Three Essays in Empirical Corporate Finance

Poorya Kabir

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2020

Abstract

Three Essays in Empirical Corporate Finance

Poorya Kabir

This dissertation presents three essays in empirical corporate finance. The essays discuss how financial markets affect the real economy. The first essay studies how a change in credit supply affects firms' decisions to create new products or destroy the existing ones. It provides reduced form causal evidence that a reduction in credit supply reduces product creation substantially. The second essay studies the effect of less product creation on consumer welfare. I find that the effect on consumer welfare is smaller relative to a "naive" interpretation of the reduced form estimate, due to equilibrium responses. The third essay studies how financially constrained firms reduce total investment costs. It provides suggestive evidence that when reducing total investment cost, they do so by lowering the expansion of output capacity and choosing cheaper investment options.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I owe gratitude towards my advisors Charles Calomiris, Olivier Darmouni, Xavier Giroud, David Weinstein, and Daniel Wolfenzon. Their mentorship, enthusiasm for research, and support have been instrumental in completing this dissertation.

I am also thankful to my friends and, in particular, Mohammadreza Bolandnazar, Seyed Mohammad Mansouri, and Roya Arab Loodaricheh. There was less laughter without them, and the Ph.D. wouldn't have been the same.

I am grateful to my family members. First and foremost, to my loving parents, Parvin Lotfi and Mohammadreza Kabir, to whom this dissertation is dedicated. Their unconditional love and support have been key in keeping me push forward. And second, to my brother Pooyan, who has always been supportive and encouraging.

Dedicated to my beloved parents

Mohammadreza Kabir and Parvin Lotfi

Chapter 1

Effect of Credit Supply on Product Creation

Poorya Kabir

I show that firms that face a reduction in credit supply reduce product creation, by using variation from US banks' exposure to the mortgage market to instrument for credit supply. The magnitude is substantial: firms facing a one-standard-deviation decrease in credit supply offered 10% fewer products. The reduction in product offering is permanent and takes six quarters to realize fully. Furthermore, I show that the reduction in product offerings derives from the limited creation of new products rather than the destruction of existing ones. Moreover, I show that the results are robust to alternative outcome variable definitions, alternative weighting methods, [1]

[1]I am grateful to Olivier Darmouni, Xavier Giroud, and David Weinstein for their invaluable advice during this project. I am also grateful to Charles Calomiris, Tano Santos, Patrick Bolton, Daniel Wolfenzon, Neng Wang, Kairong Xiao, Tania Babina, Gur Huberman, Giorgia Piacentino, Matt Backus, and Kent Daniel for their useful suggestions. Researcher's own analyses are calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

1.1 Introduction

Understanding consumer well-being is of interest to both policymakers and academics in general, and especially so during financial crises since the marginal utility of consumption is higher. An understudied dimension of consumer well-being during financial crises is the set of available products to choose from. The availability of products depends on firms' decisions to create/destroy products. In this paper, I study how credit supply, an essential function of financial markets that is disrupted during financial crises, affects firms' decisions to create/destroy products.

I study the question in the context of the consumer packaged goods market (the market for goods sold in grocery stores, drug stores, and mass merchandise stores). Why do we care about product creation in this market? First, it is a large component of household expenses. It accounts for 40% of consumption of goods, and 16% of total household expenditure. Second, product creation benefits both firms and consumers. Consumers benefit from the creation of new products since it either caters to heterogeneous tastes or is a quality upgrade. Firms benefit from new products by becoming more productive and profitable.

I combine two micro-level datasets and use US banks' exposure to the mortgage market as an instrument for credit supply. The first dataset is the Nielsen Home Scan Dataset (HMS), which includes micro-level data on millions of different products offered in the consumer packaged goods market. It is used to identify products offered by firms. The second dataset is the Dealscan dataset, which includes information about loans made to firms by a syndicate of lenders. It includes the names of the lender and the borrower, and is used to construct changes in credit supply.

