NBER WORKING PAPER SERIES

WHO BEARS THE WELFARE COSTS OF MONOPOLY? THE CASE OF THE CREDIT CARD INDUSTRY

Kyle F. Herkenhoff
Gajendran Raveendranathan

Working Paper 26604
http://www.nber.org/papers/w26604

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2020, Revised May 2021

We thank Bob Adams, Matteo Benetton, Satyajit Chatterjee, Dean Corbae, Kyle Dempsey, Pablo D'Erasmo, Bob Hunt, Igor Livshits, Daniel Grodzicki, Aubhik Khan, Simon Mongey, Jaromir Nosal, Ned Prescott, Larry Santucci, Jim Schmitz, Julia Thomas, and Toni Whited for helpful comments. We thank seminar participants at the AEAs, Board of Governors, Bonn, Cleveland Fed, Philadelphia Fed, NBER-Household Finance, Minneapolis Fed, McMaster University, MIT Sloan, Notre Dame, Ohio State University, PUC Chile, Queen's University, Stanford-SITE, University of Toronto, and the University of Saskatchewan. We thank Jacob Adenbaum for excellent research assistance. The authors thank the National Science Foundation (Award No. 1824422) and Washington Center for Equitable Growth for funding. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York, the Federal Reserve System, or the National Bureau of Economic Research.

Who Bears the Welfare Costs of Monopoly? The Case of the Credit Card Industry
Kyle F. Herkenhoff and Gajendran Raveendranathan
NBER Working Paper No. 26604
January 2020, Revised May 2021
JEL No. D14,D43,D60,E21,E44,G21

ABSTRACT

We measure the distribution of welfare losses from non-competitive behavior in the U.S. credit card industry during the 1970s and 1980s. The early credit card industry was characterized by regional monopolies that excluded competition. Several landmark court cases led the industry to adopt competitive reforms that resulted in greater, but still limited, oligopolistic competition. We measure the distributional consequences of these reforms by developing and estimating a heterogeneous agent, defaultable debt framework with oligopolistic lenders. Welfare gains from greater lender entry in the late 1970s are equivalent to a one-time transfer worth $3,400 (in 2016 dollars) for the bottom decile of earners (roughly 50% of their annual income) versus $900 for the top decile of earners. As the credit market expands, low-income households benefit more since they rely disproportionately on credit to smooth consumption. We find that greater lender entry resulting from these reforms delivers 65% of the potential gains from competitive pricing.

Kyle F. Herkenhoff
University of Minnesota
Department of Economics
4-101 Hanson Hall
1925 Fourth Street South
Minneapolis, MN 55455
and IZA
and also NBER
kfh@umn.edu

Gajendran Raveendranathan
Department of Economics
MacMaster University
1280 Main Street West
Hamilton Ontario L8S 4M4
Canada
raveeg1@mcmaster.ca

1 Introduction

The long-standing view that the deadweight losses from monopoly are small – for example, Harberger (1954)'s study of American manufacturing in the 1920s – has been refuted repeatedly (see a summary of arguments by Schmitz (2012)). Recent work by Schmitz (2016) argues that across a number of industries, the costs of monopoly are large and borne by low-income households. We contribute to this literature by quantifying the welfare costs of monopoly in the credit card industry. We first document that the U.S. credit card industry during the 1970s and early 1980s was regionally concentrated, engaged in non-competitive behavior such as blocking entrants and colluding on interest rates, and charged interest rates that significantly exceeded those implied by zero-profits. Because of repeated lawsuits for antitrust violations and the 1978 *Marquette* decision, the industry transitioned to an oligopoly with increased competition in the 1980s. We analyze the distributional consequences of these reforms by integrating oligopolistic lenders into a Bewley-Huggett-Aiyagari framework with default. We find that increased lender entry during the late 1970s and early 1980s disproportionately benefited low-income households. Among those in the lowest decile of earnings, welfare gains from lender entry are equivalent to a one time transfer worth 50% of annual earnings. The reform is worth four times less to the top decile of earners. Further, the reforms that took place in the 1970s and early 1980s led to large aggregate gains that achieved roughly 65% of potential gains from competitive pricing.

We begin by documenting several features of the U.S. credit card industry that relate to market power and competitive reforms during the 1970s and 1980s. First, the credit card industry was characterized by regional monopolies during the early 1960s and 1970s. The credit card networks (e.g., Visa and Mastercard) were regionally concentrated and owned by a collections of banks. Second, the industry engaged in openly non-competitive behavior that included blocking entry. Their use of exclusive contracts (e.g., Visa member banks could not issue Mastercard credit cards) and their explicit prevention of lower cost competitor cards from using their networks (e.g., the *MountainWest* case, which we discuss in Section 2) were challenged in courts and ultimately led to reforms. Moreover, the 1978 *Marquette* decision generated higher inter-regional competition among lenders by removing the influence of state usury laws, and the industry transitioned to oligopolistic competition in the 1980s. Third, there is evidence that lenders colluded on interest rates in the late 1970s and early 1980s. We review existing evidence of tacit collusion (Knittel and Stango (2003)), and then we provide new micro data on the interest rate setting behavior of banks. Between 1975 and 1982, roughly 50% of banks charged a common interest rate of 18%, which was also their rate cap, and there was almost no interest rate dispersion. Fourth, we combine our new micro data with a novel identification strategy to document near-zero pass-through from surprise monetary rate cuts to credit card interest rates. Fifth, even after adjusting for default risk and operational costs, credit card issuers charged interest rates that exceeded zero-profit interest rates by 3 to 9 percentage points.

To measure the distributional consequences of competitive reforms that emerged from court rulings on exclusivity, competitor entry, and regional usury laws, we depart from existing consumer credit models. These models typically assume competitive, zero-profit lenders (e.g., Livshits, MacGee, and Tertilt

(2007, 2010) and Chatterjee, Corbae, Nakajima, and Ríos-Rull (2007)). We develop a Bewley-Huggett-Aiyagari framework in which a finite number of non-atomistic credit card firms strategically compete for customers. Our model features defaultable debt and non-exclusive credit lines, complementing the work of Bizer and DeMarzo (1992), Hatchondo and Martinez (2018), and Kovrijnykh, Livshits, and Zetlin-Jones (2019).

Our pre-reform benchmark economy features a single monopolist lender with limited ability to price-discriminate. This assumption is motivated by Livshits, MacGee, and Tertilt (2016), who discuss the inability of lenders to discriminate in the 1980s. While our model does not explicitly feature geographic regions, we view this calibration as capturing the key features of the regional monopolies in the 1960s and early 1970s. After estimating our model to match key credit card market moments, we show that the benchmark non-competitive model accounts for roughly 40% of the observed excess spreads in the credit card industry in the 1970s. Our model also does well in replicating the limited evidence available on the disparate access to credit among low- and high-earning households.

We first measure the welfare gains from replacing a monopoly lender with a duopoly. This experiment is designed to capture increasing – yet still limited – competition among credit card issuers following the abandonment of exclusivity rules (e.g., following the *Worthen* case, which we discuss in the next section, Visa member banks were able to issue Mastercard credit cards). Based on existing evidence of collusion and our new evidence on the extreme lack of price dispersion, we view the 1970s and early 1980s as a period in which lenders colluded on interest rates but competed by issuing credit cards (i.e., lenders competed on credit limits). We call this form of competition *collusive-Cournot*, and we model it as a two-stage game. In the first stage, forward-looking lenders collude on interest rates; in the second stage they Cournot compete on limits. Our primary measure of welfare is Wealth Equivalent Variation (WEV). This is the one-time asset transfer that makes an individual in the monopoly environment indifferent to the duopoly environment. WEV provides two advantages over the traditional Consumption Equivalent Variation (CEV): (i) since WEV is in terms of dollars, it provides simple aggregation; and (ii) agents are allowed to reoptimize upon receipt of the one-time transfer. Nonetheless, we also provide CEV for all of our experiments. Among those in the lowest earnings decile, this reform is equivalent to a one-time transfer worth $1,200 (in 2016 dollars), or 18% of their annual income. The poor gain from improved consumption smoothing mostly due to higher limits and partially due to lower spreads. Those in the top earnings decile also gain for similar reasons, but by roughly three times less than the bottom decile. This reflects the disproportionate burden of monopoly power borne by low-income households, providing further support to Schmitz (2016). More generally, the reform is worth .33% of lifetime consumption for a newborn individual with zero assets. Further, the transition from monopoly to collusive-Cournot duopoly is a Pareto improvement, although lender profits distributed to households decrease along the transition path.

The implications of more competition are higher credit, higher defaults, a higher charge-off rate, and a lower excess spread (spread above the zero-profit interest rate). Moving from monopoly to collusive-Cournot duopoly increases aggregate credit to income from 1.02% to 1.49%, an increase of nearly 50%. With more indebted consumers, our model predicts that the default rate doubles from 0.06% per annum per capita to 0.10% per annum per capita. Likewise, charge-offs increase by 40%, from 0.98% to 1.41%. Consequently, the excess spread falls by 35%. While not the focus of this paper, these time trends are qualitatively consistent with observed time trends in aggregate data – for, example, those in Livshits et al. (2010). We view this as a rationale complementary to that in Athreya, Tam, and Young (2012), Livshits et al. (2016), and Sánchez (2018), who argue that technological changes in credit scoring are important in accounting for these trends during this time period.

Our second exercise is designed to capture gains from greater entry resulting from the 1978 *Marquette* decision. We measure the welfare gains from replacing a monopoly lender with a collusive-Cournot oligopoly of twenty lenders. The number of lenders is motivated by the inverse Herfindahl in the late 1980s, which implies a degree of concentration equivalent to roughly 20 equally sized banks. Among the lowest 10% of earners, this reform is equivalent to a one-time transfer worth $3,400 (in 2016 dollars), or roughly 50% of their annual income. Among the top 10% of earners, the reform is worth roughly $900 (in 2016 dollars). In comparison with high income households, the gains are roughly four times larger for low-income households. The transition from monopoly to collusive-Cournot oligopoly with twenty lenders generates significantly larger gains than the transition from monopoly to collusive-Cournot duopoly. The gains are almost three times higher for low-income households. Most of these gains are a result of an increase of almost 90% in credit limits. Spreads fall; however, the magnitudes are relatively small. This is because lenders maintain collusion on interest rates.

Our third exercise is designed to measure gains from price competition in the credit card market after the period of interest rate collusion in the early 1980s. In particular, we compute the distribution of welfare gains from replacing a monopoly with a duopoly that competes on prices. In our model with price competition, unlike the one with collusive-Cournot duopoly, consumers have a preference to borrow first from the lender that offers the credit card with the lower interest rate. This leads to debt dilution for this lender. Our assumption of non-exclusivity means that consumers may borrow from both lenders. For example, the lender that offers the lowest rate (Lender 1) understands that some of its consumers will subsequently borrow more from the other lender (Lender 2). By borrowing more from Lender 2, the consumer's default risk rises on *all* loans (for both Lender 1 and Lender 2), but Lender 1 is not compensated by Lender 2 for the additional default risk (*debt dilution*). Thus, the forward-looking Lender 1 restricts the amount of credit they extend. As a result, while welfare gains are still positive, moving from a monopoly to a duopoly with price competition yields lower welfare gains than moving to a collusive-Cournot duopoly. This suggests that while consumers may have benefited from price competition in the period after interest rate collusion, they may have also been hurt by restricted lending due to debt dilution.

Our fourth and final exercise is a counterfactual exercise that computes the welfare gains from forcing a single lender monopoly to price competitively and earn zero profits. This exercise serves two purposes. First, it provides an upper bound for the welfare costs borne by households. Second, it acts as a benchmark against which we can measure the relative effectiveness of the previous competitive reforms. The upper bound for welfare costs borne by the lowest 10% of earners is equivalent to a one-time transfer worth nearly $4,700 (in 2016 dollars), or roughly 70% of annual income. These households benefit from lower interest rates, which fall by 50%, and higher credit limits, which double. Relative to the top 10% of earners, the bottom 10% have WEV welfare gains from competitive pricing that are three times larger. This further substantiates the point that low-income households disproportionately suffer welfare losses from monopolistic credit card pricing. Comparing these welfare gains to the gains from lender entry discussed above, we find that collusive-Cournot duopoly generates 26% of the potential welfare gains from competitive pricing, whereas collusive-Cournot oligopoly with twenty lenders generates 65% of the potential welfare gains from competitive pricing. Therefore, despite lender collusion, the landmark legal decisions that led to greater entry in the late 1970s, such as the *Marquette* decision, achieved many of the potential gains from competitive pricing.

We next review the literature and then proceed as follows. Section 2 provides our narrative discussion and empirical analysis of the competitive structure of the U.S. credit card market. Section 3 lays out the model environment. Section 4 discusses our estimation of the model, and Section 5 provides the main results from our simulated reforms. Section 6 provides a number of robustness exercises, and Section 7 concludes.

Literature. Our paper is related to competitive consumer credit models (Livshits et al. (2007, 2010) and Chatterjee et al. (2007)) as well as recent models that generate lender market power via search and bargaining (e.g., Wasmer and Weil (2004), Drozd and Nosal (2008), Petrosky-Nadeau and Wasmer (2013), Bauducco and Janiak (2015), Galenianos and Nosal (2016), Herkenhoff (2019), and Raveendranathan (2020)). What makes the search models of the credit market tractable are the assumptions of atomistic lenders, free entry, and a small open economy. There is also a relatively new and innovative Industrial Organization literature that uses discrete choice frameworks to generate monopoly power among credit card lenders (Grodzicki (2015), Nelson (2018), and Galenianos and Gavazza (2019)). What makes Grodzicki (2015) and Nelson (2018) tractable is the partial equilibrium nature of the models and exogenous default. Our paper is closest to Galenianos and Gavazza (2019)'s which develops and estimates a static search model of lending that they use to study welfare implications of interest rate caps. We contribute to both of these literatures by developing a dynamic model of credit market oligopoly with defaultable debt.

Another class of models relates improvements in screening technology to greater credit limits and greater competition in the credit market (e.g., Livshits et al. (2016), Grodzicki (2019), and Sánchez (2018)). Whereas early contributions such as Ausubel (1991) document a lack of competition in the credit market throughout the 1970s and 1980s, Grodzicki (2019) makes a strong empirical argument that there has recently been an increase in competition in the credit market. Drozd and Nosal (2008) and Galenianos and Nosal (2016) also argue that reductions in entry costs – and thus increased competition – are quantitatively consistent with the rise in debt and defaults from the 1980s to the 1990s. In our framework, as additional oligopolists enter the credit market (i.e., as the market moves from monopoly and duopoly in the 1970s to oligopoly in the 2000s), our competitive structure generates increases in credit access and defaults that likely complement screening technology improvements.

Our paper relates to theoretic and quantitative models of credit lines (Drozd and Nosal (2008), Mateos-Planas and Seccia (2006), Mateos-Planas and Seccia (2013), Drozd and Serrano-Padial (2013), Drozd and Serrano-Padial (2017), Raveendranathan (2020), and Braxton, Herkenhoff, and Phillips (2018)). Drozd and Nosal (2008), Raveendranathan (2020), and Braxton et al. (2018) incorporate long-term credit lines into models with imperfect competition (via search and bargaining) in the credit market. Others, including Bizer and DeMarzo (1992), Hatchondo and Martinez (2018), and Kovrijnykh et al. (2019), consider borrowing without commitment. Closest to us are Hatchondo and Martinez (2018), who consider one-period loans without commitment. In these environments, the borrowing contracts resemble credit lines because the amount borrowed – capped by a limit – does not affect a consumer's interest rate. Our contribution to this literature is to incorporate non-exclusive credit lines into a Bewley-Huggett-Aiyagari economy with default.

Our model is also related to the quantitative banking literature (e.g., Corbae and D'erasmo (2011), Corbae and D'Erasmo (2019), Jamilov (2021)) and the partial equilibrium industrial organization banking literature (e.g., Egan, Hortaçsu, and Matvos (2017), Wang, Whited, Wu, and Xiao (2018), and Benetton (2018)). Of particular note, Corbae and D'Erasmo (2019) consider a Stackelberg Cournot bank that faces a competitive fringe. They build on the earlier work of Allen and Gale (2004) and Boyd and De Nicolo (2005), which considers Cournot competition. Lastly, our model contributes to a recently growing macroeconomic literature that attempts to quantify the welfare consequences of market power and strategic interactions (e.g., Mongey (2019), Edmond, Midrigan, and Xu (2018), Baqaee and Farhi (2017), and Berger, Herkenhoff, and Mongey (2019)).

2 Competition in the U.S. credit card industry

In this section, we document several characteristics related to market power and competitive reforms in the credit card industry during the 1970s and 1980s: (i) the credit card industry was highly concentrated and characterized by regional monopolies in the 1960s, after which the industry transitioned to an oligopoly in the 1980s; (ii) the industry blocked entrants, for which it was sued repeatedly, and colluded on interest rates; (iii) using new quarterly bank-level data on interest rate setting, we show that the majority of banks reported near-zero interest rate dispersion across and within their credit card plans; (iv) we document extreme price rigidity and zero pass-through to credit card interest rates in a period of large surprise movements in the cost of funds; and (v) even after adjusting for default risk and operational costs, credit card issuers charged interest rates that exceeded zero-profit interest rates.

2.1 Market concentration

The first defining feature of the credit card market is high concentration. We begin by with a narrative history of the early credit card industry based largely on Evans and Schmalensee (2005) and Wildfang and Marth (2005), and then we turn to call reports to document concentration in the 1980s.

During its beginning in the late 1950s and early 1960s, the credit card industry was characterized by regional monopolies. Scholars who have studied the early years of the credit card industry describe a highly non-competitive (and very "cooperative") environment. In particular, Evans and Schmalensee (2005) document that the early years of the credit card industry were characterized by limited competition. Visa's predecessor, National BankAmericard Inc., was founded in California in 1958 and only began franchising the program to other banks in 1966. Part of the franchise agreement was the restriction that banks could only issue Visa cards (a rule we call *exclusivity*). This severely limited competition and prompted several antitrust suits in the early 1970s (see below). American Express, which up to that point had been only a travel and entertainment card, began its own credit card program via franchising. The Interbank Card Association, which was the predecessor to Mastercard, was a "cooperative" of banks that began expanding around the same time period. Nonetheless, the programs were regionally concentrated: American Express was concentrated in New York, New England, New Jersey, and Pennsylvania; BankAmericard was concentrated in the West; and the Interbank Card Association was concentrated in the Midwest (Evans and Schmalensee (2005), p. 62-63).

Unfortunately, no systematic data on regional market concentration for credit card receivables exist for the 1970s or 1980s. Instead, we use bank call reports to describe concentration in the post-reform 1980s. To measure market concentration of credit card issuers (e.g., banks such as Citigroup, JP Morgan, Capital One, and Bank of America), we use their share of credit card receivables. Table 1 shows that the top 10 card issuers accounted for 45% of outstanding revolving credit in 1984-Q1, the earliest quarter for which we have credit card receivables data. The top 20 card issuers accounted for 56% of outstanding revolving credit. The maximal Herfindahl–Hirschman Index (HHI) during the 1980s occurs in 1988-Q4, when the HHI for credit card receivables reaches 0.0494. This HHI implies a degree of competition equivalent to 20.2 [=1/.0494] equally sized banks.[1] There are two caveats to the interpretation of the credit card receivables Herfindahl. First, since the banks owned the networks during this period, whether the effective number of competitors in the credit card market is 20 issuing banks or a smaller number of collusive issuer-payment networks is ambiguous. Second, we lack credit card receivable shares by region. Therefore, we view Table 1 as a conservative representation of market concentration in the 1980s.

2.2 Entry barriers and collusion

The second defining feature of the credit card market in the 1970s and 1980s was legal scrutiny over non-competitive behavior. We discuss four landmark court cases regarding (i) the lack of competition of networks across regions and banks' inability to issue competing network cards; (ii) network rules preventing competitive new-entrant credit cards from being issued within particular regions; (iii) the *Marquette* decision which generated further inter-regional competition, and (iv) interest rate collusion among issuing banks. Lastly, we discuss academic evidence of interest rate collusion during the 1980s by Knittel and Stango (2003).

We first discuss the *Worthen* and *MountainWest* cases based on analysis by Wildfang and Marth (2005).

[1]Here we use the common interpretation of the inverse Herfindahl–Hirschman Index. If all banks were equally sized $1/HHI = 1/(N \times (1/N)^2) = N$.

Table 1: Revolving credit share by issuer, 1984-Q1 (source: Call Reports)

	Share
Citibank	12.70%
Seattle-First National Bank	9.33%
First Chicago International	4.93%
Bank of Hancock City	4.53%
Chase Manhattan Bank	3.76%
First Interstate Bank	2.73%
Westminster Bank	2.20%
Wells Fargo Bank	1.80%
Security Pacific	1.67%
Hong Kong Bank	1.58%
Top 10 Share	45.2%
Top 20 Share	56.0%

Notes: 1984 Quarter 1, Consolidated Report of Condition for a Bank and its Domestic and Foreign Subsidiaries, nominal credit card receivables (*RCFD2008*) aggregated to highest holding company.

The *Worthen* case establishes that credit card networks effectively blocked entry and banned issuance of competing cards. In 1970, Worthen Bank was a member of Visa (previously the National BankAmericard network). Worthen wanted to issue Mastercard (previously the Interbank Card Association network) credit cards. Visa rules prohibited Worthen from doing so. Worthen sued over this *exclusivity* rule (the prohibition of issuing both networks cards), arguing that banning the bank from issuing other credit cards was the strongest form of anti-competitive behavior described in the Sherman Act (*per se illegality*). The federal district court agreed with Worthen. After an appeal and review by the Department of Justice, Visa abandoned *exclusivity* rules and by 1976 there were no barriers preventing banks from dually issuing Visa and Mastercard credit cards.

The second important case is the *MountainWest* case. In the 1980s, Sears attempted to enter the credit card market. The company wanted to issue an aggressive *Prime Option* card on the Visa network. The Sears card would have no annual fee, thus a lower implicit interest rate and potentially higher total credit limits. Although Sears owned MountainWest Financial, a Visa member bank, Visa prohibited Sears from issuing the card. Visa adopted a new rule banning Visa membership to any institution that was "deemed competitive."[2] As Wildfang and Marth (2005) summarize, "Sears claimed the rule was designed to, and did, exclude an aggressive price-discounting new entrant, which would have benefited consumers" (p. 682). Sears was initially affirmed in court, but the judgment was then reversed on appeal by the Tenth Circuit. According to Wildfang and Marth (2005), legal scholars believe the reversal was based largely on a misunderstanding of the facts by the jury. At the time of the trial, several news outlets interviewed Visa

[2] Wildfang and Marth (2005) write: "Visa then adopted a new rule that prohibited membership in Visa to any financial institution that issued or was affiliated with an institution that issued Discover cards or American Express cards or any other card "deemed competitive" by the Visa board of directors" (p. 682).

spokespeople, who acknowledged that allowing Sears to offer lower-fee cards would benefit consumers. The New York Times News Service (1991) reports, "David Brancoli, a spokesman for Visa, based in San Mateo, Calif., said the bank association opposed the new Sears card. He said that though competition among Visa issuers would benefit consumers, consumers would benefit even more from competition among different brands of cards..." The New York Times News Service (1991) further quoted industry analyst Allen R. DeCottiis as saying, "Visa banks are extremely concerned. They paid to build the Visa infrastructure, and now others are allowed access."

