrator, notification and accountant fees. These studies do not include indirect costs such as government funding of the bankruptcy court system not covered by court fees, thus underestimating overall enforcement costs. The value of this parameter is crucial, as it limits the ability of the planner to accumulate precautionary enforcement capacity. For these values, as we have shown, ex post enforcement capacity is a scarce resource.

Rents from the lack of enforcement (γ) The parameter γ is set equal to 33%, which in the model determines the recovery rates on defaulted loans when X falls below the critical threshold $\chi(w^*)$, i.e., when enforcement is imperfect ($P < 1$). To relate this aspect of the model to the data, we use the cross-country database on credit enforcement constructed by Djankov et al. (2008) to estimate the drop in recovery rates associated with an increase in enforcement delays illustrated in Fig. 1. To do so, we focus on the sub-sample of 20 countries with English legal origin in their dataset and run an OLS regression of the average recovery rate ($arecov$ in the dataset) on the time lag in years between default and receiving a payment ($atimepay$) and GDP per capita, which we include as a control for the level of economic development.[30] The regression predicts a 3.7 percentage point drop in recovery rates (from 85% to 81%) associated with the 38% increase in bankruptcy timelines experienced in the U.S. between 2008 and 2015. In the model, recovery rates drop about 4 percentage points (from a value of 77%) as X falls below $\chi(w^*)$.

30 We obtain a coefficient on $atimepay$ equal to -9.42, significant at the 1% level, with an adjusted R^2 of 0.84. The estimate is robust to introducing additional controls.

Appendix B. Proofs

We first prove equilibrium multiplicity under common knowledge and then present the proofs of the global games version of the model. First we prove Lemmas 1 and 2.

Proof of Lemma 1. Given the maximization problem (3), an entrepreneur with returns w finds it is optimal to repay if

$$u(1, w, 1, b, \bar{w})) \geq E(P)u(0, w, 1, b, \bar{w}) + (1 - E(P))u(0, w, 0, b, \bar{w}).$$

Substituting for $u(a, w, 1, b, \bar{w}))$, $a = 0, 1$, using (3), we get that repayment is optimal if

$$(y + b)(w - \bar{w}) \geq (1 - E(P))\gamma\mu(y + b)w.$$

Solving for $E(P)$ yields

$$E(P) \geq 1 - \frac{1}{\mu\gamma}\left(1 - \frac{\bar{w}}{w}\right) =: \theta(w, \bar{w}). \quad \square$$

Proof of Lemma 2. The first part of the Lemma is a direct consequence of the functional form of θ. To prove the second part, note that $\theta(\bar{w}, \bar{w}) = 1$. Also note that $\theta(w, \bar{w}) = 0$ if

$$1 - \frac{1}{\mu\gamma}\left(1 - \frac{\bar{w}}{w}\right) = 0 \Leftrightarrow w = \bar{w}/(1 - \mu\gamma).$$

Since $\theta(w, \bar{w})$ is strictly decreasing in w for all $\bar{w} > 0$, the set of returns for which $\theta(w, \bar{w}) \in (0, 1)$ is the interval $(\bar{w}, \bar{w}/(1 - \mu\gamma))$. For $w < \bar{w}$ the condition for repayment cannot be satisfied since $P \in [0, 1]$ but $\theta(w, \bar{w}) > 1$. Similarly, when $w > \bar{w}/(1 - \mu\gamma)$ the condition for default cannot be satisfied since $\theta(w, \bar{w}) < 0$. $\square$

Proof of Lemma 3. We simplify notation by writing $\chi(w)$ instead of $\chi(w, \bar{w})$ (same for θ). The existence and uniqueness of $\chi(w)$ directly follows from Assumption 1, given that $P(\cdot, \psi)$ is continuous, has full range in $[0, 1]$ and is strictly increasing at any X such that $P(X, \psi) < 1$. Hence, χ is a well-defined, continuous mapping from returns in $\mathcal{W} = (\bar{w}, \bar{w}/(1 - \gamma\mu))$ to enforcement capacity.

If P is given by (1) then it trivially follows that $\chi(w) = F(w)\theta(w)$ from the equilibrium condition (8). $F(w)\theta(w)$ is a continuous and real-valued function. The derivative of $\theta(w)F(w)$ is given by

$$\left(1 - \frac{1}{\mu\gamma}\left(1 - \frac{\bar{w}}{w}\right)\right) f(w) - \frac{\bar{w}}{w^2}\frac{1}{\mu\gamma}F(w),$$

which has the same sign as

$$1 - \frac{w}{\bar{w}}(1 - \mu\gamma) - \frac{F(w)}{wf(w)}.$$

If $\dfrac{F(w)}{wf(w)}$ is increasing, which follows from the first condition in Assumption 1, this expression is strictly decreasing. Now, since the second term is zero at $w = 0$ and $\dfrac{F(w)}{wf(w)}$ is increasing and since the second condition in Assumption 1 implies $\lim_{w\downarrow 0} \dfrac{F(w)}{wf(w)} < 1$, the expression – and hence the slope of $\theta(w)F(w)$ – is initially positive and eventually negative for high enough w. That is, $\theta(w)F(w)$ is single-peaked, i.e., it is strictly quasi-concave on $\mathcal{W}$. □

Proof of Proposition 1. Note that under the stated assumption $\chi(w)$ is single-peaked on $\mathcal{W}$. Let $w_{max} = \text{argmax}_{w \geq \bar{w}} \chi(w)$. First, consider the case of $X < \underline{X} := \chi(\bar{w})$. If $\bar{w} < w_{max}$, it must be that the equilibrium condition $\chi(w) = X$ has only one solution $w^* > w_{max}$ on $\mathcal{W}$. The existence of this solution follows from the fact that χ is continuous, increasing at $w = \bar{w}$ and single-peaked, and it is equal to zero at $\bar{w}/(1-\gamma\mu)$. The latter is due to the fact that $\theta(\bar{w}/(1-\gamma\mu)) = 0$ and hence the smallest capacity X satisfying $P((X, F(\bar{w}/(1-\gamma\mu))) = 0$ is $X = 0$ given property (iv) of admissible P in Definition 1. Accordingly, the equilibrium is unique in this case. The same argument applies when $\bar{w} \geq w_{max}$. If $X = \underline{X}$ then $w_1 = \bar{w}$ is also an equilibrium.

Second, let $X > \overline{X}$. In this case, $\chi(w)$ lies below X, implying that, for any given P such that all agents with default propensity less than P default, there is enough capacity so that the enforcement probability is higher than P, i.e., $\chi(w) < X$ for all w. Thus, equilibrium is unique at $w^* = \bar{w}$ and implies $\psi = F(\bar{w})$. If $X = \overline{X}$ then w_{max} is a solution of equilibrium condition $\chi(w) = X$, representing a second equilibrium.