In order to causally identify the effects of credit supply on product creation, I use changes in lending by a firm's pre-crisis lenders as a measure of the change in credit supply, and complement this with three different instrumental variable strategies following Chodorow-Reich (2013). The three instruments are: Lehman exposure (the share of loans each bank co-syndicated with Lehman), ABX exposure (banks' exposure to the toxic mortgage-backed securities market measured through the daily correlation between a bank's stock market return and the ABX return), and bank statement items (trading revenue plus deposits divided by total assets). The validity of the instruments depends on two things: a strong first-stage, which is validated in the data, and exclusion restriction. I provide evidence consistent with exclusion restriction by showing that the instruments are uncorrelated with a large set of firm, loan, and, most importantly, product and consumer characteristics. Consumer and product data are particularly useful for studying the effect of credit supply, since the main alternative explanation for the results is differences in demand. Consistent with the identification assumption, I don't find differences in consumer and product characteristics between differentially affected firms.

I find that firms that face a reduction in credit supply reduced their product creation. A one-standard-deviation decrease in credit supply results in a 10% reduction in product creation. The effect is stronger for ex-ante more financially constrained firms. I measure financial constraints as being private, not having access to bond markets, and either having fewer products or lower sales (both proxies for size). Additionally, the effect is zero before Lehman's failure, which is consistent with the "parallel trends" assumption. The effect becomes negative and significant, then increases for a few quarters and becomes stable, after credit market conditions improve. Furthermore, I show that the reduction in product offering comes from the limited creation of new products rather than the destruction of existing ones.

I do further robustness checks to make sure that the results are robust to alternative assumptions. First, I find similar results using alternative definitions of the left-hand-side variable. Second, I confirm that the results are similar using different weighting methods for observations. Third, if financial constraints derive results, we should expect to see it across different departments within the consumer packaged goods market. I show that the results hold across various departments. Fourth, I repeat the results using the Nielsen Retailer Scanner dataset, which is collected at *store* level, rather than consumer-level. The benefit of using this data is that each product roughly appears 1,000 times more, and thus, we need to worry less about a product with missing data in a quarter. Fifth, the definition of the product used in the primary analysis is very granular. For instance, in the main specification, two bottles of ketchup sauce offered by the same brand with different sizes would be considered two different products. I repeat the result with two coarser definitions of products and find similar results. Thus, the result is not driven by less creation of marginally different products.

Related work: This paper contributes to several strands of literature. It contributes to the literature that studies how financial crises, in particular the Great Recession, affected household consumption (Mian, Rao, and Sufi (2013), Di Maggio, Kermani, Keys, Piskorski, Ramcharan, Seru, and Yao (2017), Mian, Sufi, and Trebbi (2015), Mian and Sufi (2012), Agarwal, Chomsisengphet, Mahoney, and Stroebel (2017), Benmelech, Meisenzahl, and Ramcharan (2017)). While this literature focuses on consumer demand measured by spending, this paper highlights the importance of supply-side decisions made by firms regarding the availability of products.

This paper adds to the literature that studies the effects of disruptions in financial markets and reduction in credit supply on real economic outcomes. Previous papers study the effects of disruptions in financial markets on employment (Chodorow-Reich (2013), Giroud and Mueller (2017), Giroud and Mueller (2019), Greenstone, Mas, and Nguyen (2014), Huber (2018)), investment (Almeida, Campello, Laranjeira, and Weisbenner (2012), Amiti and Weinstein (2018)), pricing (Gilchrist, Schoenle, Sim, and Zakrajšek (2017), Kim (2018), Borenstein and Rose (1995), Stroebel and Vavra (2019), Chevalier and Scharfstein (1994), Chevalier and Scharfstein (1995)), likelihood of default (Khwaja and Mian (2008)), etc. I add to this literature by documenting a new channel through which credit supply affects real outcomes.

This paper fits into a large collection of literature in economics that uses barcode-level data to study consumer and firm behavior (for example: Faber and Fally (2017), Jaravel

(2018), Hottman, Redding, and Weinstein (2016a), Redding and Weinstein (2016), etc). The most closely related paper is Broda and Weinstein (2010). They document that product creation is pro-cyclical, while product destruction is weakly counter-cyclical. While they suggest that their results could be demand-driven, as prior theoretical work by Schmookler (1962) and Shleifer (1986) suggests, they do not make a causal claim. I contribute to this literature by documenting a supply-driven product creation mechanism.