In addition to the *Worthen* and *MountainWest* cases, which led to greater competition, one particularly prominent competitive reform occurred in 1978 when the Supreme Court unanimously determined that state-level usury laws (interest rate caps) were not legally binding for nationally chartered banks. This decision was the result of *Marquette Nat. Bank of Minneapolis v. First of Omaha Service Corp.* (1978). Before the *Marquette* decision, states with relatively tight usury laws faced limited competition. After the *Marquette* decision, these credit card markets became nationally contested. It facilitated the opening of inter-regional competition among lenders. In our main counterfactuals that follow, we interpret the *Marquette* decision, as well as reforms spurred by the *Worthen* and *MountainWest* cases, as increased competition from lender entry.

In the late 1980s, accusations of interest rate collusion among issuing banks were widespread. Wells Fargo and First Interstate Bank of California were sued by the California Attorney General for interest rate fixing on millions of credit cards (White (1992)). The two banks ultimately settled the case for $55 million, fearing potentially large losses in court. While the case did not go to court, academic research by Knittel and Stango (2003) provides strong evidence for interest rate collusion in the early 1980s. Using disaggregated interest rate data from historical editions of the *Quarterly Report of Interest Rates on Selected Direct Consumer Installment Loans* from 1979 to 1989, Knittel and Stango (2003) show that the average spread between credit card interest rates and the cost of funds was higher in states where firms faced relatively tight interest ceilings and lower in states with no ceilings.[3] They argue that firms colluded on interest rate ceilings up until the late 1980s. These findings will motivate our competitive structure during the 1980s, in which lenders collude on interest rates but compete on credit card limits.

2.3 Interest rate dispersion and pass-through

The third and fourth defining features of the credit card market are the lack of interest rate dispersion and pass-through from changes in the costs of funds to offered interest rates. To complement the existing legal and academic evidence of non-competitive interest rate setting behavior (e.g., Ausubel (1991) and Knittel and Stango (2003)), we analyze pass-through rates and interest rate dispersion using new quarterly bank-level data between 1975 and 1982.[4] These data are a digitized archive of *Interest Rates Charged on Selected Types of Loans* (Form *FR 2835* and its variants), created and hosted at the Board of Governors. The data include roughly 200 banks in the panel each year, resulting in 4,600 bank-quarter observations

[3] The interest rate data in Knittel and Stango (2003) correspond to the "most common" interest rate charged by banks.

[4] Note 1982 is the date that the original LIRS credit card survey was discontinued; thus, our data set ends in 1982.

spanning the seven-year period from 1975 to 1982. The data set records the lowest, highest, and most common charged interest rates on each bank's credit card plan. We refer to the data set as *LIRS*, which is the data set's abbreviation within the Federal Reserve System.

We first establish several facts on interest rate dispersion and rigidity that, to our knowledge, are new to the literature. Table 2 reports summary statistics for the pooled LIRS data set from 1975 to 1982 (see Appendix A for detailed summary statistics by year). One of the key takeaways from Table 2 is the lack of dispersion in interest rates across and within banks. The majority of banks report a most common interest rate charged of 18%. For example, the 25th and the 75th percentiles for the most common charged interest rate is 18%. These interest rates provide further support to Knittel and Stango (2003)'s hypothesis. They argue that most issuers colluded on an interest rate of 18%, which was the relevant rate cap for roughly 80% of issuers. Further, we document below that these binding caps were extremely profitable. To measure interest rate dispersion within banks, we compute the highest charged interest rate minus the lowest charged interest rate. In Table 2, we find that 68% of lenders report no interest rate dispersion, and 54% of lenders report a highest and lowest interest rate equal to 18%. The 75th percentile of interest dispersion is a mere 3 percentage points.

The lack of dispersion is further illustrated in Panels (a), (b), and (c) of Figure 1. Panel (a) plots the histogram of interest rates in 1975. While most lenders charged an interest rate of 18%, there is smaller bunching at interest rates of 15% and 12%. A minority of lenders charged rates below 10%. Panel (b) plots the histogram of the dispersion in interest rates (measured as the highest minus lowest interest rate charged by a bank) in 1975, 1978, and 1982. Upwards of 60% of lenders report no interest rate dispersion across their plans; that is, their highest charged credit card interest rate is equal to their lowest one. Panel (c) plots the fraction of lenders that report no interest rate dispersion and the fraction of lenders that charge an interest rate of 18% to all customers from 1975 to 1982. By 1981, over 60% of issuers charged an interest rate of 18% to all customers, and more than 70% of banks reported zero interest rate dispersion. Bunching worsens and interest rate dispersion declines over time despite major deregulation (e.g., the *Marquette* decision discussed above) and the adoption of new technologies.

We also observe extreme price rigidity in this period. As Table 2 shows, only 9% of banks change their lowest charged interest rate from quarter to quarter (i.e., on average, they adjust their lowest rate once every 2.7 years). Likewise, 4% change their highest rate from quarter to quarter (corresponding to an adjustment once every 6.25 years). While the non-responsiveness of the highest charged rate may reflect interest rate caps, the rigidity of the lowest charged rate in a period of large surprise monetary rate cuts is striking. We illustrate this in Panel (d) of Figure 1, which plots the times series of the lowest, highest, and the most commonly charged credit card interest rate, averaged across banks. We also overlay the Federal Reserve Board's historic interest rate series (G series), which is based on the same underlying micro data but only available annually, and the Romer and Romer (2004) monetary surprise series, updated by Wieland and Yang (2020). [5] It is clear that interest rates are not responsive to movements in the cost of

[5]The G series data are available here: `http://www.federalreserve.gov/publications/other-reports/credit-card-profitability-2012-recent-trends-in-credit-card-pricing.htm`. We include all banks that report non-missing, non-negative, and non-zero credit card interest rates. The Romer and Romer (2004) series consists of the residuals

Table 2: LIRS Summary Statistics

Variable	Mean	p25	p50	p75	SD	Min	Max
Lowest Interest Rate Charged	15.9	12	18	18	3.0	6	25.92
Highest Interest Rate Charged	17.5	18	18	18	1.8	10	30
Most Common Interest Rate Charged	17.4	18	18	18	1.9	9	25.92
Highest Minus Lowest Rate Charged	1.7	0	0	3	2.7	0	13
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 18%	0.54						
Fraction of Banks Reported No Difference Between Highest and Lowest Charged Rate	0.68						
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 15%	0.04						
Fraction of Banks Reporting No Change to Lowest Charged Interest Rate	0.41						
Quarterly Probability That Lowest Charged Rate Changes	0.09						
Quarterly Probability That Highest Charged Rate Changes	0.04						
Quarterly Probability That Most Common Charged Rate Changes	0.05						
Number of Lender-Quarter Observations	4629						
Number of Banks in sample 1975	217						
Number of Banks in sample 1982	159						

Notes. Summary statistics for pooled LIRS data set between 1975 and 1982. We require non-missing and positive lagged interest rates to be in the sample. "Lowest Interest Rate Charged" corresponds to LIRS7812 question "Lowest interest rate charged for credit card plans," from FR 2835. "Highest Interest Rate Charged" corresponds to LIRS7813 question "Highest interest rate charged for credit card plans," from FR 2835. "Most Common Interest Rate Charged" corresponds to LIRS7814 question "Most common interest rate charged for credit card plans," from FR 2835. "Highest Minus Lowest Rate Charged" is LIRS7813 minus LIRS7812.

funds. In particular, the lowest charged interest rate does not move in response to the 4% surprise rate cut in 1980-II. We formally analyze pass-through rates for the remainder of this section.

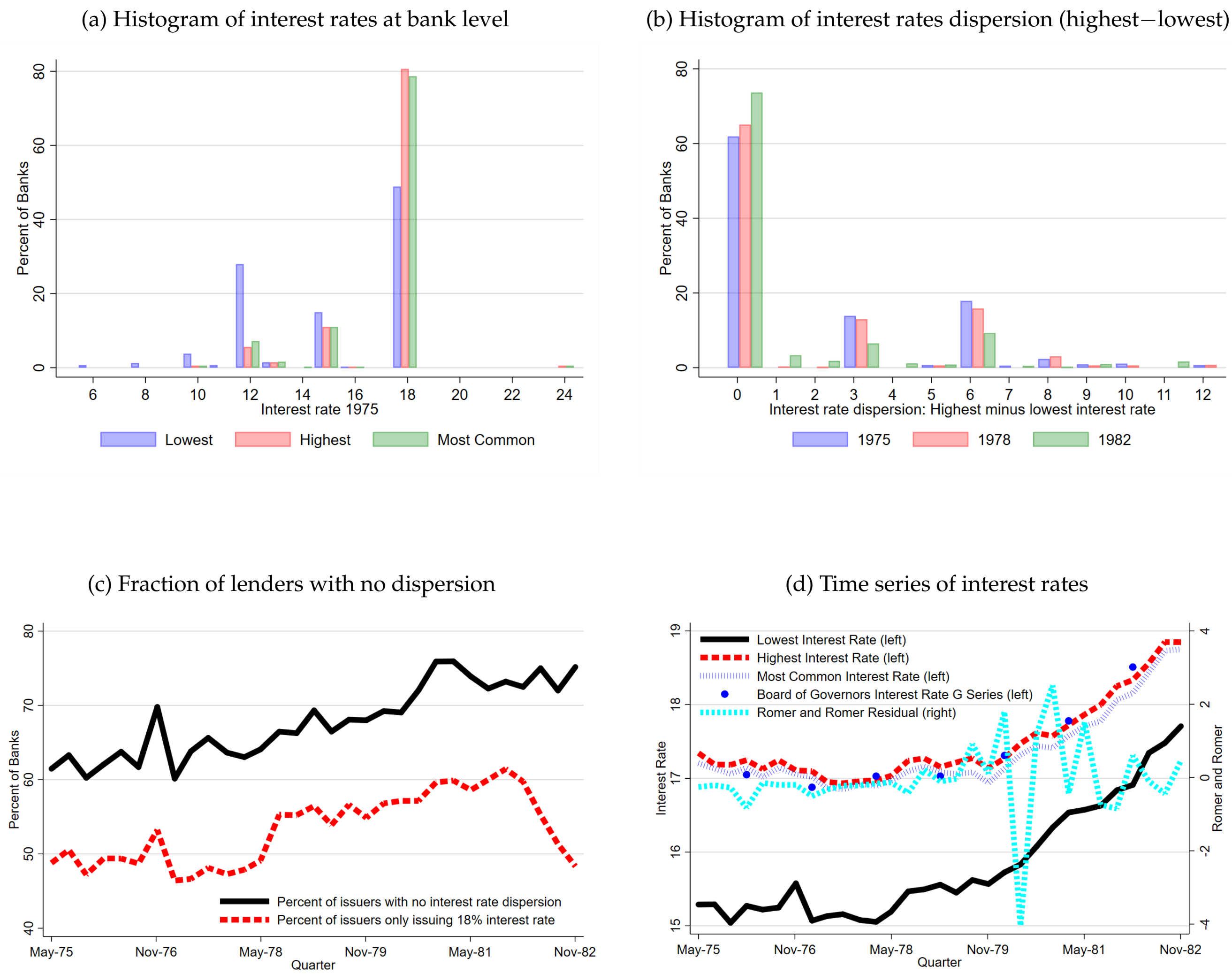
To formalize the graphical analysis in Panel (d) of Figure 1 and measure pass-through rates from surprise monetary rate cuts to credit card interest rates, we must circumvent the issue of binding interest rate caps. Our identification of pass-through addresses the issue of interest rate caps by isolating lenders whose lowest charged rate is less than the interest rate cap. Focusing on the lowest charged interest rate and restricting the sample to lenders that charge interest rates less than the cap allow us to obtain unbiased estimates of pass-through rates for two reasons. First, if the lowest rate charged is less than the interest rate cap, then the lender does not face a binding constraint for a subset of customers. Second, if the lender does not face a binding constraint, then in a competitive environment the lowest rate charged should decline following a surprise rate cut. Since there were sizable surprise rate cuts between 1975 and 1982, including one of over 4 percentage points, we would expect that in a competitive environment, issuers facing non-binding rate caps would decrease their lowest charged interest rates.

Let $i \in \mathcal{I}$ denote the set of lenders whose lowest charged interest rate is less than the cap in quarter $t-1$. Let $t \in \mathcal{T}$ denote the set of quarters in which a surprise monetary rate cut occurs. Let Δr_{it}^{Low} denote the change in the lowest offered interest rate between quarter $t-1$ and t for bank i. To recover pass-through rates, we regress Δr_{it}^{Low} on the Romer and Romer (2004) monetary surprise series in quarter t (RR_t). It is important to note that since there are multiple rate cuts, we are able to identify bank fixed effects α_i and remove non-time-varying unobserved heterogeneity among lenders. We estimate following specifications:

$$\Delta r_{it}^{Low} = \alpha_i + \beta RR_t + \epsilon_{it} \qquad \forall i \in \mathcal{I}, t \in \mathcal{T}. \tag{1}$$

from a regression of the federal funds rate on lagged Greenbook forecasts as a proxy for Federal Reserve's information set.

Figure 1: Interest rates 1975 to 1982 (Data based on FR 2835)



In Table 3, we report estimates of β in equation (1). Both Δr_{it}^{Low} and RR_t are in percentage points, and so β is interpretable as a pass-through rate.

Columns (1) and (2) of Table 3 report pass-through rates using all time periods and all lenders. The estimate in Column (1) implies a pass-through rate of 2.4% from the monetary surprise (measured in percentage points) to the change in the lowest charged interest rate (also measured in percentage points) with a relatively tight standard error. The upper bound of the 95% confidence interval is a pass-through rate of 6.1%. Column (2) adds in lender fixed effects, and the point estimate remains largely unchanged. Since the lowest rate should not necessarily rise in response to a surprise rate hike, even in a competitive environment, Column (3) restricts the sample to quarters in which there is a surprise rate *cut* of at least 15 basis points (i.e., $RR_t < -.15$). The pass-through rate remains largely unchanged, and the standard error remains relatively precise.

Table 3: Pass-through rates

	Dependent variable is Δr_{it}^{Low} (ppt), lowest interest rate charged in quarter t minus $t-1$							
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Monetary Surprise (ppt), RR_t	0.0244*	0.0239*	0.0182	0.0145	0.0599	0.0151	0.0181	0.0331
	(0.0141)	(0.0137)	(0.0145)	(0.0373)	(0.0451)	(0.0166)	(0.0255)	(0.0326)
Sample	Full	Full	Rate Cuts of At Least 15 Basis Points, All Banks	Rate Cuts of At Least 15 Basis Points, Banks with Lowest Rate Strictly Less Than 18%	Rate Cuts of At Least 15 Basis Points, Banks with Lowest Rate Strictly Less Than 15%	Rate Cuts of At Least 25 Basis Points, All Banks	Rate Cuts of At Least 25 Basis Points, Banks with Lowest Rate Strictly Less Than 18%	Rate Cuts of At Least 25 Basis Points, Banks with Lowest Rate Strictly Less Than 15%
Lender Fixed Effect	N	Y	Y	Y	Y	Y	Y	Y
Observations	4,629	4,593	2,405	1,088	759	1,494	613	420
R-squared	0.000	0.014	0.040	0.150	0.165	0.094	0.194	0.247

*Notes. LIRS data series, quarterly from 1975-I to 1982-IV. Standard errors clustered by quarter in parentheses, *** p<0.01, ** p<0.05, * p<0.1.*

As discussed above, there are several reasons for a low pass-through rate, including the presence of interest rate caps. In Columns (4) and (5), we rerun these regressions in samples of lenders that report that their lowest charged interest rate is below 18% ($r_{it-1}^{Low} < 18.0$) or below 15% ($r_{it-1}^{Low} < 15.0$), respectively, and thus are not facing binding interest rate caps on at least a subset of their credit cards. We continue to restrict the sample to quarters in which there is a surprise rate cut of at least 15 basis points (i.e., $RR_t < -.15$).

In Column (4), the pass-through rate from the monetary surprise (measured in percentage points) to the change in the lowest charged interest rate (also measured in percentage points) is 1.5%. While the standard error is wider, the upper bound of the 95% confidence interval for the estimated pass-through rate is 8.8%. When we restrict the set of lenders to be those with rates strictly less than 15% ($r_{it-1}^{Low} < 15.0$) in Column (5), the pass-through rate rises to approximately 6%. The upper bound of the 95% confidence interval in Column (5) is a pass-through rate of 14.8%.

In Columns (6) through (8), we repeat the same exercise using surprise rate cuts of at least 25 basis points (i.e., $RR_t < -.25$) . We obtain very similar results. The largest 95% confidence interval across all columns implies a pass-through rate of 9.7% (Column (8)). The very low degree of pass-through is robust to clustering at the bank level, differencing the outcome variable between quarters $t-1$ and $t+1$, alternate surprise rate cut thresholds, and removing fixed effects.

There are four potential explanations for the lack of interest rate pass-through: (1) interest rate caps; (2) technology; (3) risk sharing (lender commitment/household insurance); (4) imperfect competition. Our identification strategy rules out interest rate caps. Lack of technology such as credit scoring is an important determinant of price *dispersion*. However, it is unlikely that technology restricts the overall *level* of interest rates being charged to customers. For example, suppose the bank does not have the technology to price discriminate and charges one interest rate of 14%. Then, the cost of funds falls by 4%. In this example, even though the bank does not have credit scoring technology, we would still expect the bank to charge a lower rate of 10% in a competitive environment. We do not see this behavior in the data. Without systematic evidence on the length of credit relationships, risk sharing and imperfect competition are difficult to distinguish. In the model exercise that follows, we assume both forces are present (commitment and imperfect competition), and it is the ability of lenders to commit to credit lines that implies short-term pass-through rates of zero.

Our analysis makes three conceptual improvements upon existing estimates of pass-through in Ausubel (1991). First, Ausubel (1991) ignores interest rate caps and simply regresses the "most common" interest rate on lagged federal funds rates. Ignoring interest rate caps biases the pass-through rates towards zero and does not necessarily reflect a lack of competition. Our identification strategy relies on analysis of the lowest charged rate and circumvents the downward biases arising from interest rate caps by conditioning on banks that report offering non-binding interest rates. Second, Ausubel (1991) regresses interest rates on the raw federal funds rate, including forecastable components. Third, Ausubel (1991) uses an extremely small, non-representative sample of approximately 50 banks that responded to his own bank credit card survey (BCCS).

2.4 Excess spreads

Despite facing interest rate caps, the fifth defining feature of the credit card industry in the 1970s and 1980s is non-competitive interest rate spreads. To measure how far the credit card industry is from competitive pricing, we will focus on what we call *excess spreads*. We compute excess spreads as the difference between the *actual spread* and the *zero-profit spread*. The *actual spread* (τ_{actual}) is computed as the difference between the (cross-sectional) average credit card interest rate and the Moody's Aaa rate. The *zero-profit spread* is defined as the spread that credit card firms must charge on interest income to break even. Let τ_{zero} denote the zero-profit spread, D denote the charge-off rate, B denote outstanding revolving credit, r denote the Moody's Aaa rate, and τ_o denote the transaction cost net of non-interest income. We compute τ_o as follows: (operational cost + rewards and fraud - fee income - interchange income)/(outstanding revolving credit). Note that interchange income accrues to the issuing bank (e.g., Bank of America), not to the network (e.g., Visa).[6] Given D, B, r, and τ_o, we estimate the zero-profit spread from the following break even equation:

$$(1-D)B(1+r+\tau_{zero}) = B(1+r+\tau_o). \tag{2}$$

[6]Visa and Mastercard earn profits from *network fees* (also called *credit association fees*) that are typically 0.5% of transaction volume. Visa and Mastercard also set a separate fee called an *interchange fee*. Interchange fees are paid directly to the issuer banks and are typically equal to 1.5% to 3.0% of the transaction price. A common misconception is that these fees go to Visa and Mastercard because Visa and Mastercard set these fees. These interchange fees are tied to the generosity of the rewards program that the issuing banks choose. Cards that provide greater rewards can charge higher interchange fees. The fact that these fees, set by networks, scale with rewards suggests a lack of separation between networks and issuing banks. Since banks that choose to offer rewards can charge more interchange fees (which are borne by *merchants* who sell goods and accept credit card payments), reward cards do not yield lower profits to issuer banks (merchants are typically not allowed to price discriminate by card, although recent legal changes have relaxed these rules). Hence, by construction of the interchange fees, rewards do not lower the excess profits and excess spreads of issuers. See Hunt (2003) for more discussion.

The left hand side of (2) is total interest income net of charge-offs, and the right hand side is total cost net of non-interest income. Table 4 first presents the average excess spreads (1974-2016) under the assumption of zero transaction costs and zero non-interest income ($\tau_o = 0$). We then progressively add in the components of the transaction cost (e.g., operational costs, rewards and fraud) and components of the non-interest income (e.g., fee income, interchange income). We use estimates of the components of τ_o from Agarwal, Chomsisengphet, Mahoney, and Stroebel (2015), and in many instances, including our preferred specification, τ_o is negative. A negative τ_o implies that the credit card industry makes profits even if we ignore interest income.

Table 4 shows that the average spread on credit cards is 3.42 percentage points above breaking even if we ignore transaction costs net of non-interest income. If we include all components of non-interest income, the excess spread reaches 8.84 percentage points above breaking even. This implies a range of markups between 44% and 115% on the Moody's Aaa average rate from 1974 to 2016 (7.72%). Our preferred specification adjusts for all components of non-interest income.