Finally, if $X \in (\underline{X}, \overline{X})$ there are three equilibria, given that the quasi-concavity of χ yields two solutions on $\mathcal{W}$ to $\chi(w) = X$, and that there is an equilibrium at $w^* = \bar{w}$ with $\psi = F(\bar{w})$ since $\chi(\bar{w}) = \underline{X} < X$. □

Proof of Proposition 2. The "only if" part follows from the following argument. When χ is decreasing only three things can happen. If $\chi(\bar{w}) \leq X$ then $\chi(w) < X$ for all $w > \bar{w}$ and the unique equilibrium involves $\psi = F(\bar{w})$. If $\chi(w) = X$ has a unique solution $\hat{w}$, then neither $\chi(\bar{w}) \leq X$ nor $\chi(\bar{w}/(1-\gamma\mu)) \geq X$ hold, so the unique equilibrium implies $\psi = F(\hat{w})$. Finally, if $\chi(\bar{w}/(1-\gamma\mu)) \geq X$ then $\chi(w) > X$ for all $w < \bar{w}/(1-\gamma\mu)$ and $\psi = F(\bar{w}/(1-\gamma\mu)$. To complete the argument, consider the case in which $\chi(\cdot)$ is constant on some interval of returns. Any such interval is associated with a value of capacity X for which there exists a continuum of equilibria. However, since χ is monotone there are at most a countable number of such intervals, and thus the set of capacities at which there exist multiple equilibria has Lebesgue measure zero.

Now consider the "if" part. If $\chi(\cdot)$ is strictly increasing in some interval $(w_1, w_2) \subseteq (\bar{w}, \bar{w}/(1-\gamma\mu))$ then, the interval $(\chi(w_1), \chi(w_2))$ is non-degenerate, and for any $X \in (\chi(w_1), \chi(w_2))$ the following is true by the continuity of χ: (a) there exists $w \in (w_1, w_2)$ such that $\chi(w) = X$; (b) either there is $w' \in (\bar{w}, w_1)$ such that $\chi(w') = X$ or $\chi(\bar{w}) \leq X$; and (c) either there is $w'' \in (w_2, \bar{w}/(1-\gamma\mu))$ such that $\chi(w'') = X$ or $\chi(\bar{w}/(1-\gamma\mu)) \geq X$. Fact (a) implies there is always an equilibrium with $\psi = F(w)$. Fact (b) implies that there is at least another equilibrium with either $\psi = F(w')$ or $\psi = F(\bar{w})$. Finally, fact (c) involves at least a third equilibrium with $\psi = F(w'')$ or $\psi = F(\bar{w}/(1-\gamma\mu))$. □

B.1. Equilibrium of the global game

To prove the results, we proceed as follows. First, we present equilibrium existence, selection, and characterization results for the model with a generic discrete distribution of returns. Then, we provide the proofs of the results in the paper by deriving the limit equilibrium of sequence of games with discrete distributions that uniformly converge to F.

B.1.1. General discrete distribution of returns

In this economy, $\mathcal{W} \subset [0, \infty)$ is a finite set of possible returns, each with a positive mass, that are distributed according to the commonly known discrete distribution F, with probability mass function f.

Given contract $(b, \bar{w})$, we make the following assumption about agent payoffs.

Assumption 3. For all $w \in \mathcal{W}$,

(i) $u(0, w, 0) \neq u(1, w, 0)$; and
(ii) $u(0, w, 1) \neq u(1, w, 1)$.

Condition (i) says that no agent is indifferent between paying back the loan and defaulting when $P = 0$, i.e., there is no agent with $\theta(w) = 0$. Similarly, (ii) says that no agent is indifferent at $P = 1$, that is, there is no agent with $\theta(w) = 1$. This assumption is made for technical convenience, since it simplifies the proof of uniqueness by implying the existence of dominance regions for ν sufficiently small. It is not needed as F approximates a continuous distribution, since the mass of agents for which it is violated becomes arbitrarily small.

Note that agents with $w < \bar{w}$ ($\theta(w) > 1$) and those with $\theta(w) < 0$ behave in a non-strategic fashion: The former agents always choose $a = 0$, and the latter choose $a = 1$, regardless of P. Hence, our focus is on pinning down the behavior of types in the set $\mathcal{W}^* := \{w \in \mathcal{W} : \theta(w) \in (0, 1)\}$, with its lowest and highest elements respectively denoted w_l and w_h.

Abusing notation, let $\chi(w, w')$ be the enforcement capacity X satisfying $P(X, F(w')) = \theta(w)$. Note that $\chi(w, w')$ is strictly decreasing in w and increasing in w' for all $w, w' \in (\bar{w}, \bar{w}/(1 - \gamma\mu))$. Notice also that Assumption 1 implies that $\chi(w_h, \bar{w}) > 0$ since $P(0, \psi) = 0$ but $\theta(w_h) > 0$. Similarly, $\chi(w_l, \bar{w}/(1 - \gamma\mu)) < 1$ since $P(1, \psi) = 1$ by the continuity of P and condition (iii) in Assumption 1, but $\theta(w_l) < 1$.

To prove existence and characterize equilibrium with discrete distributions we assume that the noise scale factor satisfies $0 < \nu < \bar{\nu} := \min\{\chi(w_h, \bar{w}), 1 - \chi(w_l, \bar{w}/(1 - \gamma\mu))\}$.[31]

We first establish that there exists a unique equilibrium of the game with finite types, featuring cutoff strategies.

Theorem 1. *The game has an essentially unique equilibrium.*[32] *Equilibrium strategies are characterized by cutoffs $k(w)$ on signal x, such that all agents of type $w \in \mathcal{W}^*$ choose action $a = 1$ if $x \geq k(w)$ and $a = 0$ otherwise.*

Proof. The proof logic is as follows. First, we argue that the set of equilibrium strategy profiles has a largest and a smallest element, each involving monotone (cutoff) strategies. Second, we show that there is at most one equilibrium in monotone strategies (up to differences in behavior at cutoff signals). But this implies that the equilibrium is essentially unique.

The existence of a smallest and a largest equilibrium profile in monotone strategies follows from existing results on supermodular games by Milgrom and Roberts (1990), Vives (1990) and Van Zandt and Vives (2007). Consider the game in which we fix the profile $\mathbf{x}$ of signal realizations

[31] This upper bound on ν is helpful to show uniqueness of equilibrium by ensuring that boundary issues associated with signals close to 0 or 1 only arise when capacity is such that all agents have a dominant strategy.