This paper complements the financial economics literature that studies the impact of finance on growth in general (Rajan and Zingales (1998), Jayaratne and Strahan (1996)) and on innovation (Acharya and Xu (2017), Hsu, Tian, and Xu (2014) and Amore, Schneider, and Žaldokas (2013)). While patents capture both product and process innovation, new products are a direct measure of product innovation.

1.2 Data

I use the Dealscan dataset to construct the measure of change in the supply of credit. It includes information on syndicated loans, which are typically large loans given to firms by financial institutions. Syndicated loans are often originated by one or more lenders (also known as lead lenders) and accompanied by other lenders which act like investors in the loan (also known as participants). The unique feature of Dealscan relative to other firm-level financial information datasets is that it includes borrower and lender names for each loan. This information is crucial in constructing the measure of change in the supply of credit. The dataset also includes other information about the loan, such as amount, rate, maturity, type (term loan or line of credit), purpose, role of lender (lead or participant), etc.

I use the Nielsen HMS dataset to identify the number of different products offered by firms[2]. Each year Nielsen keeps track of purchases made by a representative sample of US households numbering between 40,000 and 60,000. These purchases are made at grocery

[2]The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

stores, drug stores, mass merchandise stores, and similar retail outlets across the United States. Each household uses a scanner to record the price, quantity, and Universal Product Code (UPC) of items purchased. The UPC is a 12-digit number, and the barcode is a machine-readable version of that number[3]. The first 6 to 10 digits of the UPC function as an identifier assigned to a firm, and are used to merge Nielsen HMS data with the GS1 Data Hub containing the name of the firm associated with the identifier. The data span a twelve-year period from 2004 to 2015, including purchases associated with over 2 million different UPCs.

There are several advantages to using the Nielsen dataset to identify firms' products. First, the barcode is the most granular definition of a product in the retail market[4]. Second, product entry and exit can be identified in the data at high frequency, which is helpful in understanding the dynamics of firm response. Additionally, it enables me to do a placebo test at high frequency. Third, each UPC belongs to a specific department, product group, and product module (for example, a can of caffeine-free diet Coke belongs to the low-calorie soft drink product module, the carbonated beverage product group, and the dry grocery department). This is useful in controlling for differences in demand across different product groups. Fourth, the data contain information on the characteristics of the households, which is used in addressing some identification challenges. Furthermore, consumer characteristics

[3]10 barcodes can be obtained from GS1 for an initial fee of $250 and an annual license renewal fee of $50. There is a significant discount associated with more barcodes. For more information on pricing please look at https://www.gs1us.org/education-training/frequently-asked-questions/gs1-company-prefix-barcodes-and-identification.

[4]Broda and Weinstein (2010) provide a detailed explanation of why it is reasonable to assume each UPC represents a different product. "Although it is difficult to enforce how a company uses a bar code, most industry experts strongly caution firms not to use the same bar code on more than one product. Doing so could cause confusion among retailers who would have trouble knowing what they were selling and for consumers whose receipts would not match their actual purchases. Similarly, firms typically do not use multiple UPCs for the same product because that makes it very difficult for retailers to reorder out of stock items. As a result, manufacturers tend to use other bar code systems for internal use and reserve the UPC for tracking products that are identical to the consumer. Therefore, it is reasonable to assume that all goods with different UPCs differ in some way that might cause consumers to pay a different price for them and that it is rare for a meaningful quality change to occur that does not result in a change of UPC. For example, changing the slogan on a Heinz ketchup bottle does not require a new bar code, but changing the size of the bottle does. In other words, it is safe to assume that if the bar code changes, it is likely that some noticeable characteristic of the product has changed."

data are useful in understanding firm response to different consumer clienteles.

Firms in the Dealscan and Nielsen datasets don't have common identifiers. I use the Orbis online database to match firms with their Orbis identifiers for both databases, and I use the matched identifier to merge firms from the Nielsen and Dealscan datasets. Additionally, I use the Fixed Income Securities Database (FISD) to identify firms that have access to the bond market.