Figure 2 graphically depicts the times series of excess spreads. Do excess spreads compensate for risk? Clearly, after adjusting for non-interest income (red line), excess spreads are positive every year from 1974 to 2016. Thus, compensation for risk is an unlikely explanation for the large deviation from break even spreads. While excess spreads remain significant in the 2000s, they have generally declined over time. Although other factors outside of the scope of this paper may simultaneously be increasing or decreasing spreads, this downward trend in spreads is consistent with increased competition in the credit card market.[7]

Table 4: Credit card industry excess spreads (source: authors' calculations, see text)

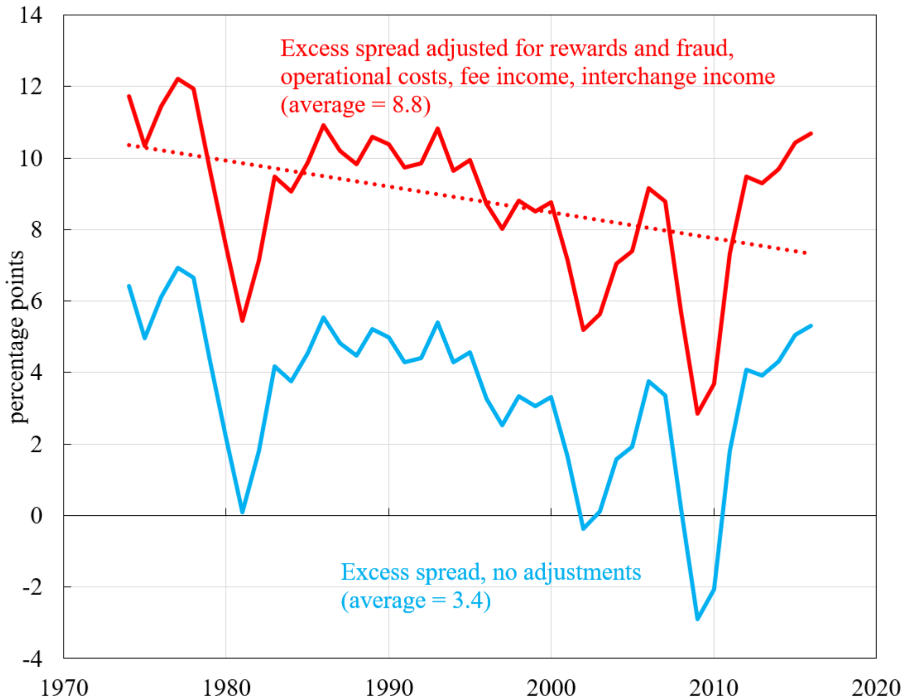
Excess spread definition	**Avg. 1974-2016 (percentage points)**
Excess spread: no adjustments	3.42
Excess spread: operational costs, interchange income	3.21
Excess spread: operational costs, fee income, interchange income	11.13
Excess spread: operational costs, fee income, interchange income, rewards and fraud	8.84

Notes. 'Excess spread: no adjustments' is defined to be $\tau_{actual} - \tau_{zero}$ where τ_{zero} is calculated from equation 2 using $\tau_o = 0.0$. 'Excess spread: operational costs, interchange income' is defined to be $\tau_{actual} - \tau_{zero}$ where τ_{zero} is calculated from equation 2 using $\tau_o = 0.002$. 'Excess spread: operational costs, fee income, interchange income' is computed similarly using $\tau_o = -0.076$. 'Excess spread: operational costs, fee income, interchange income, rewards and fraud' is computed similarly using $\tau_o = -0.052$. Operational costs, fee income, interchange income, rewards and fraud expenses as a fraction of average daily balances are taken from Agarwal et al. (2015) Table III, Column 1, Rows 9 to 16.

Summary. The U.S. credit card industry was characterized by regional monopolies in the 1950s and 1960s and concentration levels remained elevated throughout the 1980s. In the 1970s and 1980s, the

[7]See contemporaneous work by Dempsey and Ionescu (2019), who also document large excess spreads post-2012 in Y-14 micro regulatory data.

Figure 2: Credit card industry excess spreads (source: authors' calculations, see text)



Notes."Excess spread, no adjustments" is defined to be $\tau_{actual} - \tau_{zero}$ where τ_{zero} is calculated from equation 2 using $\tau_o = 0.0$."Excess spreads adjusted for rewards and fraud, operational costs, fee income, interchange income" is computed similarly using $\tau_o = -0.052$. Operational costs, fee income, interchange income, and rewards and fraud expenses as a fraction of average daily balances are taken from Agarwal et al. (2015) Table III, Column 1, Rows 9 to 16.

industry was sued repeatedly for antitrust violations such as blocking entrants, and there is evidence of interest rate collusion. The majority of banks reported zero interest rate dispersion across their credit card plans in 1975, and pass-through rates from surprise monetary rate cuts to the lowest offered credit card interest rate are a precisely estimated zero. Even after we adjust for default risk and operational costs, credit card issuers charged interest rates that exceeded zero-profit interest rates.

We use these facts to justify our departure from standard competitive models of the consumer credit market and, instead, model a finite number of non-atomistic credit card firms that issue non-exclusive credit lines. We use the model to quantify the welfare gains and losses from competitive reforms in the credit card industry. Based on the above narrative, we simulate (i) increased lender entry with price collusion, and (ii) price competition.

3 Model

Our model economy shares many elements with existing competitive models of consumer credit, in particular those of Livshits, MacGee, and Tertilt (2007) and Chatterjee, Corbae, Nakajima, and Ríos-

Rull (2007). We build on their work by integrating a lender oligopoly into a small open economy with heterogeneous consumers. We also depart from the existing literature by allowing lenders to issue non-exclusive long-term credit lines.[8]

3.1 Environment

Time is discrete and runs forever ($t = 0, 1, \ldots$). We assume a small open economy with risk-free rate r_f. For ease of exposition, we focus on a recursive exposition of the steady state, omitting time subscripts. However, when we compute transition paths in later sections of this paper, the value functions, policy functions, and profits are time dependent. The economy is populated by a unit measure of infinitely-lived heterogeneous consumers and N credit card firms (which we will also refer to as *lenders*). Consumers differ ex-ante with respect to their permanent earnings ability. They face idiosyncratic earnings shocks as well as taste shocks over their decision to default/repay. They make savings/borrowing and default/repayment decisions to maximize utility. Lenders imperfectly compete to issue non-exclusive credit lines. Lenders commit to their interest rates and limits, yielding zero pass-through from risk-free rates to credit card interest rates by assumption (see Section 2). Motivated by the limited dispersion in rates in the 1970s, we assume that lenders discriminate with respect to permanent earnings ability. This implies that all interest rates, limits, and lender profit functions are indexed by permanent earnings ability.

3.1.1 Consumers

Consumers have discount factor $\beta \in (0, 1)$. They make savings/borrowing and default/repayment decisions to maximize the present value of their flow utility over consumption, $U(c)$, as well as any utility gain or loss associated with default. There are three preference parameters associated with default. Consumers have independent and identically distributed Gumbel taste shocks over default and repayment $\zeta_R \sim^{iid} F(\zeta_R) = e^{-e^{-\kappa\zeta_R}}$ and $\zeta_D \sim^{iid} F(\zeta_D) = e^{-e^{-\kappa\zeta_D}}$, respectively (e.g., Auclert and Mitman (2018) and Chatterjee, Corbae, Dempsey, and Rios-Rull (2020)). The Gumbel scaling parameter κ is common for both shocks. We view these taste shocks as unmodeled sources of default such as divorce, health costs, and other lawsuits (Chakravarty and Rhee (1999)). If the consumer chooses to default, they incur an additional one-time utility penalty of χ (stigma).

The consumer's idiosyncratic state is given by their credit standing $i \in \{g, b\}$, permanent earnings ability $\theta \in \{\theta_L, \theta_H\} \equiv \Theta \subset \mathbb{R}_+$, a persistent earnings shock $\eta \in \mathbb{R}_+$, an iid earnings shock $\epsilon \in \mathbb{R}_+$, and net assets $a \in \mathbb{R}$. If the consumer is in good credit standing, then $i = g$, and the consumer may borrow. Otherwise, the consumer is in bad credit standing ($i = b$) and cannot borrow. Permanent earnings ability θ is fixed, and thus we refer to *type-θ consumers* when referencing permanent earnings ability. The earnings shock η is persistent and follows a Markov chain, whereas ϵ is perfectly transitory.

[8] Note that this is available in a previous version of this paper: `https://www.nber.org/system/files/working_papers/w26604/w26604.pdf`, we analyzed implications under general equilibrium. The general equilibrium effects were minimal. Hence, we assume a small open economy in this version.

Positive values of a indicate saving, whereas negative values of a indicate borrowing. The state of a consumer is therefore given by the tuple $(i, \theta, \eta, \epsilon, a)$. We define the distribution of consumers across states as $\Omega(i, \theta, \eta, \epsilon, a)$, where $\Omega : \{g, b\} \times \Theta \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \to [0, 1]$.

In order to exposit the consumer's problem, we must briefly discuss the credit card market (more details about the formation of the credit lines appear in Section 3.1.2). If a type-θ consumer chooses to borrow, they borrow from a set of credit lines $S_\theta = \{(r_1(\theta), \bar{l}_1(\theta)), \ldots, (r_N(\theta), \bar{l}_N(\theta))\} \in (\mathbb{R}_+, \mathbb{R}_+)^N$. A credit line is a long-term defaultable debt contract that specifies an interest rate $r \in \mathbb{R}_+$ and a borrowing limit $\bar{l} \in \mathbb{R}_+$. The collection of credit line interest rates and borrowing limits offered by the N lenders to type-θ consumers is denoted by S_θ.

In the 1970s, as discussed above, there was limited price discrimination, and hence all type-θ consumers in good credit standing have access to the same set of credit lines S_θ. Furthermore, credit lines are non-exclusive. If there are N credit lines available in equilibrium, the consumer will first borrow from the cheapest credit line, independent of the lender that issues it.[9] Let j denote the credit card interest rate ranking of a credit line, where $j = 1$ is the lowest credit card interest rate and $j = N$ is the highest credit card interest rate. Therefore, the credit lines can be sorted in ascending order with respect to the interest rate $(r_1(\theta) \leq r_2(\theta) \leq \ldots r_j(\theta) \leq \ldots \leq r_N(\theta))$ and the corresponding borrowing limits $(\bar{l}_1(\theta), \bar{l}_2(\theta), \ldots, \bar{l}_N(\theta))$, ignoring the issuing credit card firm's identity. With this notation, the set of credit lines available to type-θ consumers is $S_\theta = \{(r_1(\theta), \bar{l}_1(\theta)), \ldots, (r_N(\theta), \bar{l}_N(\theta))\} \in (\mathbb{R}_+, \mathbb{R}_+)^N$. For any net asset level a (recall $a < 0$ implies debt), let $a_j(a, \theta) \leq 0$ denote the balance on the credit line with credit card interest rate ranking $j \in \{1, ..., N\}$ for a type-θ consumer:

$$a_j(a, \theta) = \begin{cases} -\bar{l}_j(\theta) & \text{if} \quad a \leq -\sum_{k=1}^{j} \bar{l}_k(\theta) \\ \min\{a + \sum_{k=1}^{j} \bar{l}_k(\theta) - \bar{l}_j(\theta), 0\} & \text{if} \quad a > -\sum_{k=1}^{j} \bar{l}_k(\theta). \end{cases}$$

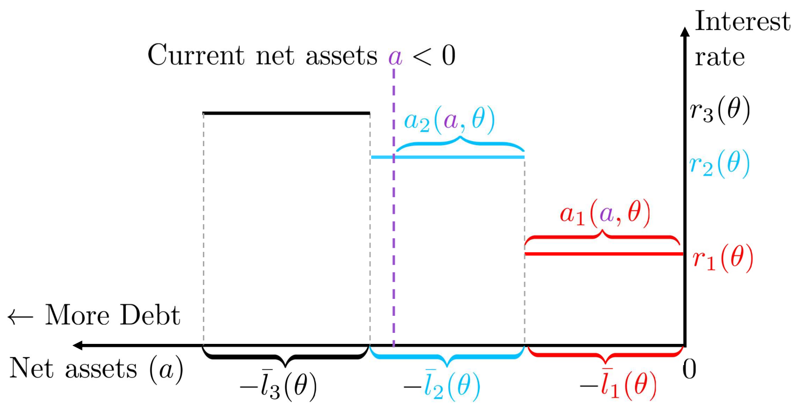
If net assets are less than or equal to the sum of the borrowing limits on credit lines $\{1, ..., j\}$, then the consumer has reached the limit on credit line j. Otherwise, if net assets are greater than the sum of the borrowing limits on credit lines $\{1, ..., j-1\}$ and net assets are negative, then the balance on credit line j is $a + \sum_{k=1}^{j} \bar{l}_k(\theta) - \bar{l}_j(\theta)$. Lastly, if net assets are positive, then the balance on credit line j (and all other credit lines) is zero. Figure 3 provides an example of the interest rates and limits consumers face with three lenders, $N = 3$. They borrow from the lowest interest rate first, $r_1(\theta)$, then the next lowest, $r_2(\theta)$, and lastly $r_3(\theta)$. The total principal and interest expense incurred on the lowest interest rate credit card is $(1 + r_1(\theta))a_1(a, \theta)$, the next lowest interest rate $(1 + r_2(\theta))a_2(a, \theta)$, and lastly, $(1 + r_3(\theta))a_3(a, \theta)$. More generally, since $\sum_{j=1}^{N} a_j(a, \theta) = a$, the principal and interest expense of a household can be written as $(1 + r_f)a + \sum_{j=1}^{N}(r_j(\theta) - r_f)a_j(a, \theta)$. When making contact with the data, it will be useful to define lender j's *spread*:

$$\tau_j(\theta) \equiv r_j(\theta) - r_f.$$

Using this notation for credit lines, we now describe the consumer's value functions. Let $V(i, \theta, \eta, \epsilon, a)$

[9] This is an equilibrium outcome in our model because there are no switching costs. This keeps the model tractable.

Figure 3: Example of credit lines available to consumer with three lenders, $N = 3$.



Notes: Non-calibrated example with three lenders and three credit lines. More negative net asset positions imply greater debt. The function $a_j(a, \theta)$ allocates net assets a most efficiently across the credit lines ordered by interest rates $r_j(\theta)$. Consumers first max out credit card 1 with the lowest interest rate $r_1(\theta)$. If they borrow more than $\bar{l}_1(\theta)$, they then begin borrowing from the credit card with the second lowest interest rate, $r_2(\theta)$, and so on.

denote the consumer's continuation value at the start of the period. Let $V^D(\theta, \eta, \epsilon)$ be the value of default and $V^R(i, \theta, \eta, \epsilon, a)$ be the value of repayment. The first choice the consumer makes is between default and repayment:

$$V(i, \theta, \eta, \epsilon, a) = E_{\zeta_D, \zeta_R} \max\{V^D(\theta, \eta, \epsilon) + \zeta_D, V^R(i, \theta, \eta, \epsilon, a) + \zeta_R\}. \tag{3}$$

Since ζ_D and ζ_R were assumed to be Gumbel with a common inverse scaling parameter κ, we can express the default probability as follows:

$$p(i, \theta, \eta, \epsilon, a) = \frac{exp(\kappa V^D(\theta, \eta, \epsilon))}{\exp(\kappa V^D(\theta, \eta, \epsilon)) + \exp(\kappa V^R(i, \theta, \eta, \epsilon, a))}. \tag{4}$$

Given our assumptions about default penalties, default is universal. That is, the consumer repays credit card debt on all credit lines or defaults on all credit lines. The policy functions for repayment/default, consumption, and savings/borrowing — $p(\cdot), c(\cdot), a'(\cdot)$ — are functions of $(i, \theta, \eta, \epsilon, a)$. However, for ease of exposition, we omit this state dependence of policy functions.

A consumer who defaults consumes labor earnings and profits, $\theta\eta\epsilon + \Pi$, where Π refers to the profits uniformly transferred to consumers from credit card firms (in Section 6, we consider alternate distributions of profits). Furthermore, the consumer cannot save or borrow ($a' = 0$) and incurs a one-time disutility cost (stigma χ) only during the default period. In the next period, the consumer may regain good credit standing with probability ϕ or stay in bad credit standing with probability $1 - \phi$. The continuation value of defaulting is given by

$$V^D(\theta,\eta,\epsilon) = U(\theta\eta\epsilon + \Pi) - \chi + \beta E_{\epsilon',\eta'|\eta}\left[\phi V(g,\theta,\eta',\epsilon',0) + (1-\phi)V(b,\theta,\eta',\epsilon',0)\right].$$

A type-θ consumer who chooses to repay and is in good credit standing $(i = g)$ may borrow from the set of credit lines or save $(a' \geq -\sum_{j=1}^{N} \bar{l}_j(\theta))$. Furthermore, this consumer retains good credit standing for the next period. The value of repayment when $i = g$ is given by

$$\begin{aligned} V^R(g,\theta,\eta,\epsilon,a) = & \max_{c,a'} \; U(c) + \beta E_{\epsilon'\eta'|\eta} V(g,\theta,\eta',\epsilon',a') \\ & s.t. \\ c + a' = & \theta\eta\epsilon + (1+r_f)a + \sum_{j=1}^{N}(r_j(\theta) - r_f)a_j(a,\theta) + \Pi \\ a' \geq & -\sum_{j=1}^{N} \bar{l}_j(\theta). \end{aligned} \tag{5}$$

Because of the taste shock for default, consumers in bad standing $(i = b)$ may redefault, a common occurrence in the data (Athreya, Mustre-del Río, and Sánchez (2019)). Because of the taste shocks, defaults may occur with a balance of zero net assets or greater. We interpret the data analogue of these defaults to be shocks that are not modeled explicitly in our framework, such as divorce, health shocks, or lawsuits. A consumer who chooses to repay and is in bad credit standing can only save $(a' \geq 0)$. Furthermore, the consumer regains good credit standing in the next period with probability ϕ and stays in bad credit standing with probability $1-\phi$. The value of repayment when $i = b$ is given by

$$\begin{aligned} V^R(b,\theta,\eta,\epsilon,a) = & \max_{c,a'} \; U(c) + \beta E_{\epsilon',\eta'|\eta}\left[\phi V(g,\theta,\eta',\epsilon',a') + (1-\phi)V(b,\theta,\eta',\epsilon',a')\right] \\ & s.t. \\ c + a' = & \theta\eta\epsilon + (1+r_f)a + \Pi \\ a' \geq & 0. \end{aligned} \tag{6}$$

Compared with the problem for the consumer in good standing (5), the budget constraint for those in bad standing drops the term $\sum_{j=1}^{N}(r_j(\theta) - r_f)a_j(a,\theta)$ because the consumer in bad credit standing cannot hold debt in equilibrium, regardless of the repayment choice.

3.1.2 Lenders

There are N lenders in the economy. We assume that lenders price discriminate across type-θ consumers. Furthermore, lenders observe the default status of individuals, $i \in \{g, b\}$, and they only issue credit lines to those in good standing.

Each lender may issue one credit line to each type-θ consumer. Therefore there are N credit lines for each consumer. We assume lenders commit to the terms of their lines of credit, consistent with the limited pass-through found in Section 2. Consider lender $k \in \{1, \dots, N\}$. We will use the convention that superscripted k refers to a lender's identity and does *not* reflect any ranking of lenders, and subscripted j refers to the lenders' credit card interest rate ranking. For a type-θ consumer, lender k's objective is to choose the terms of their credit line, $(\tau^k(\theta), \bar{l}^k(\theta))$, to maximize their net present value of profits, $\pi_t^k(\theta)$, discounted at rate r_f:

$$\sum_{t=0}^{\infty} \left(\frac{1}{1+r_f}\right)^t \pi_t^k(\theta). \tag{7}$$

For the rest of this section, we will omit time subscripts from the lender's problem and focus on steady states of the model economy. Since lenders commit to credit lines, the lender's steady state objective is equivalent to maximizing per-period profits $\pi^k(\theta)$.

A lender's credit card interest rate ranking is denoted by j, where $j = 1$ is the lowest credit card interest rate and $j = N$ is the highest one. Let $r_j(\theta)$ and $\bar{l}_j(\theta)$ denote the interest rate and borrowing limit of the lender offering the j^{th} highest credit card interest rate. The flow profits resulting from offering the j^{th} highest credit card interest rate are given by $\Pi_j(\theta)$:

$$\begin{aligned} \Pi_j(\theta) = \int \Big[& -(1 - p(g, \theta, \eta, \epsilon, a))(r_j(\theta) - r_f)a_j(a, \theta) \\ & + p(g, \theta, \eta, \epsilon, a)(1 + r_f)a_j(a, \theta)\Big] d\Omega(g, \theta, \eta, \epsilon, a) \end{aligned} \tag{8}$$

Lenders borrow from households and since households can costlessly access capital markets, the lenders must offer a riskless savings rate r_f. The resulting profits consist of two components: the first term, $-(r_j(\theta) - r_f)a_j(a, \theta) > 0$, captures the gains from repayment; the second term, $(1 + r_f)a_j(a, \theta) < 0$, captures the losses from default (lenders must repay their depositors). Total profits are computed as $\Pi = \sum_{\theta \in \Theta} \sum_{j=1}^{N} \Pi_j(\theta)$. As mentioned above, these are uniformly transferred to consumers (an assumption we relax in Section 6).

Suppose lender $k \in \{1, 2, ..., N\}$ chooses interest rate $r^k(\theta)$ and borrowing limit $\bar{l}^k(\theta)$. Let $j(r^k(\theta), r^{-k}(\theta))$ be a function that maps a lender's own interest rate $r^k(\theta)$ and the interest rates of their competitors $r^{-k}(\theta) \equiv (r^1(\theta), ..., r^{k-1}(\theta), r^{k+1}(\theta), ..., r^N(\theta))$ to the rank of $r^k(\theta)$ when the interest rates are sorted in ascending order, $j : \mathbb{R}_+ \times \mathbb{R}_+^{N-1} \to \{1, \dots, N\}$. Then, the set of credit lines can be written $S_\theta = \{(r_{j(r^k(\theta), r^{-k}(\theta))}, \bar{l}_{j(r^k(\theta), r^{-k}(\theta))})\}_{k=1}^{N}$, and the profits to credit card firm k are given by

$$\pi^k = \Pi_{j(r^k(\theta), r^{-k}(\theta))}. \tag{9}$$

To understand the notation, consider the following example. If there is one firm (monopolist), then the monopolist chooses the interest rate $r^1(\theta)$ and the borrowing limit $\bar{l}^1(\theta)$ to maximize total profits, $\pi^1(r^1(\theta), \bar{l}^1(\theta)) = \Pi_1(r_1(\theta), \bar{l}_1(\theta))$, where the first expression refers to profits by the lender's identity and the second expression refers to profits using the (degenerate) credit card interest rate ranking.

Forms of lender competition. When we analyze competitive reforms, we consider four forms of competition motivated by the empirical evidence in Section 2: (i) *monopoly*, (ii) *collusive-Cournot* competition, which is a two-stage game where lenders collude on interest rates in the first stage and then Cournot compete on limits in the second stage, (iii) *Stackelberg-Cournot* competition, which is two-stage game where lenders Stackelberg compete on interest rates in the first stage and then Cournot compete on limit in the second stage, and (iv) *competitive pricing*. We numerically characterize lender behavior for each of these forms of competition in Section 5.

Lender entry costs. When we consider the transition path, we must make assumptions regarding lender entry costs. Our benchmark model assumes zero lender entry costs. However, in Section 6, we impose up front lender entry costs equal to the net present value of profits. Since there was profitable lender entry in the 1970s and 1980s, we view this robustness exercise as providing an upper bound on lender entry costs. We show that these costs have second-order effects on welfare, compared with the gains from increased competition.