[32] In the sense that equilibrium strategies may differ in zero probability events.

and agents choose actions {0, 1} after observing their own signals. It is straightforward to check that the game satisfies the conditions of Theorem 5 in Milgrom and Roberts (1990), which states that the game has a smallest and a largest equilibrium. That is, there exist two equilibrium strategy profiles, $\underline{\mathbf{a}}(\mathbf{x})$ and $\bar{\mathbf{a}}(\mathbf{x})$ such that any equilibrium profile $\mathbf{a}(\mathbf{x})$ satisfies $\underline{\mathbf{a}}(\mathbf{x}) \leq \mathbf{a}(\mathbf{x}) \leq \bar{\mathbf{a}}(\mathbf{x})$. Moreover, if we fix the action profile of all agents, the difference in expected payoff from choosing $a = 0$ versus $a = 1$ for any given agent is increasing in $\mathbf{x}$ since default rates are the same across signal profiles while X is higher (in expectation) at higher signal profiles, thus implying a higher expected enforcement probability. That is, expected payoffs exhibit increasing differences w.r.t. $\mathbf{x}$, and Theorem 6 in Milgrom and Roberts (1990) applies: $\underline{\mathbf{a}}(\mathbf{x})$ and $\bar{\mathbf{a}}(\mathbf{x})$ are nondecreasing functions of $\mathbf{x}$. But, because an agent's strategy can only depend on her own signal, all agents must be following cutoff strategies.

To show that there is at most one equilibrium in monotone strategies, we make use of the next two lemmas. The first lemma shows that equilibrium cutoffs are bounded away from zero and one. The second lemma uses these bounds to establish the following translation result: When all cutoffs are shifted by the same amount Δ, expected enforcement probabilities go up. Equipped with such results we will show that enforcement probabilities go up as we move from the smallest to the largest equilibria, implying that there must be a unique profile of cutoffs at which indifference conditions (20) are satisfied.

Let $\mathbf{k} + \mathbf{\Delta} = (k(w) + \Delta)_{w \in \mathcal{W}^*}$, while $\underline{\mathbf{k}}$ and $\bar{\mathbf{k}}$ represent the profile of cutoffs associated with the smallest and largest equilibrium, respectively. Abusing notation, let $E[P|\mathbf{k}; x]$ represent the expected enforcement probability of an agent receiving signal x when agents use cutoff profile $\mathbf{k}$.

Lemma 5. *If* $\mathbf{k}$ *is a profile of equilibrium cutoffs then* $k(w) \in [\chi(w, \bar{w}) - \nu/2, \chi(w, \bar{w}/(1 - \gamma\mu)) + \nu/2]$ *for all* $w \in \mathcal{W}^*$.

Proof. Note that k is an equilibrium if it solves the following set of indifference conditions:

$$E[P|\mathbf{k}; k(w)] = \theta(w) \qquad \forall w \in \mathcal{W}^*. \tag{20}$$

Note also that the value of X conditional on x is at least $x - \nu/2$ and that $\psi \leq F(\bar{w}/(1 - \gamma\mu))$. Given this and the fact that P is increasing in X and decreasing in ψ we have that

$$E[P(X, \psi)|\mathbf{k}; k(w)] \geq P(k(w) - \nu/2, F(\bar{w}/(1 - \gamma\mu))).$$

But this implies that $E[P|\mathbf{k}; k(w)] > \theta(w)$ when $k(w) > \chi(w, \bar{w}/(1 - \gamma\mu)) + \nu/2$, violating the indifference condition of type w. This is because $P(\chi(w, \bar{w}/(1 - \gamma\mu)), F(\bar{w}/(1 - \gamma\mu))) = \theta(w)$ and, by Assumption 1, P is strictly increasing at $(\chi(w, \bar{w}/(1 - \gamma\mu)), F(\bar{w}/(1 - \gamma\mu)))$ since $\theta(w) < 1$. Likewise, the value of X conditional on x is at most $x + \nu/2$ and $\psi \geq F(\bar{w})$. Accordingly,

$$E\left[P(X, \psi)|\mathbf{k}; k(w)\right] \leq P(k(w) + \nu/2, F(\bar{w})),$$

which yields the above lower bound on $k(w)$ when we replace $E[P|\mathbf{k}; k(w)]$ with $\theta(w)$. □

Lemma 6. *If* $\mathbf{k}$ *is a profile of equilibrium cutoffs then* $E[P|\mathbf{k}; k(w)] < E[P|\mathbf{k} + \mathbf{\Delta}; k(w) + \Delta]$ *for all* $\Delta > 0$ *and all* $w \in \mathcal{W}^*$ *such that* $k(w) + \Delta \leq \bar{k}(w)$.

Proof. First, note that the density of X conditional on an agent receiving signal $x \in [\nu/2, 1 - \nu/2]$ is given by $h\left(\frac{x-X}{\nu}\right)$. Also notice that an agent of type w defaults if she receives a signal

$x < k(w)$, and thus, that the fraction of type-w agents defaulting when capacity is X is given by $H\left(\frac{k(w)-X}{\nu}\right)$. Since $\nu \leq \chi(w_h, \bar{w}) \leq \chi(w, \bar{w})$ and, by Lemma 5, $k(w) \geq \chi(w, \bar{w}) - \nu/2$, we have that $k(w) \geq \nu/2$. Likewise, $k(w) + \Delta \leq \bar{k}(w) \leq 1 - \nu/2$ by Lemma 5 and the fact that $\nu \leq 1 - \chi(w_l, \bar{w}/(1-\gamma\mu)) \leq 1 - \chi(w, \bar{w}/(1-\gamma\mu))$. Hence, we can obtain the following inequality by a well-defined change of variables:

$$
\begin{aligned}
&E[P(X)|\mathbf{k}; k(w)] = \\
&\quad \int_{k(w)-\nu/2}^{k(w)+\nu/2} P\left(X, F(\bar{w}) + \sum_{w'} H\left(\frac{k(w')-X}{\nu}\right) f(w')\right) h\left(\frac{k(w)-X}{\nu}\right) dX \\
&\quad < \int_{k(w)-\nu/2}^{k(w)+\nu/2} P\left(X+\Delta, F(\bar{w}) + \sum_{w'} H\left(\frac{k(w')-X}{\nu}\right) f(w')\right) h\left(\frac{k(w)-X}{\nu}\right) dX \\
&\quad = \int_{k(w)+\Delta-\nu/2}^{k(w)+\Delta+\nu/2} P\left(X', F(\bar{w}) + \sum_{w'} H\left(\frac{k(w')+\Delta-X'}{\nu}\right) f(w')\right) \\
&\quad\quad \times h\left(\frac{k(w)+\Delta-X'}{\nu}\right) dX' \\
&\quad = E[P|\mathbf{k}+\boldsymbol{\Delta}; k(w)+\Delta].
\end{aligned}
$$