Table 1.1 reports summary statistics of the merged sample. The merged data include 175 firms. Most firms have products in multiple product groups. The number of firm-product group observations is 1152. On average, firms offer about 42 different UPCs in a product group, with an average price of 3.3 dollars. Among Nielsen firms, firms in the merged sample are exceptionally large. Though they consist of less than 1% of the Nielsen firms, they sell 13% of the products. Despite the fact that the merged sample is not representative of most Nielsen firms, the effect of reduction in credit supply is likely to be greater for the other firms, since they are smaller.

1.3 Empirical Methodology

1.3.1 Construction of Credit Supply Shock

I construct the measure of change in credit supply by following Chodorow-Reich (2013). During the financial crisis of 2007-2008 there was a significant drop in bank lending. However, the drop was not uniform across all banks. Chodorow-Reich (2013) compares firms borrowing from banks that had differing levels of reduction in lending. Firms borrowing from banks that reduced lending more, faced tighter credit conditions relative to their counterparts. Below, I describe how the measure is constructed.

Formally, let $L_{b,j,t}$ be equal to one if bank b makes a loan to firm j at time t, and zero otherwise. Let $s_{b,j,t}$ denote the imputed loan share of bank b in the loan made to firm j at time t[5]. Define

[5]Some loan shares are not reported in Dealscan. So, I use data on loans with a similar structure to

$$\Delta L_{-f,b} = -\frac{\sum_{j\neq i} s_{b,j,crisis} L_{b,j,crisis}}{0.5 \sum_{j\neq i} s_{b,j,normal} L_{b,j,normal}}$$

where "crisis" is the 9-month period between October 2008 and June 2009, and "normal" represents the 18 months from October 2005 to June 2006 and October 2006 to June 2007. In 2007 signs of problems in the mortgage market were beginning, thus I exclude October 2007 to June 2008 from the normal period. During the crisis, bank lending dropped significantly. Ivashina and Scharfstein (2010) document that in the fourth quarter of 2008, the number and dollar volume of loans were, respectively, 33% and 47% lower compared to the previous quarter. $\Delta L_{-f,b}$ captures the change in the number of loans made by a bank during the financial crisis, relative to before it.

To construct the measure of change in credit supply for each borrower, I aggregate the bank-level measures of credit supply ($\Delta L_{-f,b}$) by using share in the last pre-crisis loan as weights. I denote the measure of change in credit supply for firm f with ΔL_f, then:

$$\Delta L_f = \sum_{b\in S} s_{b,f,last} \Delta L_{-f,b}$$

where S denotes banks that participated in the last pre-crisis syndicate, and $s_{b,f,last}$ is the share that bank b had in the last loan that firm f received. Intuitively, ΔL_f measures the average change in lending by banks that lent to the firm prior to the crisis. There are two reasons two assume ΔL_f captures changes in availability of credit for a firm. First, if firms can substitute their pre-crisis lenders with new lenders at low cost, this measure would not capture changes in credit supply, since a firm borrowing from an unhealthy lender can switch to a healthier lender. However, as documented by Darmouni (2019) and Chodorow-Reich (2013), the firm-bank relationship is sticky, and substituting banks is costly for firms. Second, Dealscan records loans at the time of initiation. If loan covenants are violated, banks have

impute loan shares for loans with missing information on loan shares.

the option of renegotiating the existing loans. Chodorow-Reich and Falato (2017) document that banks facing a shortage of liquidity are more likely to reduce loan commitment to an existing loan, when covenants are violated. Thus, the health of previous lenders matters, even for existing loans.