3.1.3 Equilibrium

Given a risk-free rate r_f, a stationary recursive competitive equilibrium is given by a set of credit lines $S_\theta = \{(r_1(\theta), \bar{l}_1(\theta)), \ldots, (r_N(\theta), \bar{l}_N(\theta))\}$, a stationary distribution over idiosyncratic states $\Omega\,(i, \theta, \eta, \epsilon, a)$, total profits Π, a repayment/default policy function $p(i, \theta, \eta, \epsilon, a)$, a consumption policy function $c(i, \theta, \eta, \epsilon, a)$, a savings/borrowing policy function $a'(i, \theta, \eta, \epsilon, a)$, and a set of credit card firms' best response functions $\{r^k(\theta), \bar{l}^k(\theta)\}_{k=1}^N$ such that

(i) given r_f, S_θ, and Π, the allocations $p(i, \theta, \eta, \epsilon, a)$, $c(i, \theta, \eta, \epsilon, a)$, and $a'(i, \theta, \eta, \epsilon, a)$ solve the consumer's problem in (3), (5), and (6);

(ii) for $k \in \{1, 2, ..., N\}$, $\{r^k(\theta), \bar{l}^k(\theta)\}_{k=1}^N$ maximizes each credit card firm's profits in (9);

(iii) the distribution of consumers $\Omega(i, \theta, \eta, \epsilon, a)$ is consistent with the policy functions $p(i, \theta, \eta, \epsilon, a)$ and $a'(i, \theta, \eta, \epsilon, a)$ and the exogenous process for earnings;

4 Calibration

Given the computationally demanding nature of the model, we take as many standard parameters as possible from the literature, and then we calibrate the remaining parameters to target 1971-75 moments. As discussed in Section 2, the credit card industry was characterized by (1) lack of inter-regional competition, and (2) non-competitive behavior, leading to antitrust litigation regarding the blocking of competitive new entrant credit cards. Therefore, we calibrate the model assuming that there was *pure monopoly* ($N = 1$) in the 1970s as an approximation to the limited competition and regional monopolies during that time period.

We assume that each period corresponds to one year. Table 5 presents the parameters determined outside of the model equilibrium. We use a standard estimate for the risk aversion ($\sigma = 2$). We set the risk-free savings rate equal to 1.27%, which is the average real interest rate from 1971 to 1975.[10] The re-entry probability of good credit standing $\phi = .1$ is chosen so that it takes the average consumer 10 years to re-enter the credit card market upon default. The earnings process is taken from Storesletten, Telmer, and Yaron (2000) (Table 1, Panel D).[11] We assume that permanent types are distributed so that $\ln(\theta) \sim^{idd} N(0, \sigma_\theta^2)$. We approximate this on a symmetric two-point distribution, yielding equal masses of agents at $\theta_H = 1.12$ and $\theta_L = 0.88$. We assume that the persistent component of income η follows an AR(1), $\ln(\eta') = \rho \log(\eta) + u$, where $u \sim^{idd} N(0, \sigma_\eta^2)$. Lastly, the perfectly transitory component is log normally distributed, $\ln(\epsilon) \sim^{idd} N(0, \sigma_\epsilon^2)$.

Table 5: Parameters determined outside the model equilibrium

Parameter	**Description**	**Value**
σ	Risk aversion	2
r	Risk-free rate	1.265
ϕ	Re-entry prob. good credit standing (10 year exclusion)	0.1
σ_θ^2	Variance permanent component θ	0.244
σ_η^2	Variance of innovation to AR(1) component η	0.024
σ_ϵ^2	Variance transitory component ϵ	0.063
ρ	Persistence of AR(1) component η	0.977

The remaining parameters $\{\beta, \chi, \kappa\}$ are estimated jointly to target moments between 1971 and 1975. We estimate $\beta = 0.959$ to match average credit to income of 1.04% from 1971 to 1975. We calibrate stigma $\chi = 4.293$ to match the average bankruptcy rate of 0.06% from 1971 to 1975. Since there are taste shocks, defaults may occur even when the Bellman value of repayment is greater than the Bellman value of default. We attribute these defaults to *unmodeled shocks*. We interpret the data analogue of defaults attributable to *unmodeled shocks* as the share of bankruptcies due to divorce, health costs, or lawsuits reported by Chakravarty and Rhee (1999). We therefore set $\kappa = 3.582$ so that the share of unmodeled defaults in the model coincides with the share of these bankruptcies.

5 Gains from competition

The 1970s credit card market was characterized by non-competitive behavior. However, landmark cases challenging exclusivity (e.g., the *Worthen* case discussed in Section 2) broke down regional monopolies, and the *Marquett* decision in 1978 facilitated national competition. While we do not explicitly have regions in our framework, we model these reforms as a gradual transition from monopoly to oligopoly.

We begin this section by analyzing the problem of the monopolist. Then, we consider four competitive reforms and compute the distribution of welfare gains along the transition path. The first two

[10]We measure this as the Moody's Aaa rate less inflation implied by the NIPA GDP Deflator.

[11]We base our estimates on the working paper version of Storesletten, Telmer, and Yaron (2004), since the working paper reports the relevant income process for our exercise.

Table 6: Parameters determined jointly in equilibrium

	Parameter	Value	Target	Data	Model
β	Discount rate	0.959	Credit to income	1.04	1.02
χ	Stigma	4.293	Bankruptcy rate	0.06	0.06
κ	Scaling parameter	3.582	Fraction of bankruptcy due to divorce, health, lawsuits	44.81	44.16

Notes: Credit data are from the Federal Reserve Board G.19 series. The bankruptcy rate is Chapter 7 flings per capita computed using data from the Historical Statistics of the U.S. Millennial Edition. The fraction of bankruptcy due to divorce, health costs and lawsuits is taken from Chakravarty and Rhee (1999) who compute these statistics in the PSID. The fraction of bankruptcy due to divorce, health costs and lawsuits within model is the fraction of defaults that occur when the continuation value of repayment is greater than the Bellman value of default. See text for more discussion.

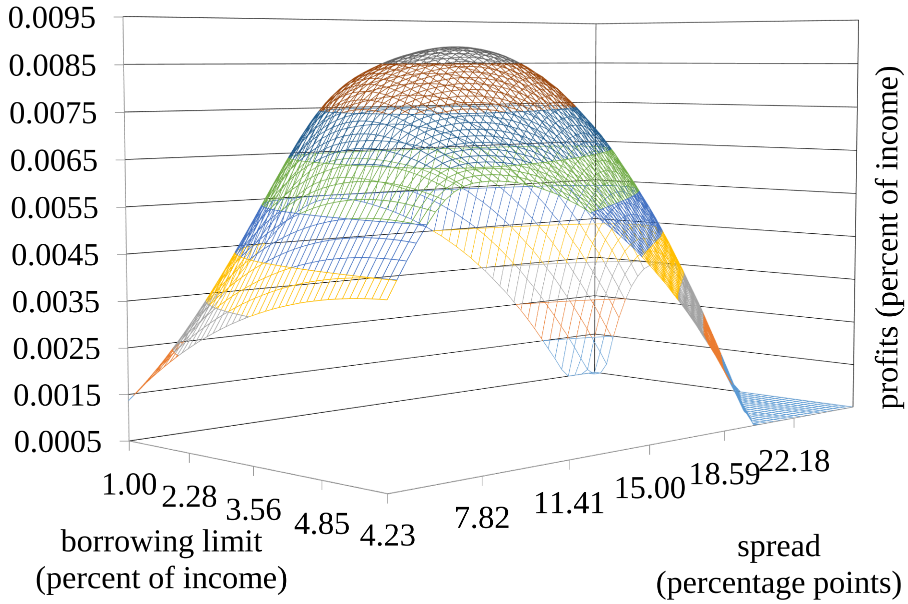
reforms we consider correspond to increased lender entry in an environment with interest rate collusion. First, we analyze a transition to a collusive-Cournot duopoly to capture increased, but limited, competition due to the abandoning of exclusive contracts following the 1970 *Worthen* case. Second, we analyze a transition to a collusive-Cournot oligopoly with twenty lenders. This transition captures increased competition from lender entry following the 1978 *Marquette* decision, while maintaining the assumption of non-competitive pricing in the early 1980s (Knittel and Stango (2003)). Standard Herfindahl indices imply competition equivalent to $N = 20$ equal sized banks in the 1980s.

The third and fourth reforms are designed to measure the benefits from increased price competition and to provide an upper bound on the gains from competition, respectively. The third reform is a transition to a Stackelberg-Cournot duopoly. This experiment of a transition from monopoly to Stackelberg-Cournot duopoly is designed to capture increased interest rate competition in the credit card market after the period of interest rate collusion in the early 1980s. The fourth reform is a counterfactual transition to competitive pricing. This transition provides an upper bound on the welfare gains from the first three competitive reforms.

5.1 Characterization of the monopolist's problem

In this section, we numerically explore the properties of the monopolist's problem. Figure 4 plots profits to the monopolist $(\pi^1(r^1(\theta), \bar{l}^1(\theta)) = \Pi_1(r_1(\theta), \bar{l}_1(\theta)))$ as a function of the spread $\tau^1(\theta) = r^1(\theta) - r_f$ and the borrowing limit $\bar{l}^1(\theta)$ for the consumer with the high permanent earnings component. The monopolist maximizes profits at an interior spread and an interior borrowing limit. This is because if the monopolist chooses a low spread, then the profit margin is low, and hence, profits are low. If the monopolist chooses a high spread, then consumers will borrow less, leading to low profits. A high spread also increases default risk as the composition of borrowers changes. The same is true for the extremes of borrowing limits. By definition, near-zero limits restrict profits, while limits that are too generous result in negative profits due to more defaults. The profit function for the consumer with low permanent earnings is qualitatively similar with lower total profits.

Figure 4: Monopolist profit function for high-earning consumer ($\theta = \theta_H$)



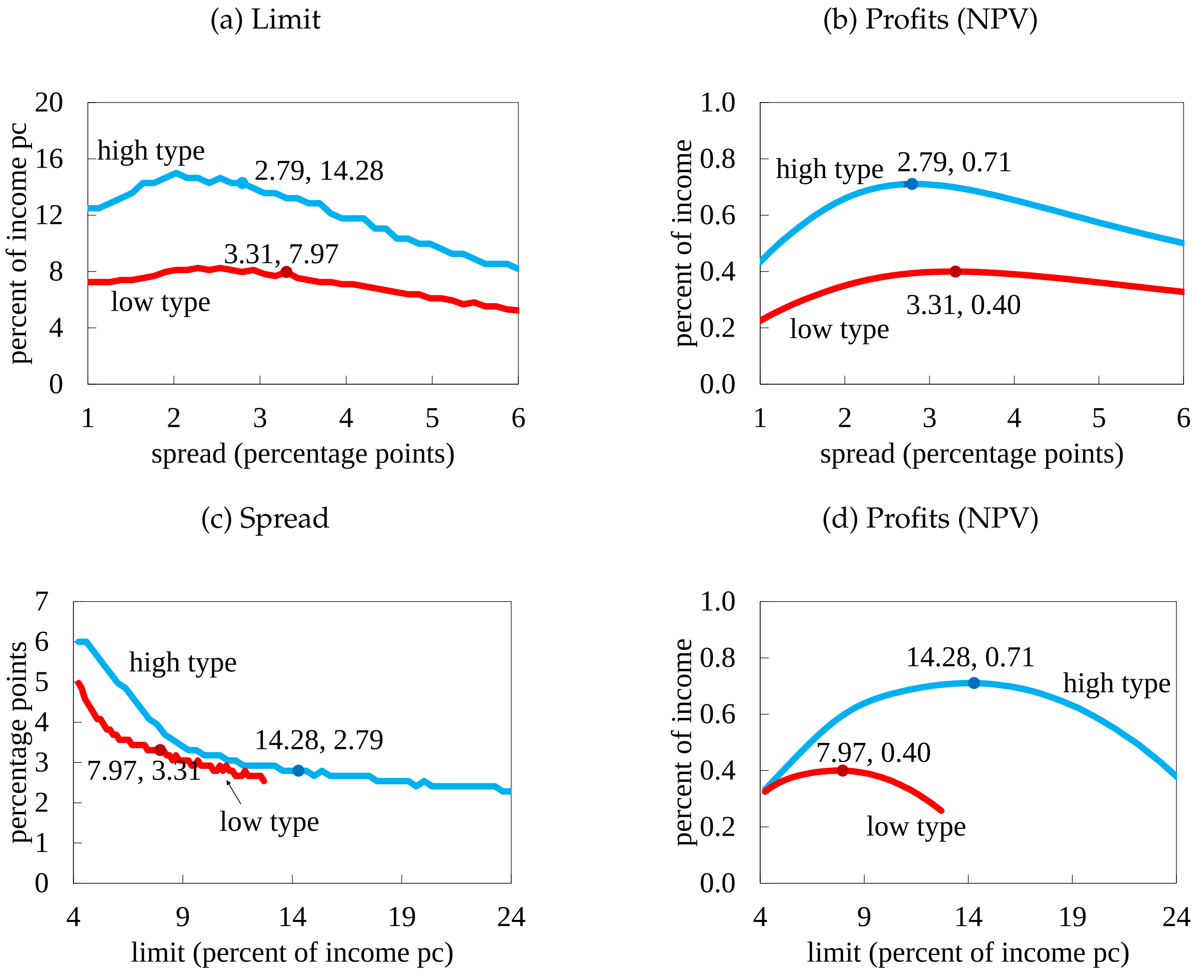
Notes: borrowing limits are expressed as a percentage of income per capita. Spreads are expressed as percentage points over the savings interest rate. Profits are expressed as a percentage of income.

To further understand why the profit function is concave and admits an interior solution, Figure 5 plots the monopolist's optimal policy functions and corresponding profits for both the high-type (blue) and low-type (red). Panels (a) and (b) plot the limits that maximize profits and the corresponding profits as a function of the spread. Hence, the spreads that maximize profits in Panel (b) are the optimal contracts for each type. We see that optimal limits are hump shaped with respect to the spreads for both types. That is, for both very low and very high values of the spread, the monopolist restricts the amount that can be borrowed by cutting limits. A low spread reduces the profit margin for each dollar of borrowing. This restricts the ability of the monopolist to take on more default risk. For high values of the spread, the only agents who borrow are those who have realized extremely low persistent and transitory earnings shocks. These agents default at a very high rate, which again incentivizes the monopolist to restrict lending.

Panels (c) and (d) plot the spreads that maximize profits and the corresponding profits as a function of the limit for both types. Hence, analogous to Panel (b), the limits that maximize profits in Panel (d) are the optimal contracts for each type. In Panel (c), spreads increase for both types as the credit limit declines. This feature is consistent with neoclassical models of monopoly where quantity restrictions raise prices.

Comparing the high-type to the low-type, we see that the policy functions are qualitatively similar except that total profits are lower for the low-type. Further, the monopolist chooses a lower limit and

Figure 5: Monopolist policy functions for both types ($\theta \in \{\theta_H, \theta_L\}$)



Notes: Panel (A) plots the lender's optimal limit when their spread is fixed at the value on the x-axis (blue denoting the high-type and red denoting the low-type). Panel (B) plots the lender's NPV profits as a percentage of income when their spread is fixed at the value on the x-axis and the limit is allowed to freely adjust. Panel (C) plots the lender's optimal spread when their limit is fixed at the value on the x-axis. Panel (D) plots the lender's NPV profits as a percentage of income when their limit is fixed at the value on the x-axis and the spread is allowed to freely adjust. Borrowing limits are expressed as a percentage of income per capita. Spreads are expressed as percentage points over the savings interest rate.

a higher spread for the low-type. Given qualitatively similar lender policy functions for both types, for the remainder of this section, we show lender policy functions for the high-type.

5.2 Non-targeted moments: Spreads and credit access by income

Before discussing the reforms, we compare the monopoly model with two non-targeted moments: spreads and relative limits among high and low earners. While our calibration uses targets from 1971 to 1975, we can only measure spreads starting in 1974. Table 7 shows that the model generates an average spread of 3.05 percentage points, accounting for roughly 40% of the observed spreads in the data. The model generates an excess spread (the spread over and above the break even spread) of 2.05 percentage points, similarly accounting for roughly 40% of the data.

Table 7: Spreads and Excess Spreads, Benchmark Monopoly Calibration vs. Data

Variable (unit=percent)	Monopoly	Data	Source
Spread	3.05	8.48	Board of Governors & Author's Calc.(1974-1975)
Excess spread: actual - zero-profit	2.05	5.69	Board of Governors & Author's Calc. (1974-1975)

Notes: See Section 2 for details on construction of the excess spread in the data. The excess spread in model is defined as $\tau_{avg} - \tau_{zero}$ where $\tau_{avg} = (\tau(\theta_H) + \tau(\theta_L))/2$ and $\tau_{zero} = D(1+r)/(1-D)$ where D is the economy-wide charge-off rate.

We also measure the relative credit limits of high and low earners in the data. Unfortunately, the first micro data available with both credit limits and income is the 1989 Survey of Consumer Finances (SCF). We proxy the high and low permanent earnings groups using college- and non-college-educated workers, respectively. We further restrict the SCF data to employed heads of household between the ages of 24 and 65.[12] While the 1989 *levels* of credit to income are non-comparable with our calibration, as a plausibility check, we compare credit limits in the model with those observed between both types of workers in the data. Table 8 shows that the relative limits of high and low-earning households in our 1970s monopoly calibration is 1.79, whereas the ratio is 2.69 in the 1989 SCF. In our 1970s monopoly calibration, the credit limit to income ratio between high earners and low earners is 1.41, whereas this ratio is 1.21 in the 1989 SCF. Since low-income households default at a higher rate, the monopolist optimally extends lower limits to low-income households. As a result, our model produces disparate access to credit among rich and poor households.

Table 8: Relative Limits of High and Low Permanent Earners, Benchmark Monopoly Calibration vs. Data

Variable (unit=percent)	Monopoly	Data	Source
High earner limit over low earner limit	1.79	2.69	SCF & Author's Calc. (1989)
High earner limit to income over low earner limit to income	1.41	1.21	SCF & Author's Calc. (1989)

Notes: Model column of "High earner limit over low earner limit" corresponds to $l_1(\theta_H)/l_1(\theta_L)$. Data column of "High earner limit over low earner limit" corresponds to the ratio of average credit limits of college-educated prime-age heads in the SCF relative to non-college-educated heads. We require all 1989 SCF heads of household to be employed and thus have well defined credit to income ratios. We further winsorize the data at the 1% level to remove outliers. Model column of "High earner limit to income over low earner limit to income" corresponds to the average credit limit to contemporaneous income ratio of high-types to low-types. Data column of "High earner limit to income over low earner limit to income" corresponds to the credit limit to gross family income ratio of college-educated prime-age heads in the SCF relative to non-college-educated heads. We require all SCF heads of household to be employed and thus have well defined credit to income ratios. We further winsorize the data at the 1% level to remove outliers.

[12] We require employment in order to remove income values of zero, and thus undefined credit to income ratios.

5.3 Monopoly to collusive-Cournot duopoly

We first measure the welfare gains from replacing a monopoly lender with a collusive-Cournot duopoly. This experiment is designed to capture greater – but still limited – competition among credit card issuers after the abandonment of exclusivity rules. Based on evidence of collusion (e.g., Knittel and Stango (2003)) and our micro evidence that shows an extreme lack of interest rate dispersion, we view the 1970s and early 1980s as a period in which lenders colluded on interest rates, but competed by issuing cards; that is, lenders competed on credit limits. As mentioned earlier, we call this form of competition *collusive-Cournot*, and we model it as a two-stage game. In the first stage, forward-looking lenders collude to set a spread ($\tau(\theta)$) for each type $\theta \in \{\theta_L, \theta_H\}$; in the second stage, we assume the lenders Cournot compete on limits in a simultaneous move game for each type, and we analyze the symmetric Nash equilibrium of that game.

We compute the welfare gains from replacing a monopoly lender with a collusive-Cournot duopoly along the transition path. We assume there is a one-time, unexpected, and permanent change from monopoly to collusive-Cournot duopoly at date $t = 1$. Solving for an unrestricted path of spreads and limits would be computationally infeasible. Therefore, we make the following restriction: at date $t = 1$, forward-looking lenders compete as collusive-Cournot duopolists and commit to their new strategies. That is, strategies are restricted to being constant over time. This is still very computationally intensive, since for each set of strategies, we must solve the transition path forward and compute profits.[13] To understand the transition from monopoly to collusive-Cournot duopoly in the 1970s, we first characterize the collusive-Cournot policy functions.

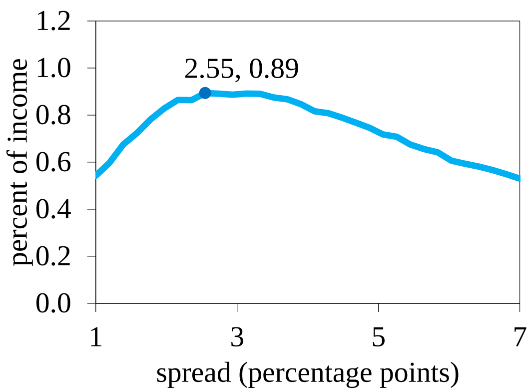
Figure 6 illustrates lender policy functions in each stage of the game. Panel (a) plots the net present value of profits as a function of the spread for type-θ_H consumers. It shows that in the first stage, lenders collude on a spread of 2.55 percentage points to maximize their total second stage profits. Taking the spread from the first stage as given, lenders simultaneously compete over borrowing limits in stage 2. Panel (b) plots the Nash equilibrium limit in stage 2 as a function of the spread. When the two lenders collude on a spread of 2.55 percentage points, the sum of the Nash equilibrium limits on both credit lines is 26.50% of income per capita. Panel (c) plots the best response function of the lenders in the second stage for the case where the spread is 2.55 percentage points. The best response in limits is downward sloping, reflecting the fact that limits are strategic substitutes. This is because when one lender increases their limit, they increase total borrowing and total default risk. Hence, the other lender is forced to tighten their limit. The symmetric Nash equilibrium is determined by the point where the best response function crosses the 45 degree line.

We begin by comparing credit market outcomes across the initial steady state ($t = 0$) and the terminal steady state. Columns (1) and (2) of Table 9 report limits and spreads for both consumer types and other credit-related summary statistics for monopoly and collusive-Cournot duopoly. Both columns (1) and (2) demonstrate that the high-type receives a higher limit and a lower spread compared with the low-

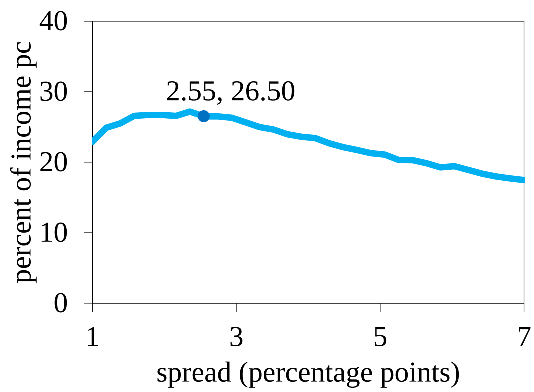
[13] In Appendix F, we show optimization across steady states yields similar strategies which suggests there are small gains from re-optimization later in the transition.