The inequality is strict because $\mathbf{k}$ being an equilibrium profile means that $E[P|\mathbf{k}; k(w)] = \theta(w) < 1$ for all $w \in \mathcal{W}^*$. Accordingly, enforcement probabilities, conditional on $x = k(w)$, are less than 1 for a positive measure of $X \in [x - \nu/2, x + \nu/2]$, and hence, expected enforcement probabilities go up strictly when capacity increases by Δ by Assumption 1. □

Equipped with Lemma 6, we can argue that $\underline{\mathbf{k}} = \bar{\mathbf{k}}$. Assume, by way of contradiction, that $\underline{k}(w) < \bar{k}(w)$ for some $w \in \mathcal{W}^*$. Denote $\hat{w} = \arg\max_{w \in \mathcal{W}^*}(\bar{k}(w) - \underline{k}(w))$ and $\hat{\Delta} = \bar{k}(\hat{w}) - \underline{k}(\hat{w})$. By Lemma 6, we have that

$$\theta(\hat{w}) = E[P|\underline{\mathbf{k}}; \underline{k}(\hat{w})] < E[P|\underline{\mathbf{k}} + \hat{\boldsymbol{\Delta}}; \bar{k}(\hat{w})] \leq E[P|\bar{\mathbf{k}}; \bar{k}(\hat{w})] = \theta(\hat{w}),$$

where the last inequality comes from the fact that default rates at $\bar{\mathbf{k}}$ are lower than at $\underline{\mathbf{k}} + \hat{\boldsymbol{\Delta}} \geq \bar{\mathbf{k}}$, and thus, the expected enforcement probability conditional on $x = \bar{k}(\hat{w})$ is higher. □

Next, we proceed to characterize the limit equilibrium as $\nu \to 0$. First, note that for threshold profile $k(\cdot)$ to be an equilibrium profile, it must satisfy the set of indifference conditions (20).

To solve this system of equations, we need to pin down agent beliefs when they receive their threshold signals. To do so, we make use of a generalized version of the original *belief constraint* (Sákovics and Steiner, 2012): On average, conditional on $x = k(w)$, agents with types in a subset $W' \subseteq \mathcal{W}^*$ believe that the default rate of agents in W' is uniformly distributed in [0, 1]. The default rate in W' when capacity is X is given by

$$\psi(X, W') = \frac{1}{\sum_{W'} f(w)} \sum_{W'} H\left(\frac{k(w)-X}{\nu}\right) f(w). \tag{21}$$

Lemma 7 *(belief constraint). For any subset $W' \subseteq \mathcal{W}^*$ and any $z \in [0,1]$,*

$$\frac{1}{\sum_{W'} f(w)} \sum_{W'} Pr\left(\psi(X, W') \le z \middle| x = k(w)\right) f(w) = z, \tag{22}$$

where $Pr\left(\cdot \middle| x = k(w)\right)$ is the probability assessment of $\psi(W', X)$ by an agent receiving $x = k(w)$.

Proof. See the *Online Appendix*. □

The previous result is instrumental in characterizing equilibrium thresholds as ν goes to zero. In particular, it allows us to derive closed-form solutions for the above indifference conditions from which we can obtain $\mathbf{k}$. In stating this result, we refer to a partition $\Phi = \{W_1, \cdots, W_I\}$ of $\mathcal{W}^*$ as being monotone if $\max W_i < \min W_{i+1}$, $i = 1, \cdots, I-1$, and denote the infimum and the supremum of W_i by $\underline{w}_i$ and $\bar{w}_i$, respectively. Also, let $F^-(w) = \sum_{w'<w} f(w')$ and $\chi^-(w)$ the capacity X satisfying $P(X, F^-(w)) = \theta(w)$.

Theorem 2. *In the limit, as $\nu \to 0$, the equilibrium cutoff strategies are given by a unique monotone partition $\Phi = \{W_1, \cdots, W_I\}$ and a unique vector $(k_1, \cdots, k_I)$ satisfying the following conditions:*

(i) $k(w) = k(w') = k_i$ *for all* $w, w' \in W_i$.
(ii) $k_i > k_{i+1}$ *for all* $i = 1, \cdots, I-1$.
(iii) $\chi^-(\underline{w}_i) \le k_i \le \chi(\bar{w}_i)$ *for all* $i = 1, \cdots, I$.
(iv) $\int_{F^-(\underline{w}_i)}^{F(\bar{w}_i)} P(k_i, z)dz = \sum_{W_i} \theta(w) f(w)$ *for all* $i = 1, \cdots, I$.

Proof. From Theorem 1 we know that for each $\nu > 0$, there exists essentially a unique equilibrium, which is in monotone strategies. Let $k^\nu(w)$ represent the equilibrium threshold of type-w agents associated with $\nu > 0$, with $\mathbf{k}^\nu$ denoting the equilibrium cutoff profile. The first step of the proof is to show that $\mathbf{k}^\nu$ uniformly converges as $\nu \to 0$, and to identify the set of indifference conditions that pin down the limit equilibrium. Let

$$A_w(z|\mathbf{k}^\nu, W') := Pr\left(\psi(X, W') \le z \middle| x = k^\nu(w)\right)$$

denote the strategic belief of an agent of type $w \in W'$ when she receives her threshold signal $x = k^\nu(w)$.

Lemma 8. *There exist a unique partition $\{W_1, \cdots, W_I\}$ and a set of thresholds $k_1 > k_2 > \cdots > k_I$ such that, as $\nu \to 0$, for all $w \in W_i$, $i = 1, \cdots, I$, $k^\nu(w)$ uniformly converges to k_i. Moreover, thresholds $\mathbf{k} = (k_1, \cdots, k_I)$ solve the system of limit indifference conditions*

$$\int_0^1 P\left(k_i, F(\bar{w}) + \sum_{\cup_{j<i} W_j} f(w') + z \sum_{W_i} f(w')\right) dA_w(z|\mathbf{k}, W_i) = \theta(w), \quad \forall w \in W_i, \ \forall i, \tag{23}$$

where $A_w(z|\mathbf{k}, W_i)$ represents the strategic beliefs of type-w agents in the limit and satisfies the belief constraint (22).

Proof. See the *Online Appendix*. □

Equipped with this set of indifference conditions, we next prove that the partition of types is monotone and that thresholds satisfy (iii) and (iv) in the theorem.