ΔL_f measures the change in credit supply only if demand for credit across firms borrowing from the same bank is uncorrelated. Otherwise, ΔL_f measures both firm demand and bank credit supply. This will bias the regression coefficient of the effect of credit supply on product creation. For example, assume that some firms face less demand for product creation during recessions, and thus have lower demand for credit. If these firms borrow from the same bank, reduced bank lending is correlated with a firm's decision to reduce product creation; however, one cannot assume the correlation is causal. Similar firms borrowing from the same bank can bias the regression coefficient upward or downward. The aforementioned example represents a downward bias. However, it could be that firms with reduced demand for product creation will increase borrowing effort in order to be able to operate during recessions. Thus we see an increase in demand for loans coming from firms that reduce product creation. Ex ante it is not clear which one of these biases dominates the estimated coefficient. Consequently, I complement the credit supply measure with three different instrumental variable strategies following Chodorow-Reich (2013).

The first instrumental variable is the share of loans that each bank co-syndicated with Lehman (Lehman instrument). Ivashina and Scharfstein (2010) show that firms that had Lehman as one of their lenders started drawing down their credit lines after Lehman's failure in September of 2008. As a result, banks that co-syndicated more loans with Lehman faced a shortage of liquidity relative to their counterparts. Consequently, they reduced more of their lending. Since Lehman's failure was unexpected, the share of co-syndicated loans with Lehman is likely to satisfy the exclusion restriction.

The second instrument measures the exposure of banks to toxic mortgage-backed securities (ABX instrument). Bank exposure to these assets is inferred from the correlation of the daily stock return of a bank with the ABX AAA 2006-H1 index between October 2007 and December 2007. The index follows the price of residential mortgage-backed securities. Banks that were more exposed to the mortgage crisis reduced more of their lending during the financial crisis. Even though mortgage and commercial lending occur via the same bank, losses in the mortgage market are likely to be uncorrelated with the health of the bank's corporate clients, since the recession originated with the housing market. Thus, exposure to mortgage-backed securities is likely to be a valid instrument.

The third instrument uses trading revenue plus bank deposits divided by total assets (Bank items instrument). On a bank balance sheet, trading revenue is where losses from mortgage-backed securities are written down. The logic of this instrument is similar to that of ABX exposure.

Instrumental variables need to satisfy two conditions to be plausible. First, instruments should have a strong first stage. In Table 1.5, I show that the results of first-stage F-statistics are greater than 10, which alleviates concerns about weak instruments. Second, instruments should satisfy the exclusion restriction. In order for instruments to satisfy the exclusion restriction, I need to show that the only channel through which the instrument is correlated with product creation is ΔL_f. I provide three pieces of evidence consistent with this assumption. First, I do a placebo test in the next section and show that the estimated coefficient is zero before Lehman's failure. Second, I show that the instruments are not correlated with a large set of observables. Table 1.2 shows the regression coefficient and p-value of a regression where the left-hand side is the Lehman instrument and the right-hand side is the variable in the table[6]. These covariates can be divided into four categories: firm (including data on whether the firm is in manufacturing, whether the firm has access to the bond market, and whether the firm is public), loan (including data on the loan amount of the last loan prior to the crisis, whether the firm had multiple lead lenders, the maturity of the last loan before the crisis, whether the loan is a term loan or revolving loan),

[6]Similar tables for the other two instruments are available in the online appendix.

product (including data on the number of products prior to the crisis, firm sales prior to the crisis, and the average price of the firm's products), and consumer (including data on the household income of the firm's consumers and the household size of the firm's consumers). The conventional alternative explanation in the literature of the effects of credit supply on firm decisions is that differences in demand for credit cause differences in demand for loans. However, the advantage of having product- and consumer-level controls is that in this paper, it allows me to ascertain whether there are differences among consumers and among products of differentially exposed firms. The table shows the p-values are above the 0.05 threshold, which insures that the instrument is not correlated with firm observables. Additional evidence consistent with a balance of covariates comes from the main regression specification results being the same, with or without product group fixed effects.

Similar to the balance of covariates table for the Lehman instrument, in Table 1.3 and 1.4, I check if the ABX and bank statement items instrument are correlated with firm, loan, consumer and product characteristics. Both tables confirm balance of covariates across these variables.