Figure 6: collusive-Cournot policy functions ($\theta = \theta_H$)

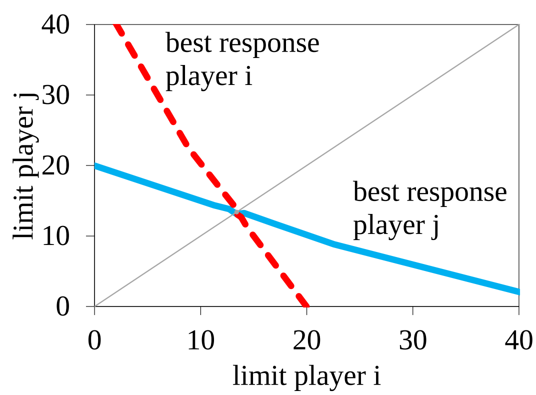
(a) Net present value of profits



(b) Nash borrowing limit in stage 2



(c) Borrowing limit best response in stage 2



Notes: Panel (a) plots the net present value of lender profits as a function of the spread in the first stage. Lenders collude in the first stage to set the spread that maximizes profits. Panel (b) plots the stage 2 Nash equilibrium limit as a function of the spread in the first stage. Panel (c) plots the lenders' best response functions in the second stage for the spread that maximizes profits in the first stage. The symmetric Nash equilibrium is the point where the limits cross the 45 degree line.

type. From monopoly to collusive-Cournot duopoly, total borrowing limits increase by almost 30% for both the high-type and low-type (20.29% to 26.50% of income per capita for the high-type and 11.32% to 14.57% of income per capita for the low-type). However, the fall in spreads is small (2.79 to 2.55 percentage points for the high-type and 3.31 to 2.94 percentage points for the low-type). This is not surprising because lenders collude on spreads and compete on limits.

A side effect of greater competition is a significant expansion of the credit market. Total credit increases by roughly 50% from 1.02% of income to 1.49% of income. Furthermore, both the default rate and charge-off rate increase significantly. Defaults nearly double, rising from a default rate of 0.06% per annum per capita to 0.10% per annum per capita. Likewise, charge-offs increase by 40%, from 0.98% to 1.41%. An implication of (1) roughly constant spreads for both the high and low-type and (2) increasing

Table 9: Comparison of steady state outcomes

Summary Statistics	**(1)** Monopoly	**(2)** collusive-Cournot (N=2)	**(3)** collusive-Cournot (N=20)	**(4)** Stackelberg-Cournot (N=2)	**(5)** Competitive Pricing
High-type					
Line 1: Borrowing limit to initial income pc	20.29	26.50	37.83	5.01	38.13
Line 1: Spread	2.79	2.55	2.74	0.64	1.42
Line 2: Borrowing limit to initial income pc	_	_	_	16.84	_
Line 2: Spread	_	_	_	2.96	_
Low-type					
Line 1: Borrowing limit to initial income pc	11.32	14.57	21.28	4.90	21.82
Line 1: Spread	3.31	2.94	2.74	0.91	1.66
Line 2: Borrowing limit to initial income pc	_	_	_	8.24	_
Line 2: Spread	_	_	_	3.48	_
Credit to income	*1.02*	1.49	2.22	1.33	3.31
Bankruptcy rate	*0.06*	0.10	0.18	0.07	0.16
Charge-off rate	0.98	1.41	2.45	0.91	1.45
Excess spread: actual - zero-profit	2.05	1.30	0.20	1.07	0

Notes: Table reports credit-related summary statistics for the initial monopoly steady state at $t = 0$ (Column (1)) and the steady-states at the end of the transition path after each competitive reform. Column (2) is a collusive-Cournot duopoly, in which lenders collude in the first stage on interest rates and then compete on limits in the second stage. Column (3) is a collusive-Cournot oligopoly with twenty lenders. Column (4) is a Stackelberg-duopoly, where lenders Stackelberg compete on interest rates in the first stage and Cournot compete on limits in the second stage. Column (5) is perfectly competitive pricing. We define the competitive pricing equilibrium to be the limit and interest that maximize welfare of an unborn agent (given their permanent earning ability), subject to weakly positive profits.

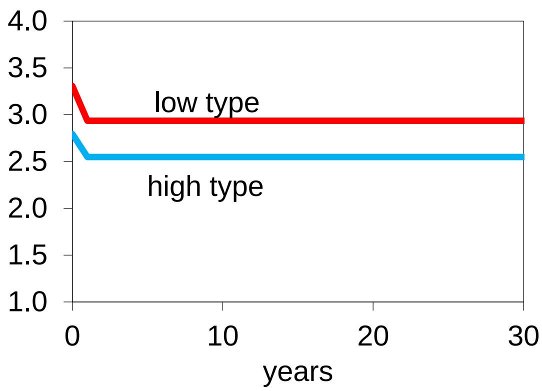
charge-offs is that the *excess spread* decreases from 2.05 percentage points to 1.30 percentage points, a reduction of roughly 35%. While not the focus of this paper, these time trends are qualitatively consistent with observed time trends in aggregate data – for example, Livshits et al. (2010). Other unmodeled factors were also important during this time period, including technology changes in credit scoring (e.g., Livshits et al. (2010), Athreya et al. (2012), Livshits et al. (2016), Sánchez (2018), Raveendranathan (2020)). We provide a complementary rationale for rising bankruptcies in the 1970s and 1980s.

Having compared steady state outcomes between a monopoly and a collusive-Cournot duopoly, we now turn to a transition analysis. Figure 7 illustrates the optimal spreads and credit limits for both the high-type and low-type consumers along the transition path to a collusive-Cournot duopoly. Panel (a) shows the paths for spreads for the low-type (θ_L) and high-type (θ_H) decline once the new entrant begins to Cournot-compete with the monopolist. Likewise, combined limits increase for both the low-type (θ_L) and high-type (θ_H).

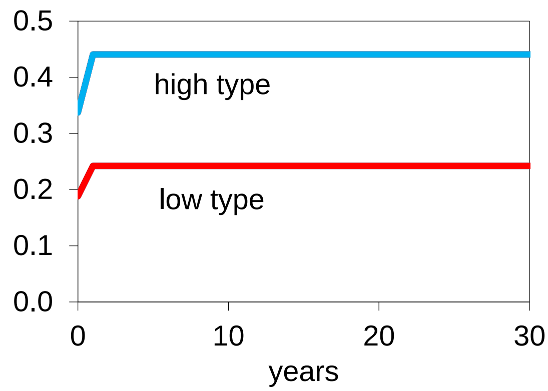
Panels (a) through (d) of Figure 8 plot key variables along the transition path from monopoly to collusive-Cournot duopoly. Since total borrowing limits increase and spreads decrease, individuals borrow more. As a result, credit rises monotonically in Panel (a) of Figure 8. Borrowing limits increase

Figure 7: Transition from monopoly to collusive-Cournot duopoly

(a) Spreads



(b) Borrowing Limits



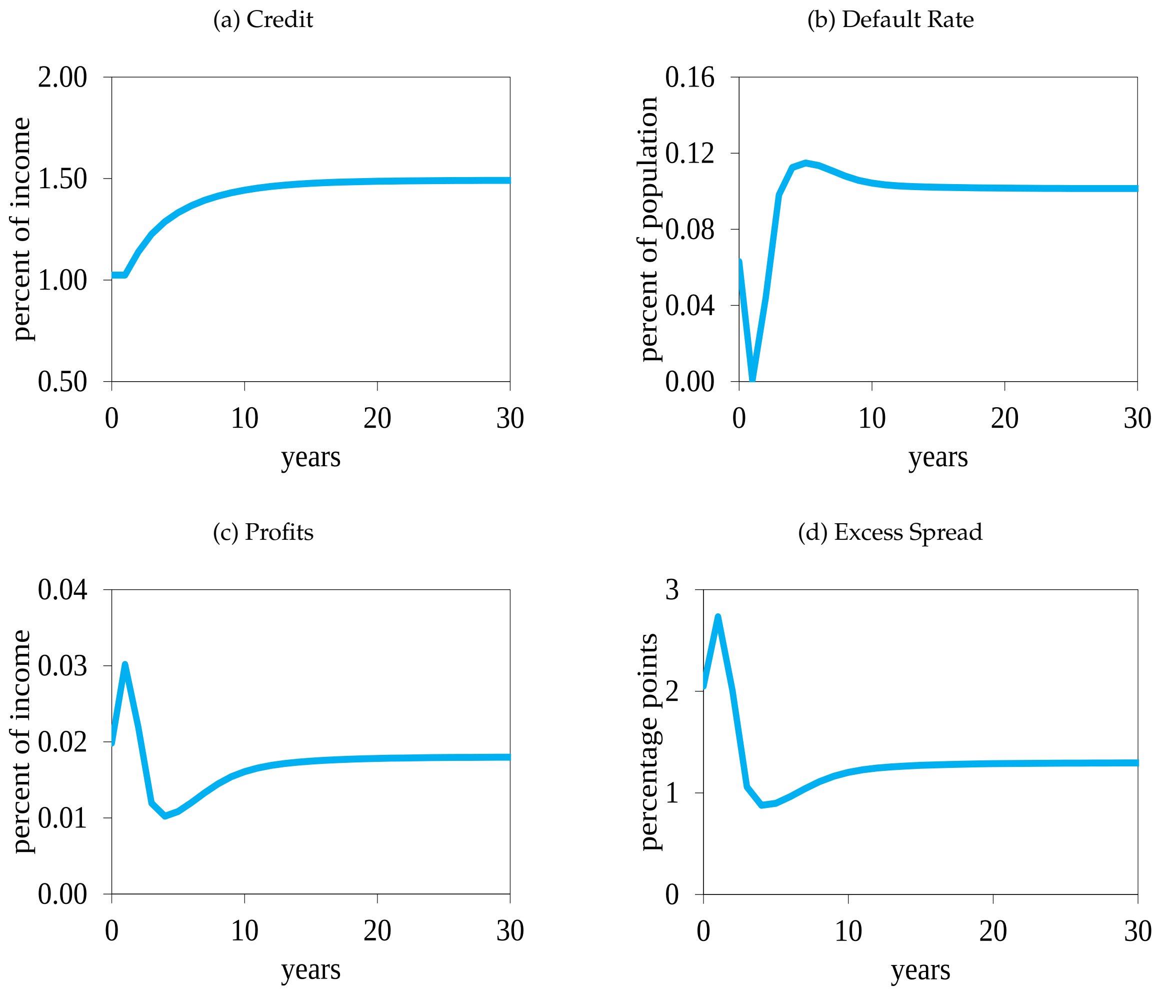
Notes: The initial steady state (t=0) is a monopoly. At date $t = 1$, the economy unexpectedly transitions to a collusive-Cournot duopoly. We assume perfect foresight for subsequent periods. Panel (a) plots the optimal spread of the high-type and low-type. Panel (b) plots the optimal limit of the high-type and low-type.

discontinuously at $t = 1$; however, consumers slowly adjust their savings and borrowing. As a result, the credit utilization rate (credit/limit) falls, and consumers initially have extra slack on their credit lines. This generates the sharp drop in defaults shown in Panel (b). In the long run, however, consumers dis-save and borrow more, which causes defaults to rise in the cross-section. Panel (c) plots the resulting path for aggregate lender profits. The discontinuous decline in defaults generates an initial positive spike in profits. Eventually competition sets in and profits decline, but the consumer credit sector remains profitable throughout the transition path. Qualitatively, a similar pattern emerges in Panel (d) for the excess spread.

Distributional consequences of collusive-Cournot duopoly. We now quantify the *distribution* of welfare gains along the transition path. Table 10 presents the welfare gains generated by the various competitive reforms we consider. We use two measures of welfare: (1) consumption equivalent variation (CEV) and (2) wealth equivalent variation (WEV). Consumption equivalent variation is a standard measure that calculates the lifetime increase of consumption in the initial monopoly steady state such that a consumer is indifferent between the economy with a monopolist and an economy with a duopoly. In our model, we compute CEV numerically, taking into account default costs and taste shocks.

Wealth equivalent variation is the one-time transfer that the consumer requires in the initial monopoly steady state to be just as well off with a duopoly. Wealth equivalent variation is our preferred measure because (1) it allows for aggregation across heterogeneous consumers and (2) it takes into account that consumers re-optimize on their decisions given the one-time transfer. Following Conesa, Costa, Kamali,

Figure 8: Transition from monopoly to collusive-Cournot duopoly



Notes: The initial steady state (t=0) is a monopoly. At date $t = 1$, the economy unexpectedly transitions to a collusive-Cournot duopoly. We assume perfect foresight for subsequent periods.

Kehoe, Nygard, Raveendranathan, and Saxena (2018), it is calculated as follows:

$$
\begin{aligned}
&\min \quad WEV \\
&\quad s.t. \\
V_0(i,\theta,\eta,\epsilon,z,a+WEV) &\geq V_t(i,\theta,\eta,\epsilon,z,a) \\
a+WEV &\geq -\bar{l}^1(\theta) \qquad \text{if } \; i=g \\
a+WEV &\geq 0 \qquad \text{if } \; i=b,
\end{aligned}
\tag{10}
$$

where $V_0(i,\theta,\eta,\epsilon,z,a+WEV)$ refers to the value at the initial steady state (monopolist) given a one-time transfer of WEV, $V_t(i,\theta,\eta,\epsilon,z,a)$ refers to the value in period t along the transition path, and $\bar{l}^1(\theta)$ refers to the borrowing limit in the initial steady state. The last two inequalities ensure that the minimization problem is well defined. When computing $V_0(i, \theta, \eta, \epsilon, z, a + WEV)$, the consumer takes into account profit changes resulting from the reform. When measuring welfare for unborn agents, we assume agents enter in good standing with zero assets and that they draw their earnings states from the ergodic earnings distribution. When aggregating wealth equivalent variation over living cohorts, we use the initial steady state distribution of agents.

Table 10: Welfare gains from monopoly to different forms of competition.

Welfare gains: Monopoly to...	**(2)** collusive-Cournot (N=2)	**(3)** collusive-Cournot (N=20)	**(4)** Stackelberg-Cournot (N=2)	**(5)** Competitive Pricing
CEV unborn at $t = 1$ (% of lifetime consumption)	0.33	0.72	0.14	0.92
WEV unborn at $t = 1$ (% of initial income pc)	2.10	5.70	1.79	8.54
WEV low-unborn at $t = 1$ (% of initial income pc)	1.60	4.70	1.83	6.68
WEV high-unborn at $t = 1$ (% of initial income pc)	2.94	7.38	1.72	11.60
WEV alive at $t = 1$ (% of initial income)	1.55	3.86	1.08	6.03
Population better off (% of population)	100	100	100	100

Notes: This table reports welfare gains along the transition path relative to monopoly steady state. When measuring wealth or consumption equivalent variation for unborn agents, we assume agents enter in good standing with zero assets and that they draw their earnings states from the ergodic income distribution. When aggregating wealth equivalent variation over living cohorts, we use the initial steady state distribution of agents. Welfare is measured as either (a) consumption equivalent variation (CEV) for an unborn agent at the date of the transition $t = 1$, (b) the wealth equivalent variation (WEV) using equation (10) for unborn agents, or (c) WEV for the cohort that is alive at the date of the transition $t = 1$.

Column (2) of Table 10 provides both consumption and wealth equivalent gains from collusive-Cournot duopoly. At the date of the transition experiment ($t = 1$), an unborn agent requires an increase in lifetime consumption of 0.33% to be as well off living in an economy with a single monopoly lender rather than a collusive-Cournot duopoly. Equivalently, an unborn agent requires a one-time transfer at birth worth 2.10% of initial income per capita to be as well off living in an economy with a single monopoly lender rather than a collusive-Cournot duopoly. [14] Among those that are alive at the date of the transition experiment ($t = 1$), the aggregate sum of wealth equivalent variation across workers equals 1.55% of income. As we discuss below, these initial gains are a small fraction of gains from competitive pricing. Lastly, all agents are better off from a new lender entrant, although total profits transferred to the consumers fall.

Panel (a) of Figure 9 plots wealth equivalent variation by earnings decile. Individuals in the lowest earnings decile require a transfer worth $1,200 (in 2016 dollars) to be indifferent between the status quo and transitioning to collusive-Cournot duopoly. Individuals in the highest earnings decile require a transfer worth $400. Panel (b) of Figure 9 expresses the WEV as a ratio of contemporaneous earnings

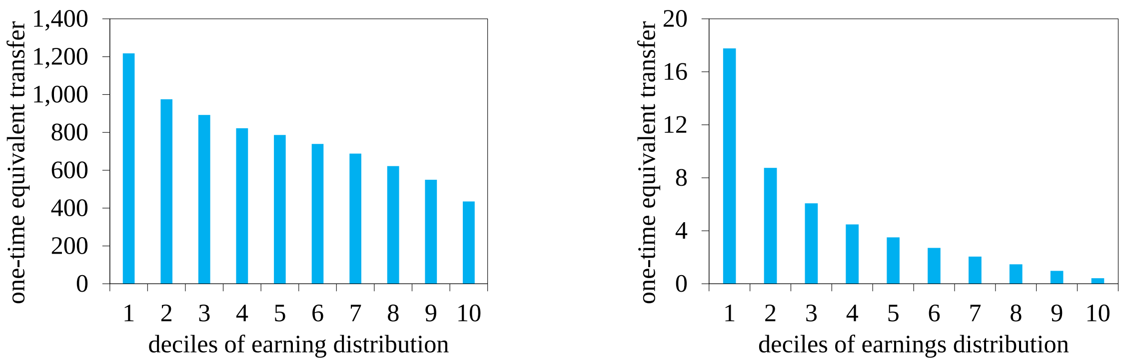
[14] WEV unborn $= E_{\theta,\eta,\epsilon}\text{WEV}(g, \theta, \eta, \epsilon, 0)$. WEV unborn $(\theta) = E_{\eta,\epsilon}\text{WEV}(g, \theta, \eta, \epsilon, 0)$. Aggregate wealth equivalent variation $= \int \text{WEV}(i, \theta, \eta, \epsilon, a) d\Omega(i, \theta, \eta, \epsilon, a)$. Total income $= \int [\theta\eta\epsilon + r\max(0, a) + \Pi] d\Omega(i, \theta, \eta, \epsilon, a)$. Population better off $= \int 1_{\{V_t(i,\theta,\eta,\epsilon,z,a) \geq V_0(i,\theta,\eta,\epsilon,z,a)\}} d\Omega(i, \theta, \eta, \epsilon, a)$.

in each decile. Individuals in the lowest decile of earnings would require a transfer worth nearly 18% of their annual earnings to be indifferent between the status quo and transitioning to collusive-Cournot duopoly. Individuals in the highest deciles would require transfers worth very little of their annual earnings (0.4 percent). This reflects the disproportionate burden of monopoly power borne by low-income households. [15]

Figure 9: Welfare gains by earnings decile along transition path from monopoly to collusive-Cournot duopoly (N=2)

(a) Wealth Equivalent Variation (WEV) by earnings decile

(b) Wealth Equivalent Variation (WEV) over income by earnings decile



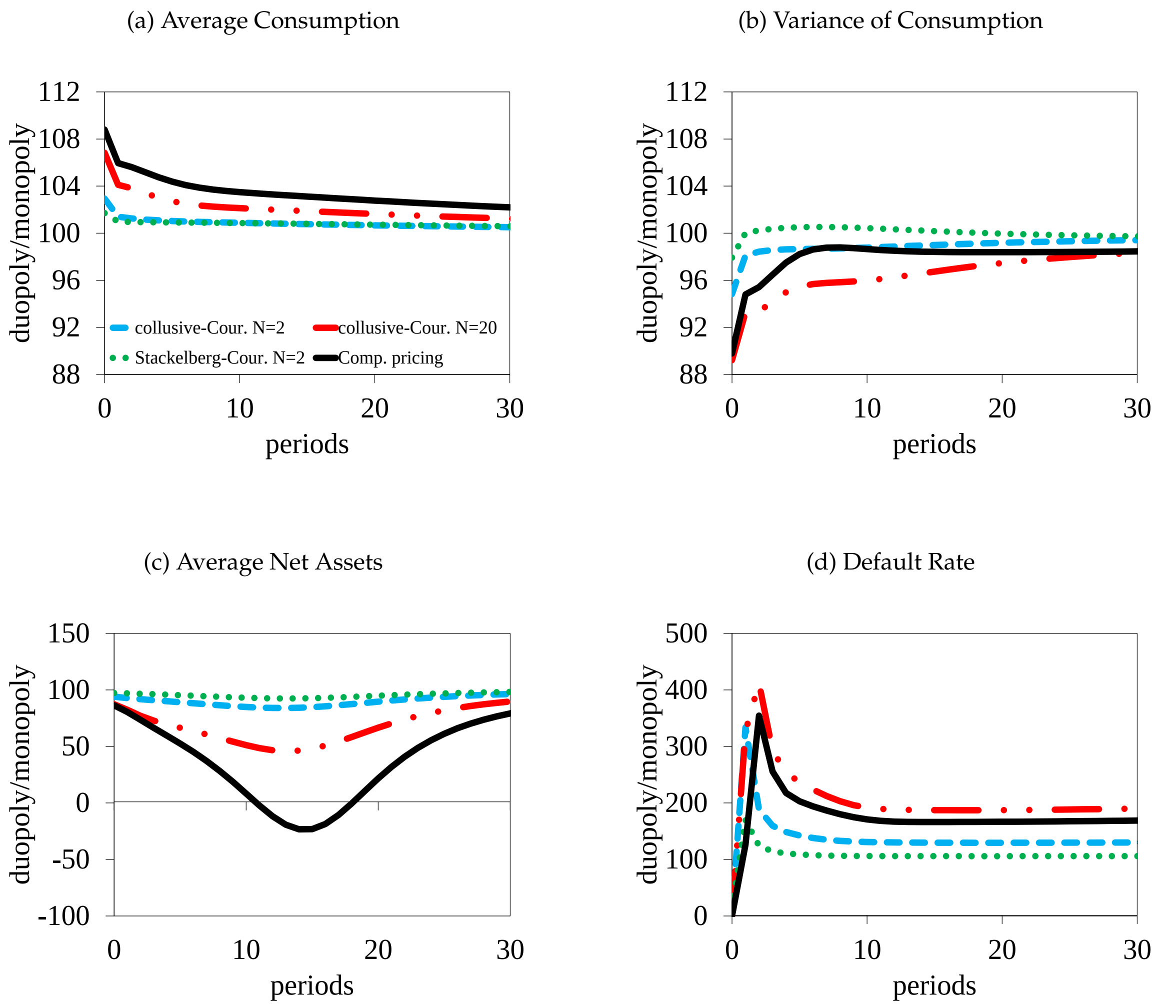
Notes: Welfare gains from the transition are measured using wealth equivalent variation in Panel (a). Panel (b) takes the ratio of the wealth equivalent variation to earnings in each decile.

To understand what drives the welfare gains among low-earning individuals, we study implications for average consumption, variance of consumption, average net assets, and default probabilities of agents in the lowest decile of earnings at the start of the transition experiment ($t = 1$). We compare implications between remaining in the monopoly economy and transitioning to collusive-Cournot duopoly, as well as to other forms of competition in subsequent experiments (i.e., collusive-Cournot oligopoly with twenty lenders, Stackelberg-Cournot duopoly, and competitive pricing, which we discuss later). We find that low-earning individuals benefit from higher average consumption, lower variance of consumption, and lower default rates along the transition path.