We show that the partition of types must be monotone by way of contradiction. Assume that there are two types $w > \hat{w}$ such that $w \in \mathcal{W}_i$ and $\hat{w} \in \mathcal{W}_m$ with $m > i$. First note that the LHS in (23) is bounded below by $P(k_i, F(\bar{w}) + \sum_{\cup_{j \leq i} W_j} f(w'))$ and bounded above by $P(k_i, F(\bar{w}) + \sum_{\cup_{j<i} W_j} f(w'))$. Given this, since $\theta(\hat{w}) < 1$ the enforcement probability when all agents with types in W_m default is strictly less than 1, i.e., $P(k_m, F(\bar{w}) + \sum_{\cup_{j \leq m} W_j} f(w')) < 1$. Otherwise, (23) would be violated. In addition, $m > i$ implies that $k_m < k_i$ by the above lemma and that $\sum_{\cup_{j<i} W_j} f(w') < \sum_{\cup_{j \leq m} W_j} f(w')$. Combining all this, we arrive at the following contradiction

$$\theta(w) \geq P\left(k_i, F(\bar{w}) + \sum_{\cup_{j<i} W_j} f(w')\right) > P\left(k_m, F(\bar{w}) + \sum_{\cup_{j \leq m} W_j} f(w')\right) \geq \theta(\hat{w}).$$

The monotonicity of the type partition implies that $F(\bar{w}) + \sum_{\cup_{j \leq i} W_j} f(w') = F(\overline{w}_i)$ and that $F(\bar{w}) + \sum_{\cup_{j<i} W_j} f(w') = F^-(\underline{w}_i)$. Given this, it is straightforward to check that the above bounds on the LHS in (23) lead to condition (iii) in the theorem.

Finally, to obtain condition (iv) from (23), we make use of the belief constraint in the limit, which can be written as

$$\frac{1}{\sum_{W_i} f(w)} \sum_{W_i} A_w(z|\mathbf{k}, W_i) f(w) = z. \tag{24}$$

Multiplying both sides of (23) by $\frac{f(w)}{\sum_{W_i} f(w)}$ and summing over $w \in W_i$ we get

$$\int_0^1 P\left(k_i, F^-(\underline{w}_i) + z \sum_{W_i} f(w')\right) d\left(\frac{1}{\sum_{W_i} f(w)} \sum_{W_i} A_w(z|\mathbf{k}, W_i) f(w)\right)$$
$$= \frac{\sum_{W_i} \theta(w) f(w)}{\sum_{W_i} f(w)}.$$

Using the belief constraint (24) to substitute for the last term in the LHS, we obtain

$$\int_0^1 P\left(k_i, F^-(\underline{w}_i) + z \sum_{W_i} f(w')\right) dz = \frac{\sum_{W_i} \theta(w) f(w)}{\sum_{W_i} f(w)}. \tag{25}$$

Note that $F^-(\underline{w}_i) + z \sum_{W_i} f(w') \sim U[F^-(\underline{w}_i), F(\bar{w}_i)]$ with density $\frac{1}{\sum_{W_i} f(w)}$ since $z \sim U[0, 1]$. Hence, we can rewrite (25) as

$$\frac{1}{\sum_{W_i} f(w)} \int_{F^-(\underline{w}_i)}^{F(\bar{w}_i)} P(k_i, z)dz = \frac{\sum_{W_i} \theta(w) f(w)}{\sum_{W_i} f(w)},$$

yielding condition (iv). □

B.1.2. Continuous distribution of returns

We next prove Propositions 3 and 4 by showing that, for any sequence of discrete distributions that converges to F, the limit equilibrium converges to the unique equilibrium characterized by Proposition (4).

Proof of Propositions 3 and 4. To prove both propositions we show that (uniform) convergence of discrete return distributions to F implies that the equilibrium characterization in Theorem 2 leads to the limit characterization in Proposition 4. Since Theorem 2 pins down the unique equilibrium satisfying agents' indifference conditions $E(P|\mathbf{k}, k(w)) = \theta(w)$ while P and θ are Lipschitz continuous and bounded in $[\bar{w}, \bar{w}/(1-\gamma\mu)]$, the sequence of equilibrium thresholds must converge, with its limit being the solution to the limit equilibrium conditions in Proposition 4.

Let $\{F^n\}_{n=1}^{\infty}$ be any sequence of discrete distributions satisfying Assumption 3 that uniformly converges to F and thus $\lim_{n\to\infty} F^n(w) = F(w)$ and $\lim_{n\to\infty} F^{n-}(w) = F(w)$ for all $w \in [0, \infty)$. This implies that $\chi^-(w, w') \to \chi(w, w')$ for all w, w' since $P(\cdot)$ and $\chi(\cdot)$ are continuous and bounded.

The limit characterization in Proposition 4 is derived as follows. Part (a) follows from conditions (i)–(ii) in Theorem 2, meaning that k is decreasing, so we can partition the space of types into a collection of successive intervals in which k alternates between being strictly decreasing and constant. Part (b) follows from (ii)–(iii): a strictly decreasing k in a given interval of types is approximated by a (growing) collection of consecutive, singleton W_i in the discrete economy since $\chi^-(w) \to \chi(w)$ and thus $\underline{w}_i \to \overline{w}_i$ in condition (iii) whenever $\chi(\cdot)$ is strictly decreasing in $(\underline{w}_i, \overline{w}_i)$. But then, as the mass associated with each of these singletons goes to zero, F^- approximates F, and condition (iii) implies that k converges to $\chi(w)$.

Part (c) follows from parts (a) and (b) and conditions (iii) and (iv). Since $\chi(w)$ is continuous and bounded, parts (a) and (b) imply that $k(w) = \chi(w)$ at the boundaries of an interval in which k is constant, except possibly when $\underline{w}_i = \bar{w}$, in which case condition (iii) requires that $k_j \geq \chi(\underline{w}_j)$. Expression (10) is the limit of (iv). □

Proof of Proposition 5. We now argue that $\chi(w)$ being increasing at $\bar{w}$ and quasi-concave leads to a partition consisting of two intervals, the first one in which k is constant and the second one in which it is strictly decreasing.