Panel (a) of Figure 10 plots the relative average consumption profile of agents in the bottom decile of earnings at the time of the transition ($t = 1$). More specifically, Panel (a) plots the ratio of average consumption along the transition path divided by the *status quo* average consumption if the agents remained in the monopoly economy. Agents consume roughly 3% more along the collusive-Cournot duopoly transition path relative to remaining in steady state with a single monopoly lender (henceforth, *status quo*). In terms of consumption volatility, increased lender competition allows agents to better smooth consumption. Panel (b) shows that the variance of consumption decreases by 5% along the transition path to a collusive-Cournot with two lenders relative to status quo. Panel (c) shows that low-earning individuals

[15] In the bottom decile, 64% of consumers are type-θ_L consumers and 36% are type-θ_H consumers.

Figure 10: Average consumption, asset, and default profiles of a low-earning consumer



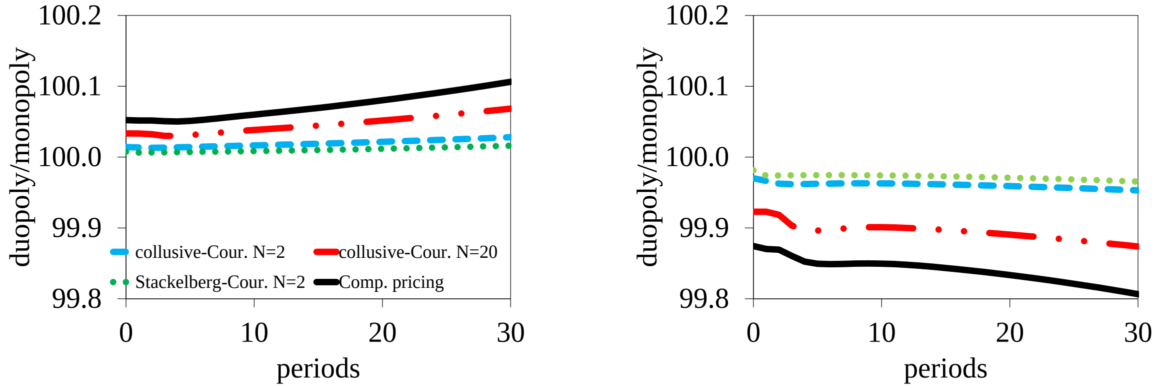
Notes: Panels (a) through (d) are derived from simulating agents in the bottom-decile of earnings at date $t = 1$ on the transition path to (1) collusive-Cournot duopoly (2) collusive-Cournot (N=20) (3) Stackelberg-Cournot duopoly, and (4) single-lender competitive pricing. Panel (a) plots the average consumption path along the transition path expressed as a ratio to average consumption if the agent remained in a world with a single-lender monopoly. Panel (b) plots the variance of consumption along the transition path expressed as a ratio to the variance of consumption if the agent remains in a world with a single-lender monopoly. Panels (c) and (d) repeat the same exercise for net assets and defaults, respectively.

are able to dis-save relatively more with greater lender competition. Because credit limits expand, default rates initially decline in Panel (d). In the long run, however, default rates rise above the monopoly case because agents accumulate more debt.

In figure 11, we show that the top 10% of earners also gain from greater competition through higher average consumption and lower consumption variance along the transition path. However, since the highest earners less frequently rely on credit to smooth consumption, their welfare gains are an order of magnitude smaller.

Figure 11: Average consumption and variance of consumption profiles of a high-earning consumer

(a) Average Consumption (b) Variance of Consumption



Notes: Panels (a) through (b) are derived from simulating agents in the top-decile of earnings at date $t = 1$ on the transition path to (1) collusive-Cournot duopoly (2) collusive-Cournot (N=20) (3) Stackelberg-Cournot duopoly, and (4) single-lender competitive pricing. Panel (a) plots the average consumption path along the transition path expressed as a ratio to average consumption if the agent remained in a world with a single-lender monopoly. Panel (b) plots the variance of consumption along the transition path expressed as a ratio to the variance of consumption if the agent remains in a world with a single-lender monopoly.

5.4 Monopoly to collusive-Cournot oligopoly

To capture greater entry resulting from the 1978 *Marquette* decision and other lawsuit resolutions in the 1980s such as the *Worthen* case, we now consider the implications of a further increase in lender entry in an environment with price collusion. We measure the welfare gains from replacing a monopoly lender with a collusive-Cournot oligopoly of twenty lenders, in which the number of lenders is motivated by the inverse Herfindahl in the late 1980s (see Section 2).

Column (3) of Table 9 reports the spreads and limits for both consumer types and credit statistics. When the number of lenders increases from one to twenty, we observe a significant increase in limits of almost 90% for both types. Recall that in the reform from monopoly to collusive-Cournot duopoly, the increase was roughly 30% for both types. The change in the spreads is small. As mentioned above, this is not surprising given the nature of competition in this transition. Lenders compete on limits, but not on spreads. [16]

Column (3) of Table 10 reports the welfare gains from collusive-Cournot oligopoly. The aggregate wealth equivalent gains from monopoly to collusive-Cournot oligopoly are 2.5 times larger than the gains from monopoly to collusive-Cournot duopoly (3.86% vs. 1.55% of income). Panel (a) of Figure 12 plots wealth equivalent variation by earnings decile for the transition from monopoly to collusive-

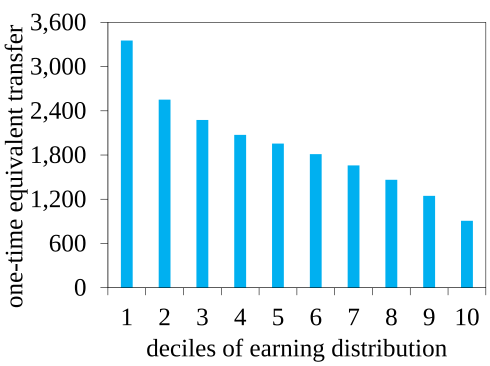
[16] When we compare collusive-Cournot (N=2) with collusive-Cournot (N=20), the spread for the high-type increases from 2.55 to 2.74 percentage points while it decreases from 2.94 to 2.74 percentage points for the low-type. This is because there are two opposing effects on spreads: (1) competitive forces lower spreads, and (2) since individuals cannot commit to borrow from only one lender, debt dilution raises spreads (e.g., Bizer and DeMarzo (1992) and Hatchondo and Martinez (2018)).

Cournot oligopoly. The gains for individuals in the lowest earnings decile are now worth $3,400 (in 2016 dollars), which is almost 50% of their income. Hence, the gains for these consumers are roughly three times larger compared with the gains from monopoly to collusive-Cournot duopoly. The 10% highest earners gain from oligopoly as they do from duopoly, but by much less. Their welfare gains, in 2016 dollars, are roughly three times smaller.

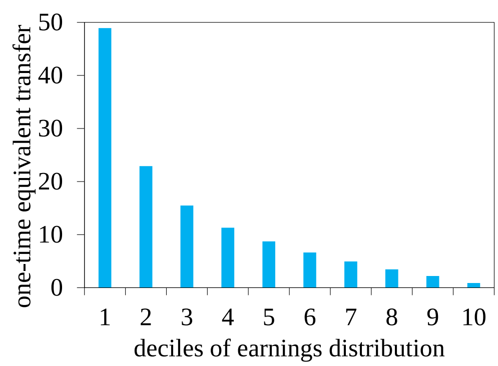
As Figure 10 demonstrates, low-income individuals gain from substantial improvements in the ability to smooth consumption. Panel (a) shows that among the poorest 10% of agents at the date of the transition, they consume roughly 7% more along the transition path to collusive-Cournot with twenty lenders as they do in the status quo (i.e., remaining in the monopoly steady state). Panel (b) shows that the variance of consumption decreases by 11% along the transition path to a collusive-Cournot with twenty lenders relative to the variance in the status quo. Over time, consumption variance remains significantly lower.

Figure 12: Welfare gains by earnings decile along transition path from monopoly to collusive-Cournot oligopoly (N=20)

(a) Wealth Equivalent Variation (WEV) by earnings decile



(b) Wealth Equivalent Variation (WEV) over income by earnings decile



Notes: Welfare gains from the transition are measured using wealth equivalent variation in Panel (a). Panel (b) takes the ratio of the wealth equivalent variation to earnings in each decile.

In summary, even with interest rate collusion, increased lender entry leads to large welfare gains due to an expansion in credit limits. Further, the welfare costs of monopoly in the lending market are borne primarily by low-earning individuals.

5.5 Monopoly to Stackelberg-Cournot

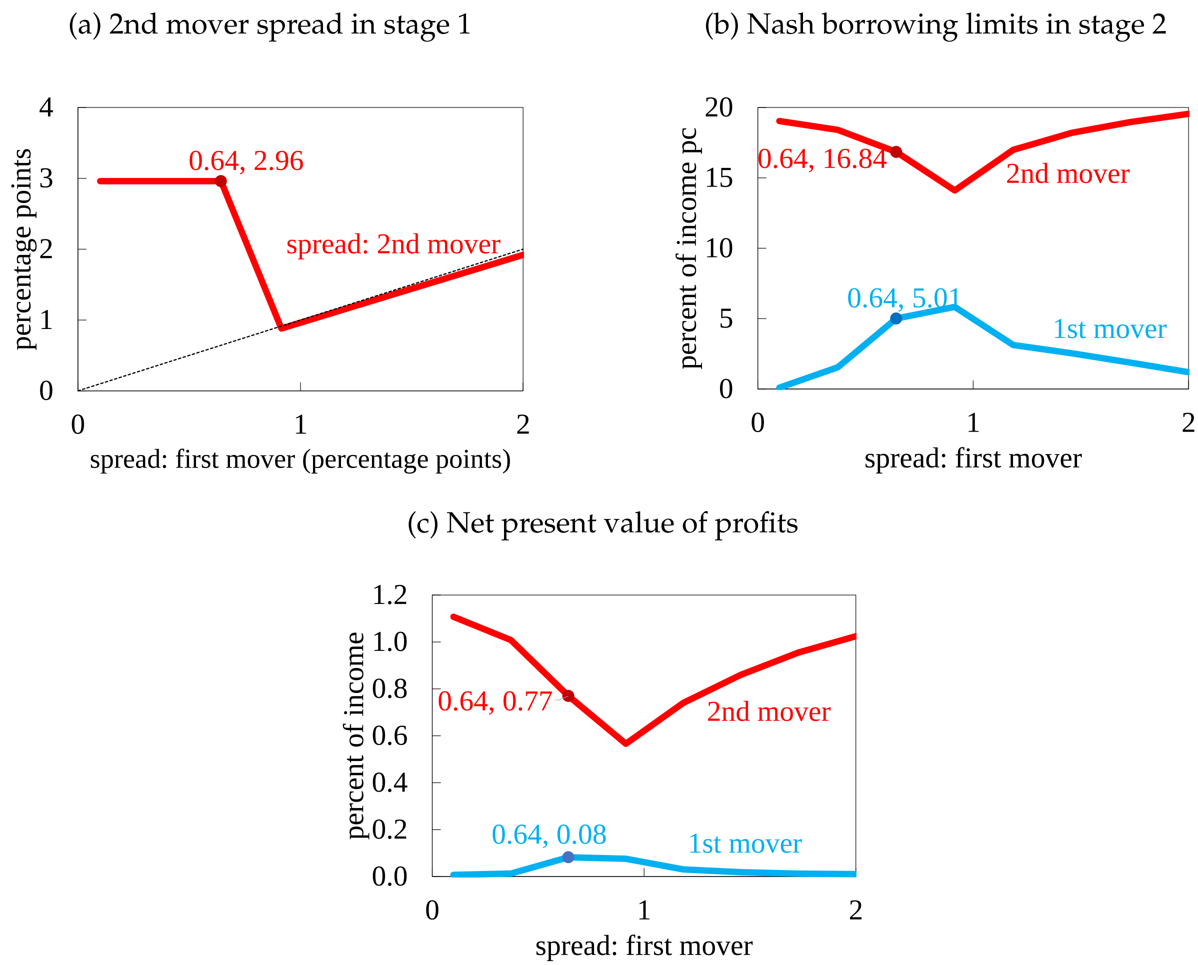
In this section, we introduce price competition by analyzing a transition from monopoly to Stackelberg-Cournot. As mentioned earlier, the lenders Stackelberg compete on spreads in the first stage and Nash compete on limits in the second stage. We assume Stackelberg competition on spreads because a pure strategy Nash equilibrium does not exist, and solving for a mixed strategy Nash equilibrium is not tractable. In contrast to collusive-Cournot analyzed in the previous section, there is now price competition in the first stage. In stage 1, the second mover chooses a profit maximizing spread given the first mover's spread. The first mover chooses a profit maximizing spread internalizing the second mover's best response. The stage 2 competition on limits is the same as collusive-Cournot, except the Nash equilibrium for limits may no longer be symmetric because the lenders can choose different spreads.

To maintain tractability, we assume there are two lenders. This experiment of a transition from monopoly to Stackelberg-Cournot duopoly is designed to capture interest rate competition in the credit card market after the period of interest rate collusion in the early 1980s. We assume the incumbent lender (the monopolist) is the first mover and the new entrant is the second mover. In the period of the transition, the profits or losses on the existing credit line are borne by the monopolist. In principle, the new spread and new limit chosen by the monopolist as a first mover or second mover could affect profits on their existing credit line in the period of the transition. However, numerically, the assumption that the monopolist moves first has no significant impact on the equilibrium spreads and limits. To understand the transition path from monopoly to Stackelberg-Cournot duopoly, we first characterize each Stackelberg-Cournot duopolist's policy functions.

For ease of exposition, we assume Lender 1 is the first mover, and Lender 2 is the second mover. Let τ^1 denote the first mover's spread. In Figure 13, panel (a) plots the second mover's best response in spreads $\tau^2(\tau^1)$ in stage 1. Panel (a) shows that if the first mover commits to a large spread in the first stage ($\tau^1 > 0.64\%$), the second mover will undercut the first mover and set τ^2 just below τ^1; that is, $\tau^2 = \tau^1 - \epsilon$ for arbitrarily small ϵ. In this region ($\tau^1 > 0.64\%$), spreads are strategic complements; that is, $\frac{d\tau^2(\tau^1)}{d\tau^1} \geq 0$. If the first mover increases their spread, the second mover also increases their spread and undercuts the first mover. This is typical in Stackelberg-Bertrand games. However, strategic complementarity of spreads does not hold for all first-mover spreads. There is a threshold at which the second mover's undercutting strategy is no longer profitable relative to the alternate strategy of charging a higher spread and becoming the second-ranked lender. For extremely low spreads, the second mover is made strictly better off by setting a high spread ($\tau^1 \leq 0.64\%$ in panel (a)). The spread at which the second mover abandons their undercutting strategy is the equilibrium interest rate. In equilibrium, the first mover, Lender 1, sets a spread of 0.64%, and the second mover, Lender 2, sets a spread of 2.96%.

Panel (b) of Figure 13 plots the stage 2 Nash borrowing limits for both lenders given the first mover's spread and the second mover's best response in spreads (i.e., $\bar{l}^1(\tau^1, \tau^2(\tau^1))$, and $\bar{l}^2(\tau^1, \tau^2(\tau^1))$, respectively). In equilibrium, the first mover offers a limit equal to 5.01% of income per capita, three times less than the second mover's limit (16.84% of per capita income). The first mover picks a lower limit because, unlike the collusive-Cournot duopoly, with Stackelberg-Cournot competition consumers will first borrow from the first lender. This is because the first mover, Lender 1, offers a lower interest rate. Thus, there is stronger debt dilution for Lender 1. Lender 1 understands that some of its consumers will subsequently borrow from Lender 2, raising default risk on Lender 1's own loans. Ideally, Lender 1 would raise rates to reflect the default risk generated by the presence of Lender 2, but the threat of being undercut constrains Lender 1's ability to charge a higher rate. Thus, lender 1 picks a lower spread and restricts the amount of credit they extend to mitigate default risk.

Figure 13: Stackelberg policy functions given 1st mover spread ($\theta = \theta_H$)



Notes: Panel (a) plots the second mover's best response in spreads in stage 1 as a function of the first mover's spread. Panel (b) plots the stage 2 Nash borrowing limits for both lenders given the first mover's spread and the second mover's best response in spreads. Panel (c) plots the net present value of profits for both lenders given the first mover's spread, the second mover's best response in spreads, and Nash borrowing limits for both lenders given their respective spreads. Spreads are expressed as percentage point deviations over the risk-free rate. Borrowing limits are expressed as a percentage of income per capita. Net present value of profits are expressed as a percentage of income.

Lastly, panel (c) of Figure 13 plots the net present value of profits for both lenders given the first mover's spread, the second mover's best response in spreads, and Nash borrowing limits for both lenders given their respective spreads $(\tau^1, \tau^2(\tau^1), \bar{l}^1(\tau^1, \tau^2(\tau^1)), \bar{l}^2(\tau^1, \tau^2(\tau^1)))$. The figure clearly illustrates Lender 2 has higher profits, reflecting the second mover advantage inherent in price competition. The equilibrium is determined by the spread at which the first mover maximizes profits ($\tau^1 = 0.64\%$), which is also the spread that induces the second mover to abandon undercutting and set their spread to 2.96%.[17] To summarize, the Stackelberg-Cournot equilibrium features the first mover setting a low

[17] This logic is also why a pure strategy Nash equilibrium does not exist. For low spreads, a profitable deviation is to set a

spread ($\tau_1 = 0.64\%$) and low limit (5.01% of income pc) and the second mover setting a high spread ($\tau_2 = 2.96\%$) and high limit (16.84% of income pc).[18]

Column (4) of Table 9 reports limits and spreads for both consumer types and other credit-related summary statistics for Stackelberg-Cournot duopoly. In comparison with monopoly in Column (1), Stackelberg-Cournot duopoly allows both consumers to benefit from a lower spread on the first credit line and higher total borrowing limits. In comparison with consumers under collusive-Cournot duopoly in Column (2), consumers benefit from a low spread on the first line. However, total borrowing limits are *lower* in Stackelberg-Cournot duopoly. The reason for the lower total borrowing limit is the presence of stronger debt dilution in the equilibrium with Stackelberg-Cournot.

In Table 10, when we compare welfare outcomes from monopoly to collusive-Cournot duopoly vs. monopoly to Stackelberg-Cournot duopoly, we see that the low-type benefits more from price competition in Stackelberg-Cournot duopoly (a one-time transfer worth 1.83% of income in Stackelberg-Cournot and 1.60% of income in collusive-Cournot). The high-type does not benefit as much because of the lower total borrowing limit (a one-time transfer worth 1.72% of income in Stackelberg-Cournot and 2.94% of income in collusive-Cournot). Hence, while price competition increases welfare directly by lowering spreads, it could also dampen welfare by reducing the total borrowing limit.

5.6 Monopoly to competitive pricing

Lastly, we analyze a counterfactual transition from monopoly to competitive pricing. This exercise serves two purposes. First, it provides an upper bound for the welfare costs borne by households. Second, it acts as a benchmark against which we can measure the relative effectiveness of the previous competitive reforms.

Unlike the prior experiments, there is no lender entry in this section. Rather, we assume that there is a single lender that unexpectedly and permanently changes from monopoly pricing to competitive pricing at $t = 1$. We must define what competitive pricing means in this context. Our first condition for competitive pricing is zero profits. However, this condition alone is not sufficient to uniquely determine the equilibrium. For any given limit, Figure 4 shows that there is a zero-profit spread. Likewise, for any given spread, there is a zero-profit credit limit.

The way we resolve this issue is to define competitive pricing to be the combination of interest rates and limits that maximizes the expected utility of an unborn agent in good standing with zero assets and permanent earnings component θ, subject to weakly positive lender profits:

$$\max_{r(\theta),\bar{l}(\theta)} \quad E_{\eta,\epsilon} V_0(g,\theta,\eta,\epsilon,0)$$
$$s.t.$$
$$\sum_{t=0}^{\infty} \left(\frac{1}{1+r_f}\right)^t \pi_t(\theta) \geq 0.$$

high spread. However, if your competitor sets a high spread then a profitable deviation is to undercut, etc.

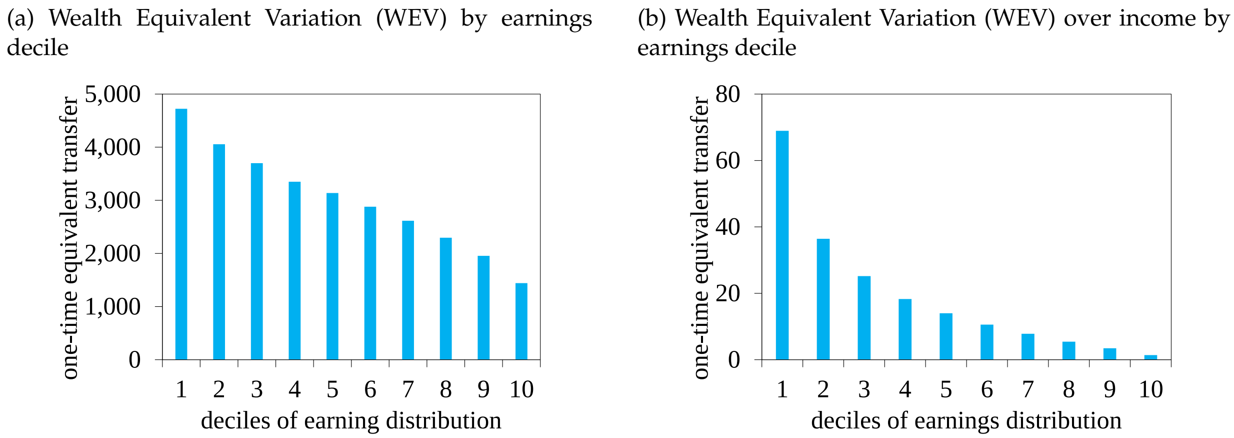
[18] Appendix C discusses the second mover's policy functions.

Column (5) of Table 9 describes the competitive pricing steady state. Compared with monopoly pricing, the credit limits are nearly twice as large (38.13% versus 20.29% for the high-type and 21.82% versus 11.32% for the low-type). Likewise, spreads fall by 50% (2.79 versus 1.42 percentage points for the high-type and 3.31 versus 1.66 percentage points for the low-type), and credit to income more than triples, increasing from 1.02 to 3.31 percent. A side effect of increased borrowing is that the default rate more than doubles, rising from 0.06 to 0.16 percent. Because the lender breaks even, excess spreads are zero by construction.