First, notice that k decreasing implies that there must be at least one pooling threshold because $\chi(w)$ is initially increasing. To show why there is only one, we use the fact that condition (c) requires that $k_1 = \chi(\bar{w}_1) \geq \chi(\underline{w}_1)$. First notice that $\chi(\bar{w}/(1-\gamma\mu)) = 0$ since $\theta(\bar{w}/(1-\gamma\mu)) = 0$ but $P(X, F(\bar{w}/(1-\gamma\mu))) > 0$ for $X > 0$ by Assumption 1. Hence, $\chi(w)$ has a single peak in $[\bar{w}, \bar{w}/(1-\gamma\mu)]$, given that it is initially increasing and strictly quasi-concave. Accordingly, we must have that $\chi(w)$ is increasing at $\underline{w}_1$ and decreasing at $\bar{w}_1$. Otherwise, either $\chi(w)$ is decreasing at $\underline{w}_1$ or $\chi(w)$ is increasing in $(\underline{w}_1, \bar{w}_1)$. The former case implies that $\chi(\bar{w}_1) < \chi(\underline{w}_1)$, violating (c). The latter case implies that $k_1 = \max_{w\in[\underline{w}_1, \bar{w}_1]} \chi(w)$. But then condition (10) is violated since

$$\int_{F(\underline{w}_1)}^{F(\bar{w}_1)} P(k_1, z)dz > \int_{F(\underline{w}_1)}^{F(\bar{w}_1)} P(\chi(F^{-1}(z)), z)dz = \int_{\underline{w}_j}^{\bar{w}_j} \theta(w) f(w)dw,$$

where the last equality comes from the change in variable $w = F^{-1}(z)$ $(dz = f(w)dw)$ and thus $P(\chi(F^{-1}(z), z) = P(\chi(w), F(w))$, which is equal to $\theta(w)$ by the definition of $\chi(w)$.

Accordingly, by single peakedness, if $\chi(w)$ is increasing at $\underline{w}_1$ and decreasing at $\bar{w}_1$ we cannot find another interval satisfying the same condition that does not intersect with $[\underline{w}_i, \overline{w}_i]$. Thus, there must be a unique interval of returns at which k is constant. Finally, since $\chi(w)$ is increasing in $[\bar{w}, \underline{w}_1]$, the monotonicity of $k(w)$ requires that $\underline{w}_1 = \bar{w}$.

We finish the proof by showing that $\bar{w}_1 = w^*$, where w^* is the unique solution to (11) in $(\bar{w}, \bar{w}/(1-\gamma\mu))$. To do so, given the above change of variable, we write (11) as follows:

$$\int_{F(\bar{w})}^{F(w^*)} P(\chi(w^*), z)dz - \int_{F(\bar{w})}^{F(w^*)} P(\chi(F^{-1}(z)), z)dz = 0. \tag{26}$$

We show that this equation must have a unique solution by arguing that the LHS is equal to zero and strictly increasing at $w^* = \bar{w}$, strictly quasi-concave in $(\bar{w}, \bar{w}/(1-\gamma\mu))$, and negative at $w^* = \bar{w}/(1-\gamma\mu)$.

Denote $w_{\max}$ the unique maximizer of $\chi(\cdot)$ in $(\bar{w}, \bar{w}/(1-\gamma\mu))$. Since $\chi(w)$ is increasing in $(\bar{w}, w_{\max})$ and $F^{-1}(z) < w^*$ for all $z < F(w^*)$, the difference $P(\chi(\cdot), z) - P(\chi(F^{-1}(z)), z)$ is positive and increasing in $(\bar{w}, w_{\max})$. Accordingly, the LHS of (26) is increasing in $[\bar{w}, w_{\max})$ from an initial value of zero. Similarly, $\chi(\cdot)$ being decreasing in $(w_{\max}, \bar{w}/(1-\gamma\mu))$ implies that the LHS is decreasing in that interval. To see why, express the LHS of (26) as

$$\int_{F(\bar{w})}^{F(w_{\max})} \Big[P(\chi(w^*), z) - P(\chi(F^{-1}(z)), z)\Big] dz + \int_{F(w_{\max})}^{F(w^*)} \Big[P(\chi(w^*), z) - P(\chi(F^{-1}(z)), z)\Big] dz.$$

Both integrals are decreasing in w^* for $w^* > w_{\max}$ since $P(\chi(\cdot), z) - P(\chi(F^{-1}(z)), z)$ is decreasing in $(w_{\max}, \bar{w}/(1-\gamma\mu))$. In addition, the second integral is negative since $\chi(w^*) < \chi(F^{-1}(z))$ due to the fact that $F^{-1}(z) < w^*$ and $\chi(\cdot)$ is decreasing in $(w_{\max}, \bar{w}/(1-\gamma\mu))$. Finally, notice that $P(\chi(w^*), z) = 0$ when $w^* = \bar{w}/(1-\gamma\mu)$ for all $z \leq F(w^*)$, implying that the LHS of (26) is strictly negative at $w^* = \bar{w}/(1-\gamma\mu)$. □

Proof of Corollary 1. We prove the corollary by showing that, when $\bar{w} < w_{max}$, w^* is the unique solution to (12). Obviously, if $\bar{w} \geq w_{max}$ then $\theta(w)F(w)$ is strictly decreasing for $w > \bar{w}$, and conditions (a)–(c) in Proposition 4 lead to $k(w) = \theta(w)F(w)$, i.e., to $w^* = \bar{w}$.

Condition (c) in Proposition 4 implies that $k_1 \geq \theta(\bar{w})F(\bar{w}) = F(\bar{w})$. Hence, solving the integral and substituting for $k_1 = \theta(w^*)F(w^*)$ and $\underline{w}_1 = \bar{w}$, we can express the LHS of (11) as

$$\int_{k_i}^{F(w^*)} \frac{k_i}{z} dz + \int_{F(\bar{w})}^{k_i} dz = k_i \log\left(\frac{F(w^*)}{k_i}\right) + k_i - F(\bar{w})$$
$$= \theta(w^*)F(w^*)(1 - \log\theta(w^*)) - F(\bar{w}).$$

Equating the RHS of the last expression to the RHS of (11) yields (12). □

We next prove Proposition 6. Before that, we establish that the lenders' problem has an interior solution.

Lemma 9. *The solution to* (15) *involves finite b for all* $X \in [0, 1]$.