The large reduction in spreads and increase in limits generates considerable welfare gains. Column (5) of Table 10 shows that an unborn agent would require an increase in lifetime consumption of 0.92% to be as well off living in an economy with a single lender that behaves monopolistically rather than a single lender that behaves competitively. The wealth equivalent gain to the same consumer, expressed as a one-time transfer, is equal to 8.54% of initial income per capita. The aggregate wealth equivalent variation across workers who are alive at the date of the transition equals 6.03% of initial income. We also find much larger distributional welfare gains. Panel (a) of Figure 14 shows that the lowest earning individuals would require a one-time transfer worth $4,700 to be indifferent between a single lender that prices monopolistically and a single lender that prices competitively. Individuals in the highest earnings decile would require a transfer worth $1,500. Expressing the wealth equivalent variation as a ratio of earnings in Panel (b), we find that individuals in the lowest decile of earnings would require a transfer worth almost 70% of their annual earnings to be indifferent between a single lender that prices monopolistically and a single lender that prices competitively. Higher mean consumption, lower consumption variance, and lower default rates drive these gains. Panel (a) of Figure 10 shows that with competitive pricing, agents on average consume roughly 9% more during the transition than they do in the status quo monopoly steady state. Panel (b) shows that consumption variance falls by approximately 10% with competitive pricing relative to status quo.

When we compare the reforms discussed in the previous sections with competitive pricing, we find that collusive-Cournot duopoly and Stackelberg-Cournot duopoly generate 18-26% of total gains from competitive pricing (aggregate WEV as a percentage of income equals 1.55 for collusive-Cournot duopoly, 1.08 for Stackelberg-Cournot duopoly, and 6.03 for competitive pricing). The collusive-Cournot oligopoly with twenty lenders generates 65% of gains from competitive pricing (3.86/6.03). Hence, the transitions from monopoly to collusive-Cournot duopoly or Stackelberg-Cournot duopoly generate a small fraction of competitive pricing welfare gains. The transition to collusive-Cournot oligopoly with twenty lenders generates a large fraction of competitive pricing welfare gains. This is because the credit limits in both collusive-Cournot oligopoly with twenty lenders and competitive pricing are roughly the same. Hence, increased lender entry and the expansion in credit limits are important to account for the welfare gains from competitive reforms in the credit card industry. Further, the main difference between collusive-Cournot oligopoly with twenty lenders and competitive pricing is that competitive pricing leads to significantly lower spreads. This also shows that of the aggregate welfare gains from competitive pricing, roughly 65% of them are explained by higher limits (3.86/6.03). Lower spreads explain roughly 35% of the gains.

Figure 14: Welfare gains by earnings decile along transition path from monopolistic single lender to a perfectly competitive single lender.



Notes: Welfare gains from the transition are measured using wealth equivalent variation in Panel (a). Panel (b) takes the ratio of the wealth equivalent variation to earnings in each decile.

6 Robustness

In this section, we discuss two robustness exercises. We consider (1) fixed costs of lender entry, and (2) an alternate redistribution of lender profits to high-earning households.

First, we assess the role of lender entry costs for our welfare analysis. We use our simulated lender profits and our aggregate wealth equivalent variation to provide bounds on what private parties and/or society would be willing to pay for a new entrant. Our first exercise recomputes household welfare if there is a fixed lender cost equal to the net present value of lender profits that must be paid up front. The net present value of profits is the maximum amount a lender would be willing to pay to enter the market. The initial fixed cost is equally distributed among households. We know that there was lender entry in the 1970s, and so the fixed cost must be weakly lower than the discounted stream of profits.[19]

Our second exercise is to compute the amount *society* would be willing to give up to have another lender enter. This amount is given by aggregating across wealth compensating variation (WCV), which we define in equation (11). WCV is negative for households that are better off with lender entry and positive for households that are worse off. In the case that WCV is a negative number, this is the amount households are willing to *give up* along the transition path with greater lender entry to be just as well off as they were in the initial monopolistic steady state. Similar to WEV, WCV can be aggregated across

[19] Appendix E presents the modifications made to the baseline model for these robustness exercises.

households to obtain the total amount individuals would be willing to give up for lender entry.[20]

$$
\begin{aligned}
\min \;\; & WCV \qquad\qquad (11)\\
& s.t. \\
V_0(i,\theta,\eta,\epsilon,z,a) \leq & V_t(i,\theta,\eta,\epsilon,z,a+WCV) \\
a + WEV \geq & -\bar{l}^1 \qquad\quad \text{if } \; i = g \\
a + WEV \geq & 0 \qquad\qquad \text{if } \; i = b.
\end{aligned}
$$

Table 11 reports our results for collusive-Cournot duopoly (Column (2)), collusive-Cournot with twenty lenders (Column (3)), and Stackelberg-Cournot duopoly (Column (4)). In all three cases, we measure welfare among the cohort that is alive at the date of the transition.

If we assume that the entry cost of the new lender is equal to the discounted flow profits of the lender (row (2) in Table 11), then the aggregate wealth equivalent variation as a percentage of income falls from 1.55 to 1.09 for collusive-Cournot duopoly, 3.86 to 3.73 for collusive-Cournot with N=20, and 1.08 to 0.17 for Stackelberg-Cournot duopoly. The fixed cost reduces the welfare by a smaller amount for collusive-Cournot and by a larger amount for Stackelberg-Cournot. This is because, in Stackelberg-Cournot, we assume the entrant is the second mover. The second mover captures a larger share of the profits. Therefore, in this analysis, a larger share of profits implies a larger fixed cost, which dampens welfare gains. Given that there was profitable lender entry in the 1970s, we view an entry cost equal to the net present value of lender profits as an upper bound, and thus our reported welfare gains in Table 11 can be viewed as lower bounds. If we compute wealth compensation variation (row (3) in Table 11), in all three cases, the welfare estimate is nearly identical to the wealth equivalent variation estimates in our benchmark reforms.

Our second robustness check is to consider an alternative distribution of lender profits. We assume that lender profits are only rebated to households in the top 0.1% of the earnings distribution. We recompute welfare gains from lender entry in row (4) of Table 11. We find that welfare gains are generally higher in this economy. Low earning households no longer have flow profits from lenders each period, and thus the fall in total profits does not affect them. As a result, individuals alive at the date of the transition from monopoly to collusive-Cournot duopoly would require a transfer worth 1.64% of initial income compared with 1.55% in the benchmark reform. Similarly, individuals would require a transfer worth 4.53% of initial income in comparison to 3.86% of initial income in the transition to collusive-Cournot with twenty lenders. Hence, the welfare gains are roughly 6% to 17% larger than our benchmark reforms. We find similar results when we consider Stackelberg-Cournot competition.

[20] Because of the timing assumption on the transition path, $\bar{l}^1$ is still the relevant borrowing constraint for WCV since previously borrowed money is owed to the prior monopoly lender. That monopoly lender committed to lend $\bar{l}^1$.

Table 11: Welfare along transition path under alternative assumption for lender fixed costs and profit distribution

Welfare gains of alive at t=1: Monopoly to...	**(2)** collusive-Cournot (N=2)	**(3)** collusive-Cournot (N=20)	**(4)** Stackelberg-Cournot (N=2)
	(unit = % of initial income)		
(1) Benchmark, WEV	1.55	3.86	1.08
(2) Fixed cost equal to NPV lender profits, WEV	1.09	3.73	0.17
(3) Benchmark, WCV	1.50	3.63	1.04
(4) Profits distributed to top 0.1%, WEV	1.64	4.53	1.33

Notes: (1) "Benchmark, WEV" taken from Table 10, expresses aggregate WEV of the cohort alive at the time of the transition over income in the initial equilibrium. (2) "Fixed cost equal to NPV lender profits, WEV" assumes new lenders initially make a loss equal to the NPV of future profits and the loss is equally distributed among households. (3) "Benchmark, WCV" (defined in equation (11)) computes the amount consumers would be willing to give up for lender entry. (4) "Profits distributed to top 0.1%, WEV" computes aggregate WEV assuming that lender profits are only distributed to households in the top 0.1% of the earnings distribution.

7 Conclusion

During the 1960s, 1970s, and 1980s, the U.S. credit card industry engaged in various forms of non-competitive behavior. The industry repeatedly faced – and repeatedly lost – lawsuits brought by the Justice Department, Federal Trade Commission, and private parties. To study the consequences of these court rulings and the implications of subsequent competitive reforms, we relax the assumption of atomistic zero-profit lenders in workhorse consumer credit models. We propose a new model that incorporates oligopoly in the consumer credit market.

We use our estimated model to measure the gains from these reforms by simulating (i) increased lender entry with price collusion, and (ii) price competition. Moving from a monopoly to a collusive twenty lender oligopoly yields significant welfare gains, especially among low-income households. In the bottom decile of earnings, welfare gains from competitive reforms in the 1970s are equivalent to a one-time transfer worth \$3,400 (in 2016 dollars), or roughly 50% of their annual income. These gains are driven primarily by increased borrowing limits, leading to increased consumption as well as lower consumption volatility along the transition path. Transitioning to an oligopoly is worth four times less to the top decile of earners. We find similar distributional consequences from price competition. Further, the oligopoly with twenty lenders captures roughly 65% of the potential gains from competitive pricing.

We contribute a new theory – and quantification of that theory – to the mounting evidence that monopolies inflict great harm on low-income households (see, e.g. Schmitz (2016)). By integrating lender monopoly in the presence of heterogeneous consumers, we are able to explore the distributional consequences of monopoly. Despite its relatively small ex-ante size in the U.S. economy, we show that the welfare costs of monopoly in the U.S. credit card industry are large and disproportionately borne by low-income households. The primary driver of these gains is that greater competition spurs large expost increases in credit access. We also show that the competitive reforms that took place in the 1970s, in particular greater lender entry due to the *Marquette* decision, generated a significant fraction of competitive pricing gains.

While our article tackles several important issues, many questions remain. Does lender market power inhibit innovation and the adoption of new lending technologies? What type and magnitude of aggregate fluctuations would generate negative profits for lenders, and what types of policies would prevent lenders from exiting various segments of the credit card market? Are non-competitive credit card lenders limiting pass-through of monetary policy to households? We believe that our framework is tractable enough for future researchers to make progress on these questions as well as other important unanswered questions in the consumer credit literature.

References

Agarwal, S., S. Chomsisengphet, N. Mahoney, and J. Stroebel (2015). Regulating consumer financial products: evidence from credit cards. *Quarterly Journal of Economics 130(1)*, 111–164.

Allen, F. and D. Gale (2004). Competition and financial stability. *Journal of Money, Credit and Banking*, 453–480.

Athreya, K., J. Mustre-del Río, and J. M. Sánchez (2019). The persistence of financial distress. *The Review of Financial Studies forthcoming*.

Athreya, K., X. S. Tam, and E. R. Young (2012). A quantitative theory of information and unsecured credit. *American Economic Journal: Macroeconomics 4*(3), 153–83.

Auclert, A. and K. Mitman (2018). Bankruptcy and aggregate demand. In *2018 Meeting Papers*, Number 1085. Society for Economic Dynamics.

Ausubel, L. M. (1991). The failure of competition in the credit card market. *American Economic Review 81(1)*, 50–81.

Baqaee, D. R. and E. Farhi (2017). Productivity and misallocation in general equilibrium. NBER Working Paper 24007, National Bureau of Economic Research.

Bauducco, S. and A. Janiak (2015). On the welfare cost of bank concentration. Technical report.

Benetton, M. (2018). Leverage regulation and market structure: An empirical model of the uk mortgage market. *Available at SSRN 3247956*.

Berger, D. W., K. F. Herkenhoff, and S. Mongey (2019). Labor market power. Technical report, National Bureau of Economic Research.

Bizer, D. S. and P. M. DeMarzo (1992). Sequential banking. *Journal of Political Economy 100*(1), 41–61.

Boyd, J. H. and G. De Nicolo (2005). The theory of bank risk taking and competition revisited. *The Journal of Finance 60*(3), 1329–1343.

Braxton, J. C., K. Herkenhoff, and G. Phillips (2018). Can the unemployed borrow? implications for public insurance. *Working paper*.

Chakravarty, S. and E.-Y. Rhee (1999). Factors affecting an individual's bankruptcy filing decision. *Available at SSRN 164351*.

Chatterjee, S., D. Corbae, K. Dempsey, and J.-V. Rios-Rull (2020). A theory of credit scoring and the competitive pricing of default risk. *Manuscript*.

Chatterjee, S., D. Corbae, M. Nakajima, and J.-V. Ríos-Rull (2007). A quantitative theory of unsecured consumer credit with a risk of default. *Econometrica 75(6)*, 1525–1589.

Conesa, J. C., D. V. Costa, P. Kamali, T. J. Kehoe, V. M. Nygard, G. Raveendranathan, and A. Saxena (2018). Macroeconomic effects of medicare. *The Journal of Economics of Ageing 11*, 27–40.

Corbae, D. and P. D'erasmo (2011). A quantitative model of banking industry dynamics. *Manuscript, University of Wisconsin, Madison, and University of Maryland*.

Corbae, D. and P. D'Erasmo (2019). Capital requirements in a quantitative model of banking industry dynamics. Technical report, National Bureau of Economic Research.

Dempsey, K. and F. Ionescu (2019). Lending standards and consumption insurance over the business cycle. In *2019 Meeting Papers*, Number 1428. Society for Economic Dynamics.

Drozd, L. and J. Nosal (2008). Competing for customers: A search model of the market for unsecured credit. *Working paper*.

Drozd, L. and R. Serrano-Padial (2013). On prepayment and rollover risk in the u.s. credit card market. *Working paper*.

Drozd, L. and R. Serrano-Padial (2017). Modeling the revolving revolution: the debt collection channel. *American Economic Review 107(3)*, 897–930.

Edmond, C., V. Midrigan, and D. Y. Xu (2018). How costly are markups? NBER Working Paper 24800, National Bureau of Economic Research.

Egan, M., A. Hortaçsu, and G. Matvos (2017). Deposit competition and financial fragility: Evidence from the us banking sector. *American Economic Review 107*(1), 169–216.

Evans, D. S. and R. Schmalensee (2005). *Paying with plastic: the digital revolution in buying and borrowing*. MIT Press.

Galenianos, M. and A. Gavazza (2019). Regulatory interventions in consumer financial markets: The case of credit cards. *Available at SSRN 3405897*.

Galenianos, M. and J. Nosal (2016). Segmented markets and unsecured credit. *Working paper*.

Grodzicki, D. (2015). Competition and customer acquisition in the us credit card market. Technical report, Mimeo.

Grodzicki, D. (2019). The evolution of competition in the credit card market. *Working paper*.

Harberger, A. C. (1954). Monopoly and resource allocation. *The American Economic Review*, 77–87.

Hatchondo, J. C. and L. Martinez (2018). Credit risk without commitment. *Working paper*.

Herkenhoff, K. F. (2019). The impact of consumer credit access on unemployment. *The Review of Economic Studies 86*(6), 2605–2642.

Hunt, R. M. (2003). An introduction to the economics of payment card networks. *Review of Network Economics 2*(2).

Jamilov, R. (2021). A macroeconomic model with heterogeneous banks. Technical report.

Knittel, C. R. and V. Stango (2003). Price ceilings as focal points for tacit collusion: Evidence from credit cards. *American Economic Review 93*(5), 1703–1729.

Kovrijnykh, N., I. Livshits, and A. Zetlin-Jones (2019). Building credit history with heterogeneously informed lenders. Technical report.

Livshits, I., J. MacGee, and M. Tertilt (2007). Consumer bankruptcy: A fresh start. *American Economic Review 97(1)*, 402–418.

Livshits, I., J. MacGee, and M. Tertilt (2010). Accounting for the rise in consumer bankruptcies. *American Economic Journal: Macroeconomics 2(1)*, 165–193.

Livshits, I., J. C. MacGee, and M. Tertilt (2016). The democratization of credit and the rise in consumer bankruptcies. *The Review of Economic Studies*, rdw011.

Mateos-Planas, X. and G. Seccia (2006). Welfare implications of endogenous credit limits with bankruptcy. *Journal of Economic Dynamics and Control 30(11)*, 2081–2115.

Mateos-Planas, X. and G. Seccia (2013). A model of credit limits and bankruptcy. *Economics Letters 121(3)*, 469–472.

Mongey, S. (2019). Market structure and monetary non-neutrality. *Working paper*.

Nelson, S. T. (2018). Private information and price regulation in the us credit card market. Technical report, Mimeo.

Petrosky-Nadeau, N. and E. Wasmer (2013). The cyclical volatility of labor markets under frictional financial markets. *American Economic Journal: Macroeconomics 5*(1), 193–221.

Raveendranathan, G. (2020). Revolving credit lines and targeted search. *Journal of Economic Dynamics and Control 118*.

Romer, C. D. and D. H. Romer (2004). A new measure of monetary shocks: Derivation and implications. *American Economic Review 94*(4), 1055–1084.

Sánchez, J. M. (2018). The information technology revolution and the unsecured credit market. *Economic Inquiry 56(2)*, 914–930.

Schmitz, J. A. (2012). New and larger costs of monopoly and tariffs. *Federal Reserve Bank of Minneapolis Research Department Staff Report 468*.

Schmitz, J. A. (2016). The cost of monopoly: A new view. *Federal Reserve Bank of Minneapolis Region*.

Storesletten, K., C. Telmer, and A. Yaron (2000). Consumption and risk sharing over the life cycle. *NBER Working Paper* (w7995).

Storesletten, K., C. I. Telmer, and A. Yaron (2004). Consumption and risk sharing over the life cycle. *Journal of Monetary Economics 51*(3), 609–633.

The New York Times News Service, . (1991). Injunction against visa lets sears issue its cards. *The Baltimore Sun*, February, 26.

Wang, Y., T. M. Whited, Y. Wu, and K. Xiao (2018). Bank market power and monetary policy transmission: Evidence from a structural estimation. *Available at SSRN 3049665*.

Wasmer, E. and P. Weil (2004). The macroeconomics of labor and credit market imperfections. *American Economic Review 94*(4), 944–963.

White, G. (1992). 2 banks settle cardholder suit for 55 million. *Los Angeles Times*, April, 4.

Wieland, J. F. and M.-J. Yang (2020). Financial dampening. *Journal of Money, Credit and Banking 52*(1), 79–113.

Wildfang, K. C. and R. W. Marth (2005). The persistence of antitrust controversy and litigation in credit card networks. *Antitrust LJ 73*, 675.

Appendix

A Data

Our regression analysis in Section 2 is based on digitized archives of *Interest Rates Charged on Selected Types of Loans* (Form *FR 2835* and its variants), created and hosted at the Board of Governors. The micro-data is not public. Therefore, we provide detailed summary statistics in this appendix. Table 12 describes the main variables in our analysis.[21] Tables 13 and 14 provide information on interest rates and interest rate dispersion from 1975 to 1982. These data can be used to inform theories of credit scoring and rate dispersion in the early years of the credit card industry. The panel is quarterly and unbalanced. The data include roughly 200 banks per year. The data set records the lowest, highest, and most common charged interest rates on each bank's credit card plan (see Table 12).

Table 12: LIRS Variable Descriptions

Variable	Start date	End date	Description	Confidential?	Reporting Form
LIRS7812	2/12/1972	11/6/1982	LOWEST INTEREST RATE CHARGED FOR CREDIT CARD PLANS	No	FR 2835
LIRS7813	2/12/1972	11/6/1982	HIGHEST INTEREST RATE CHARGED FOR CREDIT CARD PLANS	No	FR 2835
LIRS7814	2/12/1983	8/6/1994	MOST COMMON INTEREST RATE CHARGED FOR CREDIT CARD PLANS	No	FR 2835

B Computational Algorithm

- Algorithm for steady state given set of credit lines $S_\theta = \{(\tau_1(\theta), \bar{l}_1(\theta)), \ldots, (\tau_N(\theta), \bar{l}_N(\theta))\}$
  1. Guess total profits $\Pi(S)$
  2. Given set of credit lines and total profits, solve consumer's problem through value function iteration
  3. Given policy functions, simulate economy and solve for terminal stationary distribution of $\Omega(i, \theta, \eta, \epsilon, a; S)$ where $\Omega : \{g, b\} \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \to [0, 1]$
  4. Given stationary distribution, update total profits
  5. Repeat 2.-4.until convergence
- Algorithm for transition given set of credit lines from $t = 1, ...T$ where T refers to the period for the terminal steady state.
  1. Solve for initial steady state given set of credit lines in $t = 1$
  2. Solve for terminal steady state given set of credit lines in $t = T$
  3. Guess sequence of aggregate total profits $\Pi(S)$ for transition path

[21]This is provided by the Board of Governor's Micro Data Reference Manual `https://www.federalreserve.gov/apps/mdrm/`

Table 13: Summary Statistics LIRS, 1975 to 1978

Variable	p25	p50	p75	Mean	Standard Deviation	Min	Max	Observations	Banks
1975									
Lowest Interest Rate Charged	12	16.5	18	15.25173	3.058655	6	18	330	217
Highest Interest Rate Charged	18	18	18	17.25482	1.862724	10	24	330	217
Most Common Interest Rate Charged	18	18	18	17.173	1.947327	10	24	330	217
Highest Minus Lowest Rate Charged	0	0	3	2.003091	2.896519	0	12	330	217
Quarterly Probability That Lowest Charged Rate Changes	0	0	0	0.036364	0.187478	0	1	330	217
Quarterly Probability That Highest Charged Rate Changes	0	0	0	0.018182	0.133811	0	1	330	217
Quarterly Probability That Most Common Charged Rate Changes	0	0	0	0.021212	0.14431	0	1	330	217
Fraction of Banks Reported No Difference Between Highest and Lowest Charged Rate	0	1	1	0.624242	0.485053	0	1	330	217
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 18%	0	0.5	1	0.5	0.500759	0	1	330	217
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 15%	0	0	0	0.048485	0.215115	0	1	330	217
1976									
Lowest Interest Rate Charged	12	16.48	18	15.29687	3.006702	6	18	584	176
Highest Interest Rate Charged	18	18	18	17.1907	1.797162	10	24	584	176
Most Common Interest Rate Charged	18	18	18	17.09481	1.893592	10	24	584	176
Highest Minus Lowest Rate Charged	0	0	3	1.893836	2.809243	0	12	584	176
Quarterly Probability That Lowest Charged Rate Changes	0	0	0	0.054795	0.227774	0	1	584	176
Quarterly Probability That Highest Charged Rate Changes	0	0	0	0.017123	0.129842	0	1	584	176
Quarterly Probability That Most Common Charged Rate Changes	0	0	0	0.023973	0.153095	0	1	584	176
Fraction of Banks Reported No Difference Between Highest and Lowest Charged Rate	0	1	1	0.638699	0.480789	0	1	584	176
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 18%	0	0	1	0.498288	0.500426	0	1	584	176
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 15%	0	0	0	0.059932	0.237563	0	1	584	176
1977									
Lowest Interest Rate Charged	12	15	18	15.16965	3.043804	6	18	601	169
Highest Interest Rate Charged	18	18	18	17.03484	1.960559	10	18	601	169
Most Common Interest Rate Charged	18	18	18	16.96058	2.031148	9	18	601	169
Highest Minus Lowest Rate Charged	0	0	3	1.865191	2.746045	0	12	601	169
Quarterly Probability That Lowest Charged Rate Changes	0	0	0	0.071547	0.257952	0	1	601	169
Quarterly Probability That Highest Charged Rate Changes	0	0	0	0.021631	0.145595	0	1	601	169
Quarterly Probability That Most Common Charged Rate Changes	0	0	0	0.036606	0.187948	0	1	601	169
Fraction of Banks Reported No Difference Between Highest and Lowest Charged Rate	0	1	1	0.63228	0.482586	0	1	601	169
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 18%	0	0	1	0.477537	0.499911	0	1	601	169
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 15%	0	0	0	0.049917	0.217954	0	1	601	169
1978									
Lowest Interest Rate Charged	12	18	18	15.28005	3.105733	6	18	637	177
Highest Interest Rate Charged	18	18	18	17.14498	1.876022	10	18	637	177
Most Common Interest Rate Charged	18	18	18	17.04292	1.954877	10	18	637	177
Highest Minus Lowest Rate Charged	0	0	3	1.864929	2.82426	0	12	637	177
Quarterly Probability That Lowest Charged Rate Changes	0	0	0	0.073783	0.261623	0	1	637	177
Quarterly Probability That Highest Charged Rate Changes	0	0	0	0.017268	0.130372	0	1	637	177
Quarterly Probability That Most Common Charged Rate Changes	0	0	0	0.032967	0.178691	0	1	637	177
Fraction of Banks Reported No Difference Between Highest and Lowest Charged Rate	0	1	1	0.643642	0.479299	0	1	637	177
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 18%	0	1	1	0.514914	0.50017	0	1	637	177
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 15%	0	0	0	0.036107	0.186702	0	1	637	177

Table 14: Summary Statistics LIRS, 1979 to 1982

Variable	p25	p50	p75	Mean	Standard Deviation	Min	Max	Observations	Banks
1979									
Lowest Interest Rate Charged	12	18	18	15.6337	2.949441	8	18	644	181
Highest Interest Rate Charged	18	18	18	17.27981	1.74522	10	18	644	181
Most Common Interest Rate Charged	18	18	18	17.12764	1.867032	10	18	644	181
Highest Minus Lowest Rate Charged	0	0	3	1.646118	2.633292	0	10	644	181
Quarterly Probability That Lowest Charged Rate Changes	0	0	0	0.069876	0.255136	0	1	644	181
Quarterly Probability That Highest Charged Rate Changes	0	0	0	0.017081	0.129673	0	1	644	181
Quarterly Probability That Most Common Charged Rate Changes	0	0	0	0.026398	0.160439	0	1	644	181
Fraction of Banks Reported No Difference Between Highest and Lowest Charged Rate	0	1	1	0.680124	0.466791	0	1	644	181
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 18%	0	1	1	0.569876	0.495478	0	1	644	181
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 15%	0	0	0	0.029503	0.169343	0	1	644	181
1980									
Lowest Interest Rate Charged	12	18	18	15.99783	2.838035	8	22	636	182
Highest Interest Rate Charged	18	18	18	17.53349	1.709019	11	24	636	182
Most Common Interest Rate Charged	18	18	18	17.36969	1.825551	11	24	636	182
Highest Minus Lowest Rate Charged	0	0	3	1.53566	2.664269	0	13	636	182
Quarterly Probability That Lowest Charged Rate Changes	0	0	0	0.121069	0.326464	0	1	636	182
Quarterly Probability That Highest Charged Rate Changes	0	0	0	0.051887	0.221973	0	1	636	182
Quarterly Probability That Most Common Charged Rate Changes	0	0	0	0.073899	0.261813	0	1	636	182
Fraction of Banks Reported No Difference Between Highest and Lowest Charged Rate	0	1	1	0.709119	0.454526	0	1	636	182
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 18%	0	1	1	0.578616	0.494169	0	1	636	182
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 15%	0	0	0	0.050314	0.218765	0	1	636	182
1981									
Lowest Interest Rate Charged	15	18	18	16.62226	2.601491	10	22	614	170
Highest Interest Rate Charged	18	18	18	17.94324	1.593033	12	30	614	170
Most Common Interest Rate Charged	18	18	18	17.76094	1.553755	12	24	614	170
Highest Minus Lowest Rate Charged	0	0	1.84	1.320977	2.483644	0	12	614	170
Quarterly Probability That Lowest Charged Rate Changes	0	0	0	0.112378	0.316088	0	1	614	170
Quarterly Probability That Highest Charged Rate Changes	0	0	0	0.07329	0.260824	0	1	614	170
Quarterly Probability That Most Common Charged Rate Changes	0	0	0	0.076547	0.266088	0	1	614	170
Fraction of Banks Reported No Difference Between Highest and Lowest Charged Rate	0	1	1	0.734528	0.441944	0	1	614	170
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 18%	0	1	1	0.600977	0.490097	0	1	614	170
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 15%	0	0	0	0.04886	0.215751	0	1	614	170
1982									
Lowest Interest Rate Charged	18	18	18	17.3569	2.598594	10	25.92	583	159
Highest Interest Rate Charged	18	18	19.5	18.61828	1.691271	10.5	25.92	583	159
Most Common Interest Rate Charged	18	18	19	18.5076	1.67938	10.5	25.92	583	159
Highest Minus Lowest Rate Charged	0	0	1.8	1.261389	2.498304	0	12	583	159
Quarterly Probability That Lowest Charged Rate Changes	0	0	0	0.149228	0.356619	0	1	583	159
Quarterly Probability That Highest Charged Rate Changes	0	0	0	0.128645	0.335094	0	1	583	159
Quarterly Probability That Most Common Charged Rate Changes	0	0	0	0.125214	0.331246	0	1	583	159
Fraction of Banks Reported No Difference Between Highest and Lowest Charged Rate	0	1	1	0.737564	0.440336	0	1	583	159
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 18%	0	1	1	0.538593	0.498936	0	1	583	159
Fraction of Banks Reporting Highest and Lowest Charged Rate Equal to 15%	0	0	0	0.010292	0.101011	0	1	583	159

4. Given set of credit lines and sequence of total profits, solve consumer's problem in each period through backward induction starting from $T-1$
5. Given policy functions from previous step and distribution of consumers from initial steady state, simulate economy and solve for new sequence of aggregate profits
6. Update guess for sequence of total profits
7. Repeat 3.-6. until convergence

- Solve for monopoly, collusive-Cournot, Stackelberg-Cournot, and perfectly competitive pricing
  - Monopoly:
    1. Define grid on spreads and borrowing limits: $(\tau_1(\theta), \bar{l}_1(\theta)), \ldots, (\tau_N(\theta), \bar{l}_N(\theta)) \in (\mathbb{R}_+, \mathbb{R}_+)^N$
    2. For each point on grid for spreads and borrowing limits, solve for steady state
    3. Pick spreads and borrowing limits that maximize total profits across steady states
  - Collusive-Cournot
    1. Define grid on spreads for each type θ
    2. Given number of lenders N and spread $\tau(\theta)$, solve for best response limit function
       * Define grid on limits for every θ. Let each limit on this grid denote the identical limit of players $\{1, 2, ..., N-1\}$
       * For each spread and grid point on limits, solve for profit maximizing limit of player N given $\tau(\theta)$ and sum of limits of players $\{1, 2, ..., N-1\}$. For this step, the lenders internalize the transition path for each set of credit lines. The initial steady state is the equilibrium outcome from the monopolist problem.
    3. For every spread $\tau(\theta)$, using best response limit function from previous step, use Bisection to solve for the symmetric Nash equilibrium limit (stage 2 outcome). This is the point where the best response function intersects the 45 degree line.
    4. Given stage 2 outcomes, pick the spreads that maximize profits in stage 1 for every θ
  - Stackelberg-Cournot duopoly
    1. Define grid on spreads for both players for each type θ
    2. For every spread of first mover $\tau^1(\theta)$ and spread of second mover $\tau^2(\theta)$, solve for stage 2 best response limit function of both the players
       * Define grid on limits for every θ for first mover
       * For each spread combination and grid point on limits from previous step, solve for profit maximizing limit of the second mover. This gives the best response limit function of the second mover. For this step, the both lenders internalize the transition path for each set of credit lines. The initial steady state is the equilibrium outcome from the monopolist problem. The best response limit function for the first player is analogous.

3. For every spread $\tau^1(\theta)$ and $\tau^2(\theta)$, given best response limit functions from previous step, solve for the Nash equilibrium limit (stage 2 outcome)
4. Given stage 2 outcomes and $\tau^1(\theta)$, second mover picks the spread that maximizes their net present value of profits in stage 1
5. First mover picks spread that maximizes their net present value of profits given second mover's spread best response function and Nash equilibrium limits

– Perfectly competitive pricing

1. Define grid on spreads and borrowing limits: $(\tau_1(\theta), \bar{l}_1(\theta)), \ldots, (\tau_N(\theta), \bar{l}_N(\theta)) \in (\mathbb{R}_+, \mathbb{R}_+)^N$
2. Solve for spreads and borrowing limits that maximize the welfare of unborn agent (given θ) subject to weakly positive profits

C Stackelberg-Cournot: Second mover's policy functions

This section provides further economic intuition for the equilibrium discussed in Section 5.5. We discuss the policy functions of the second mover given the first mover's optimal spread $\tau^1 = 0.64\%$.

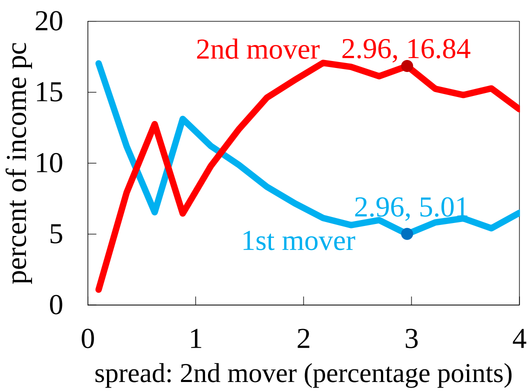
Panel (a) of Figure 15 plots stage 2 Nash borrowing limits for both lenders as a function of the second mover's spread fixing the first mover's optimal spread (0.64%). Panel (b) plots the net present value of profits for both lenders given the first mover's optimal spread, the second mover's spread, and the Nash borrowing limits associated with the respective spreads. As panel (b) shows, the second mover maximizes profits by setting their spread to 2.96% (red marker). Panel (a) shows that for the region where the second mover sets a spread higher than that of the first mover ($\tau^2 > 0.64\%$), the second mover's Nash limit initially increases and then decreases. Their profits are maximized in the region where their limit starts to decrease. By choosing a high spread, the second mover dampens their own profits because they will attract a riskier pool of borrowers. However, they benefit by forcing the first mover with a lower spread to set a low limit. Hence, the higher spread of the second mover allows them to set a higher limit in comparison to the first mover. However, they optimally choose a spread so high that they also restrict their own limit. This mechanism is consistent with what we saw for the monopolist in panel (a) of Figure 5 and collusive-Cournot for panel (b) of Figure 6. The optimal borrowing limit is hump-shaped as a function of the spread.

D Lender policy functions for the low type

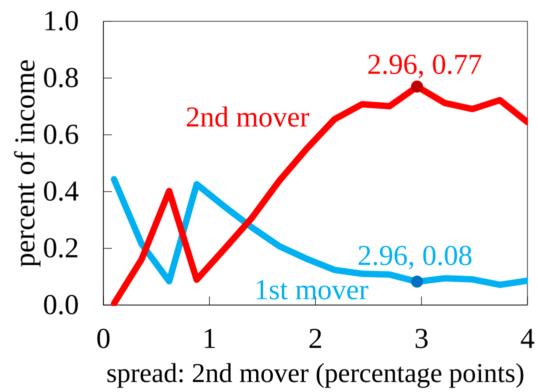
This section presents lender policy functions for the low-type that are analogous to those presented in Figures 6 and 13, which were for the high-type. Figure 16 presents the policy functions in the case of collusive-Cournot duopoly. Figure 17 presents the policy functions in the case of a Stackelberg-Cournot duopoly. In both cases, the policy functions are qualitatively similar to those of the high-type. The difference is that the low-type receives a lower limit and a higher spread.

Figure 15: Stackelberg policy functions given 2nd mover spread at 1st mover optimal spread $(\theta = \theta_H)$

(a) Nash borrowing limit in stage 2



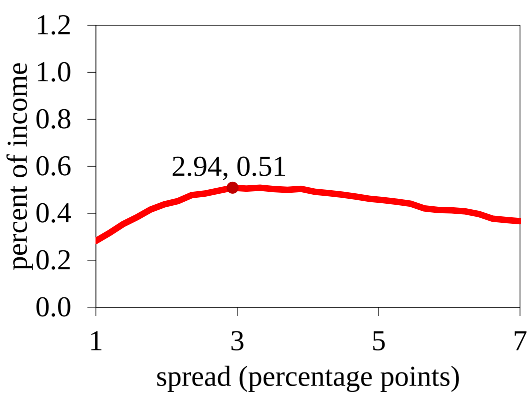
(b) Net present value of profits in stage 1



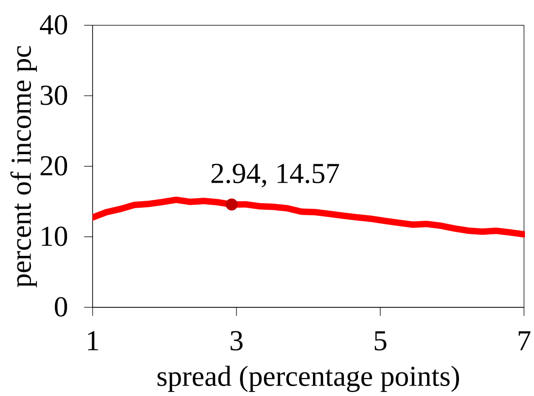
Notes: Panel (a) plots stage 2 Nash borrowing limits for both lenders as a function of the second mover's spread fixing the first mover's optimal spread (0.64%). Panel (b) plots the net present value of profits for both lenders given the first over's optimal spread, the second mover's spread, and the Nash borrowing limits associated with the respective spreads.

Figure 16: collusive-Cournot policy functions ($\theta = \theta_L$)

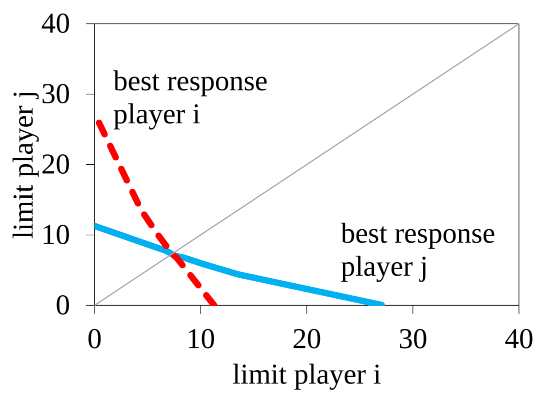
(a) Net present value of profits



(b) Nash borrowing limit in stage 2

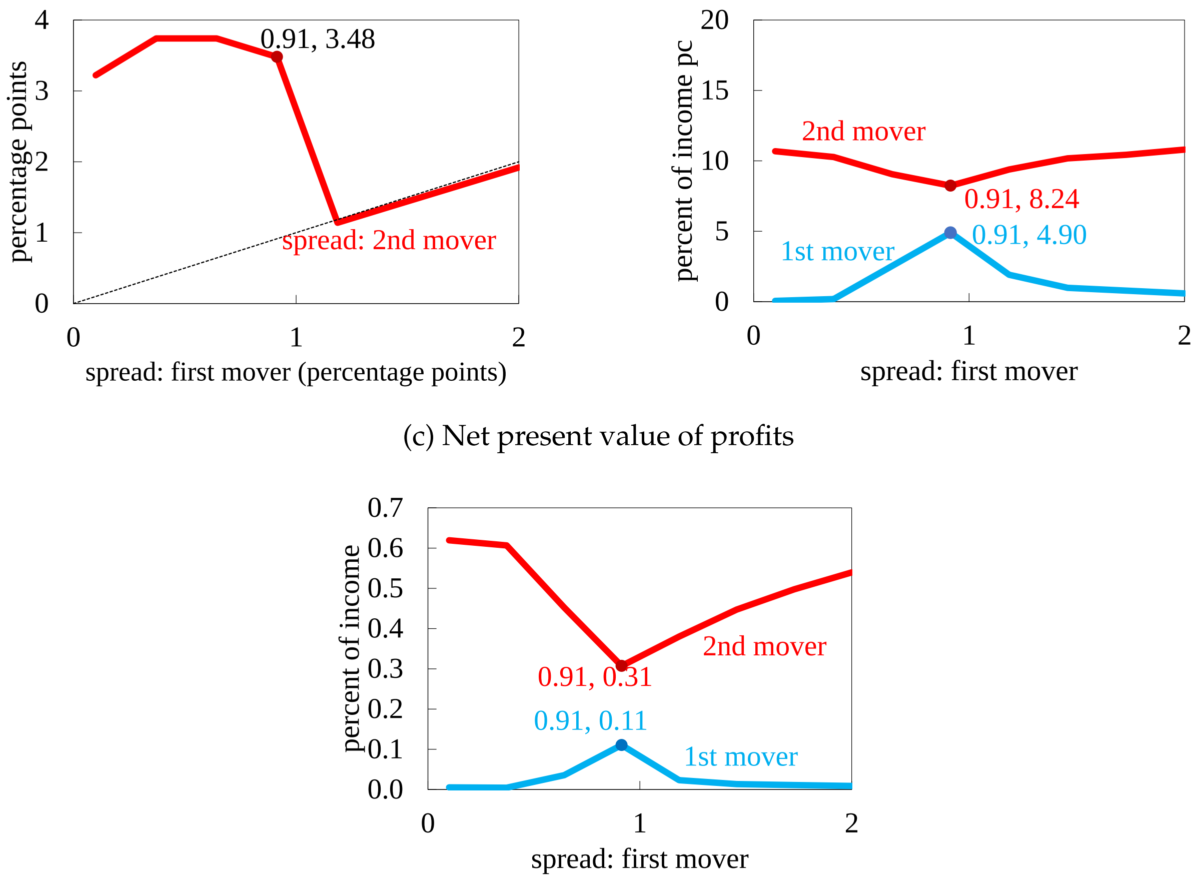


(c) Borrowing limit best response in stage 2



Notes: Panel (a) plots the net present value of lender profits as a function of the spread in the first stage. Lenders collude in the first stage to set the spread that maximizes profits. Panel (b) plots the stage 2 Nash equilibrium limit as a function of the spread in the first stage. Panel (c) plots the lenders' best response functions in the second stage for the spread that maximizes profits in the first stage. The symmetric Nash equilibrium is the point where the limits cross the 45 degree line.

Figure 17: Stackelberg policy functions given 1st mover spread ($\theta = \theta_L$)



(a) 2nd mover spread in stage 1

(b) Nash borrowing limits in stage 2

(c) Net present value of profits

Notes: Panel (a) plots the second mover's best response in spreads in stage 1 as a function of the first mover's spread. Panel (b) plots the stage 2 Nash borrowing limits for both lenders given the first mover's spread and the second mover's best response in spreads. Panel (c) plots the net present value of profits for both lenders given the first mover's spread, the second mover's best response in spreads, and Nash borrowing limits for both lenders given their respective spreads. Spreads are expressed as percentage point deviations over the risk-free rate. Borrowing limits are expressed as a percentage of income per capita. Net present value of profits are expressed as a percentage of income.

E Model economies for robustness exercises

This section presents the modifications made to the baseline model for the robustness exercises (Section 6). The first robustness exercise allows for fixed costs of lender entry. We compute the fixed cost as equal to the net present value of profits for lenders $k \in \{2, ..., N\}$, which serves as an upper bound. The total fixed cost is computed as follows:

$$F = \sum_{\theta \in \Theta} \sum_{k=2}^{N} \sum_{t=1}^{\infty} \left(\frac{1}{1 + r_f} \right)^t \pi_t^k(\theta), \tag{12}$$

where F denotes the total fixed cost. F is equally distributed among households in the first period of the transition. This leads to a change in the budget constraint of the household in the first period of the transition. For example, the budget constraint of the consumer in good standing who chooses to repay is given by the following equation:

$$c + a' = \theta\eta\epsilon + (1 + r_f)a + \sum_{j=1}^{N} (r_j(\theta) - r_f) a_j(a, \theta) + \Pi - F. \tag{13}$$

The modifications to the budget constraints for consumers in good standing who default or consumers in bad standing are analogous. The rest of the baseline model remains unchanged.

Our second robustness exercise allows for an alternate redistribution of lender profits to high-earning households. We assume that lender profits are only rebated to households in the top 0.1% of the earnings distribution. This leads to a change int the budget constraint for the consumer depending on whether they belong to the top 0.1% of the earnings distribution or not. The budget constraint for the consumer in good standing who chooses to repay is given by the following equation:

$$c + a' = \theta\eta\epsilon + (1 + r_f)a + \sum_{j=1}^{N} (r_j(\theta) - r_f) a_j(a, \theta) + 1_{\{\theta\eta\epsilon \geq P99\}} \frac{\Pi}{0.01}, \tag{14}$$

where $1_{\{\theta\eta\epsilon \geq P99\}} \dfrac{\Pi}{0.01}$ is an indicator function equal to 1 if total earnings $\theta\eta\epsilon$ is greater or equal to the 99th percentile of the earnings distribution ($P99$) and 0 otherwise. The modifications to the budget constraints for consumers in good standing who default or consumers in bad standing are analogous. The rest of the baseline model remains unchanged.

F Gains from re-optimization along transition path

Lenders commit to strategies at date 1 along the transition path, fully understanding the future movement of the distribution of agents across states. In this section, we conduct an exercise that allows us to assess whether lenders would potentially gain from re-optimization. To measure the approximate gains from re-optimization, we compare the $t = 1$ transition path strategies to the strategies that lenders choose in steady state. Despite lenders committing to strategies at date 1 along the transition path, Table 15 shows that if lenders are able to reoptimize across steady states, they would choose very similar strategies. We begin with the $N = 2$ collusive-Cournot case. Consider high-type consumers. Columns (1) and (2) show that optimal borrowing limits are 26.50% of income when lenders reoptimize and commit at $t = 1$ versus 26.31% when lenders optimize across steady states. Columns (1) and (2) imply spreads of 2.55% when lenders reoptimize and commit at $t = 1$ versus 2.94% when lenders optimize across steady states. The same is true for the low-type. Columns (3) and (4) yield very similar results for the $N = 20$ collusive-Cournot. Optimization across steady states yields similar strategies suggesting small gains from re-optimization later in the transition.

Table 15: Comparison of steady state outcomes

Summary Statistics	**(1)** collusive-Cournot (N=2) transition	**(2)** collusive-Cournot (N=2) steady state	**(3)** collusive-Cournot (N=20) transition	**(4)** collusive-Cournot (N=20) steady state
High-type				
Borrowing limit to initial income pc	26.50	26.31	37.83	38.56
Spread	2.55	2.94	2.74	2.16
Low-type				
Borrowing limit to initial income pc	14.57	14.57	21.28	21.52
Spread	2.94	3.32	2.74	2.35

Notes: Table reports optimal transition path strategies in Columns (1) and (3), as well as the optimal interest rate and limits in the new duopoly steady state (Column (2)) and oligopoly steady state (Column (4)).