Proof. To prove that lenders always set $b < \infty$, first note that Proposition 5 implies the existence of a type $\hat{w} \geq \bar{w}$ such that $a = 0$ if $w < \hat{w}$ and $a = 1$ otherwise. Therefore, we can express the budget constraint (16) as follows:

$$\frac{b}{y+b} \leq \int_{\hat{w}}^{\infty} \bar{w} dF(w) + \mu (P + (1-\gamma)(1-P)) \int_{0}^{\hat{w}} w dF(w). \tag{27}$$

The LHS converges to one as $b \to \infty$. Hence, we need to show that the RHS is strictly less than 1 for all $\bar{w}$. Since $\hat{w} \geq \bar{w}$ and the RHS is increasing in P for fixed $\hat{w}$, the RHS is bounded above by

$$\int_{\hat{w}}^{\infty} \hat{w} dF(w) + \mu \int_{0}^{\hat{w}} w dF(w), \tag{28}$$

which is strictly less than one for all $\hat{w}$ by Assumption 2. □

Proof of Lemma 4. We need to show that $\chi(w^*)$ is increasing in $\bar{w}$. Note that the propensity to default θ goes up with $\bar{w}$ and that $\chi(w)$ is decreasing at w^*. Given this, if we can show that w^* goes down after an increase in $\bar{w}$ then we would have proven that $\chi(w^*)$ increases with $\bar{w}$.

In the proof of Proposition 5, we showed that the LHS of indifference condition (26) is decreasing in $[w_{\max}, \bar{w}/(1-\gamma\mu)]$ and thus differentiable almost everywhere w.r.t. w^* in that interval of returns. In addition, we argued that $P(\chi(\cdot), z) - P(\chi(F^{-1}(z)), z)$ is positive in $(\bar{w}, w_{\max})$. Hence, the partial derivative of the LHS w.r.t. $\bar{w}$ is negative, since the only effect of an infinitesimal increase in $\bar{w} < w_{\max}$ is to increase the lower limit of the integral of $P(\chi(\cdot), z) - P(\chi(F^{-1}(z)), z)$. Given this, we can implicitly differentiate (26) to obtain

$$\frac{\partial LHS}{\partial w^*}\frac{dw^*}{d\bar{w}} + \frac{\partial LHS}{\partial \bar{w}} = 0.$$

Since both $\frac{\partial LHS}{\partial w^*}$ and $\frac{\partial LHS}{\partial \bar{w}}$ are negative, it must be that $\frac{dw^*}{d\bar{w}} < 0$. □

Proof of Proposition 6. Having shown that $\chi(w^*)$ is increasing in $\bar{w}$ (Lemma 4), we just need to prove that the equilibrium $\bar{w}$ is strictly increasing in b when $X = \chi(w^*)$. This is because in such a case, a drop from X to $X' < X$ must lead to either a drop in b so that $X' \geq \chi(w^{*'})$ in the new equilibrium, or the equilibrium default rate necessarily satisfies $\psi' > F(w^{*'}) > F(\bar{w}')$ because $X' < \chi(w^{*'})$.

To do so, we first note that the budget constraint (27) must hold with equality when $X = \chi(w^*)$. To see why, notice that the propensity to default $\theta(\cdot)$ and hence $k(\cdot)$ does not depend on b by Proposition 1. Accordingly, for any given $\bar{w}$, the objective function (15) is strictly increasing in b since P, $\{w : a = 0\}$ and $\{w : a = 1\}$ are constant for all b, whereas payoffs under both repayment and default are strictly increasing in b. Hence, since the LHS of budget constraint (27) is strictly increasing in b, for any given $\bar{w}$, lenders choose b so that the budget constraint binds.

Given that the budget constraint holds with equality and the LHS of (27) is increasing in b, we only need to show that the RHS of (27) is increasing in $\bar{w}$ at the equilibrium $\bar{w}$ when $X = \chi(w^*)$. We do so by contradiction. Assume that the RHS is strictly decreasing in $\bar{w}$ at the equilibrium contract associated with X. Since $X = \chi(w^*)$, we must have that all agents with $w < \bar{w}$ default and those with $w \geq \bar{w}$ repay, implying that $\hat{w} = \bar{w}$. In this context, lowering $\bar{w}$ while increasing b is feasible, since lenders' revenue goes up after a reduction in $\bar{w}$ (the RHS is strictly decreasing by assumption) while the cluster threshold goes down since $\chi(w^*)$ is decreasing in $\bar{w}$, so $X \geq \chi(w^*)$ is still satisfied at the new contract. But notice that such an alternative contract strictly increases agents' expected payoffs since it raises gross returns while reducing the deadweight loss of defaults, which are reverted back to agents' payoffs given that the zero profit condition binds. Accordingly a lender can deviate and offer a contract close to this alternative contract and make positive profits, contradicting that the original contract is an equilibrium.[33] Hence, the equilibrium $\bar{w}$ must go up with b when $X = \chi(w^*)$. □

Appendix C. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jet.2018.09.002.

References

Agarwal, Sumit, Amromin, Gene, Ben-David, Itzhak, Chomsisengphet, Souphala, Piskorski, Tomasz, Seru, Amit, 2017. Policy intervention in debt renegotiation: evidence from the home affordable modification program. J. Polit. Econ. 125 (3), 654–712.
Angeletos, George-Marios, Werning, Iván, 2006. Crises and prices: information aggregation, multiplicity, and volatility. Am. Econ. Rev., 1720–1736.
Angeletos, George-Marios, Hellwig, Christian, Pavan, Alessandro, 2006. Signaling in a global game: coordination and policy traps. J. Polit. Econ. 114, 452–484.
Arellano, Cristina, Kocherlakota, Narayana, 2014. Internal debt crises and sovereign defaults. J. Monet. Econ., S68–S80.
Bae, Kee-Hong, Goyal, Vidhan K., 2009. Creditor rights, enforcement, and bank loans. J. Finance 64 (2), 823–860.
Bank of Italy. Financial Stability Report. Technical Report 5 April 2013.
Bassetto, Marco, Phelan, Christopher, 2008. Tax riots. Rev. Econ. Stud. 75 (3), 649–669.
Bernanke, Ben S., Gertler, Mark, Gilchrist, Simon, 1999. The financial accelerator in a quantitative business cycle framework. Handb. Macroecon. 1, 1341–1393.
Bond, Philip, Rai, Ashok S., 2009. Borrower runs. J. Dev. Econ. 88 (2), 185–191.
Bond, Philip, Hagerty, Kathleen, 2010. Preventing crime waves. Am. Econ. J. Microecon. 2 (3), 138–159.
Bris, Arturo, Welch, Ivo, Zhu, Ning, 2006. The costs of bankruptcy: Chapter 7 liquidation versus Chapter 11 reorganization. J. Finance 61 (3), 1253–1303.
Carlsson, Hans, van Damme, Eric, 1993. Global games and equilibrium selection. Econometrica 61 (5), 989–1018.

[33] Formally, note that the enforcement probability in equilibrium is $P(\chi(w^*), F(\bar{w})) = 1$, given that $\chi(w^*) \geq \chi(\bar{w})$ and $\theta(\bar{w}) = 1$. Hence, agents' payoffs are given by $\int_{\bar{w}}^{\infty}(y + b)(w - \bar{w})dF$, which strictly go up if we increase b and lower $\bar{w}$.

Carpinelli, Luisa, Cascarino, Giuseppe, Giacomelli, Silvia, Vacca, Valerio Paolo, 2016. The management of non-performing loans: a survey among the main Italian banks. Quest. Econ. Finanza (311).

Carrasco, Vinicius, Salgado, Pablo, 2014. Coordinated strategic defaults and financial fragility in a costly state verification model. J. Financ. Intermed. 23 (1), 129–139.

Chan, Sewin, Haughwout, Andrew, Hayashi, Andrew, Van der Klaauw, Wilbert, 2016. Determinants of mortgage default and consumer credit use: the effects of foreclosure laws and foreclosure delays. J. Money Credit Bank. 48 (2–3), 393–413.

Christiano, Lawrence J., Motto, Roberto, Rostagno, Massimo, 2014. Risk shocks. Am. Econ. Rev. 104 (1), 27–65.

Cordell, Larry, Geng, Liang, Goodman, Laurie, Yang, Lidan, 2015. The cost of foreclosure delay. Real Estate Econ. 43 (4), 916–956.

Corsetti, Giancarlo, Dasgupta, Amil, Morris, Stephen, Shin, Hyun Song, 2004. Does one Soros make a difference? A theory of currency crises with large and small traders. Rev. Econ. Stud. 71 (1), 87–113.

Dávila, Eduardo, Walther, Ansgar, 2018. Does Size Matter? Bailouts with Large and Small Banks. NBER Working Paper 24132.

Dávila, Eduardo, Korinek, Anton, 2018. Pecuniary externalities in economies with financial frictions. Rev. Econ. Stud. 85 (1), 352–395.

Djankov, Simeon, McLiesh, Caralee, Shleifer, Andrei, 2007. Private credit in 129 countries. J. Financ. Econ. 84 (2), 299–329.

Djankov, Simeon, Hart, Oliver, McLiesh, Caralee, Shleifer, Andrei, 2008. Debt enforcement around the world. J. Polit. Econ. 116 (6), 1105–1149.

Enoch, Charles, Baldwin, Barbara, Frécaut, Olivier, Kovanen, Arto, 2001. Indonesia: Anatomy of a Banking Crisis. Two Years of Living Dangerously, 1997–99. IMF Working Paper No. 0152.

Favara, Giovanni, Schroth, Enrique, Valta, Philip, 2012. Strategic default and equity risk across countries. J. Finance 67 (6), 2051–2095.

Frankel, David M., Morris, Stephen, Pauzner, Ady, 2003. Equilibrium selection in global games with strategic complementarities. J. Econ. Theory 108 (1), 1–44.

Gale, Douglas, Hellwig, Martin, 1985. Incentive-compatible debt contracts: the one-period problem. Rev. Econ. Stud. 52 (4), 647–663.

Gomes, Joao F., Jermann, Urban, Schmid, Lukas, 2016. Sticky leverage. Am. Econ. Rev. 106 (12), 3800–3828.

Herkenhoff, Kyle F., Ohanian, Lee, 2015. The Impact of Foreclosure Delay on U.S. Unemployment. NBER Working Paper No. 21532.

Ippolito, Filippo, Peydró, José-Luis, Polo, Andrea, Sette, Enrico, 2016. Double bank runs and liquidity risk management. J. Financ. Econ. 122 (1), 135–154.

Iverson, Benjamin, 2017. Get in line: Chapter 11 restructuring in crowded bankruptcy courts. Manag. Sci. https://doi.org/10.1287/mnsc.2017.2808.

Jappelli, Tullio, Pagano, Marco, Bianco, Magda, 2005. Courts and banks: effects of judicial enforcement on credit markets. J. Money Credit Bank. 37 (2), 223–244.

Krueger, Anne, Tornell, Aaron, 1999. The Role of Bank Restructuring in Recovering from Crises: Mexico 1995–98. NBER Working Paper 7042.

La Porta, Rafael, Lopez de Silanes, Florencio, Shleifer, Andrei, Vishny, Robert W., 1998. Law and finance. J. Polit. Econ. 106 (6), 1113–1155.

Mayer, Christopher, Morrison, Edward, Piskorski, Tomasz, Gupta, Arpit, 2014. Mortgage modification and strategic behavior: evidence from a legal settlement with countrywide. Am. Econ. Rev. 104 (9), 2830–2857.

Milgrom, Paul, Roberts, John, 1990. Rationalizability, learning, and equilibrium in games with strategic complementarities. Econometrica 58 (6), 1255–1277.

Morris, Stephen, Shin, Hyun Song, 1998. Unique equilibrium in a model of self-fulfilling currency attacks. Am. Econ. Rev. 88 (3), 587–597.

Morris, Stephen, Shin, Hyun Song, 2003. Global games: theory and applications. In: Proceedings of the Eighth World Congress of the Econometric Society.

Ponticelli, Jacopo, Alencar, Leonardo S., 2016. Court enforcement, bank loans and firm investment: evidence from a bankruptcy reform in Brazil. Q. J. Econ. 131 (3), 1365–1413.

Rodano, Giacomo, Serrano-Velarde, Nicolas, Tarantino, Emanuele, 2016. Bankruptcy law and bank financing. J. Financ. Econ. 120 (2), 363–382.

Sákovics, József, Steiner, Jakub, 2012. Who matters in coordination problems? Am. Econ. Rev. 102 (7), 3439–3461.

Schiantarelli, Fabio, Stacchini, Massimiliano, Strahan, Philip E., 2016. Bank Quality, Judicial Efficiency and Borrower Runs: Loan Repayment Delays in Italy. NBER Working Paper 22034.

Shleifer, Andrei, Vishny, Robert, 2011. Fire sales in finance and macroeconomics. J. Econ. Perspect. 25 (1), 29–48.
Townsend, Robert, 1979. Optimal contracts and competitive markets with costly state verification. J. Econ. Theory 21 (2), 265–293.
Vives, Xavier, 1990. Nash equilibrium with strategic complementarities. J. Math. Econ. 19 (3), 305–321.
Woo, David, 2000. Two Approaches to Resolving Nonperforming Assets During Financial Crises. IMF Working Paper.
Van Zandt, Timothy, Vives, Xavier, 2007. Monotone equilibria in Bayesian games of strategic complementarities. J. Econ. Theory 134 (1), 339–360.
Zhu, Shuang, Pace, R. Kelley, 2015. The influence of foreclosure delays on borrowers' default behavior. J. Money Credit Bank. 47 (6), 1205–1222.