The most likely alternative explanation for the findings in this paper is that differences in firm demand cause firms to reduce product creation and borrowing at the same time. Nielsen's detailed product-level data is especially helpful in alleviating concerns about demand. While most studies control for firm-level variables, I also control for product groups. In the regression specification below, I control for product group by adding product group fixed effects. The inclusion of product group fixed effects means I compare two firms producing in the *same product category*. Thus demand at the level of product group cannot influence the results. Additionally, when the same regression is run without product group fixed effects the result is unchanged, which is further evidence consistent with a balance of covariates.

I determine the number of products by counting the number of UPCs that have positive sales in at least one of the following eight consecutive quarters. The entry date for a product is the first date that the product has positive sales. The exit date is the first date that the product has zero sales for the following eight consecutive quarters[7]. I assign the exit date in this way because some products have zero sales for a few quarters and positive sales afterwards. Since the product will have positive sales at a later date, I assume that the product is still being offered in quarters that zero sales are recorded. This step implicitly assumes that a firm is producing a product even if it doesn't appear in the dataset during a particular quarter[8].

1.3.2 Regression Specification

I use the following regression specification to measure the effect of change in credit supply on the number of products offered by firms:

$$\Delta \ln(1 + N_{fg}) = \alpha_g + \beta \Delta L_f + \gamma X_{fg} + \varepsilon_{fg} \tag{1.1}$$

where N_{fg} is the number of products offered by firm f in product group g[9], Δ is the difference in number of products in post-period 2008:Q4 - 2009:Q3 relative to pre-period 2007:Q2 - 2008Q1, β is the coefficient of interest and measures the effect of change in credit supply on net product creation, ΔL_f is the change in credit supply for firm f, α_g is the product group fixed effect, and X_{fg} includes the lagged value of $\Delta \ln(1 + N_{fg})$ and a rich set of firm-level controls. I control for access to the bond market and a firm's listed status, to account for differences in the ability of a firm to substitute loans. I control for: two-digit NACS fixed effects, whether the firm is in manufacturing, log sales, and log number of products, to compare similar firms. I control for lagged value of $\Delta \ln(1 + N_{fg})$ to account

[7]The results are robust if a twelve-quarter window is used.

[8]There are products that have non-zero sales even after they have had zero sales for eight consecutive quarters. I drop these products. These products constitute a small fraction of the data. I also repeat the analysis using UPCs that have positive sales in at least four consecutive quarters, and obtain similar results. Furthermore, the results remain the same if I only use UPCs that have positive sales in every quarter from their introduction to destruction.

[9]1 is added so that if $N_{fg} = 0$, the observation is included in the regression. I perform the regression analysis with different definitions of outcome variable and find similar results, which are reported in Table 1.8.

for differences in trend prior to Lehman's failure. Additionally, I control for a rich set of loan-level characteristics to account for observable differences in pre-Lehman-failure loans, including: loan amount, maturity, whether a firm had multiple lead lenders, whether the last loan was a term loan, loan start year, and number of loan fixed effects.

1.4 Empirical Results

1.4.1 Total Number of Products

Using regression specification 1.1, I show that firms that faced a reduction in credit supply created fewer products. Table 1.5 shows the results. First, note that all first-stage F-statistics are approximately greater than 10, which alleviates concerns about weak instruments. Second, all standard errors are clustered at the firm and product group levels[10]. For ease of interpretation, ΔL_f is normalized to have a standard deviation of one. This table shows that one-standard-deviation decrease in credit supply reduces the number of products offered by 10% using the Lehman instrument, 8% using ABX, 10% using the bank items instrument, and 10% using all instruments together. All of the estimated coefficients are negative, statistically and economically significant, and quantitatively similar. It is reassuring that the estimated coefficients remain similar despite using the three different instrumental variable strategies, even though the average pairwise correlation of instruments is 0.5. How large is the estimated coefficient relative to employment? Chodorow-Reich (2013) documents that for a one-standard-deviation decrease in credit supply, firms reduce employment by 2.4% percent. Thus the product market response is about four times larger than the labor market response. Interestingly, Broda and Weinstein (2010) document that product turnover is four times labor turnover in the consumer packaged goods market. Thus, the magnitude of the reduced form estimates seems to be consistent with the prior literature.

In Table 1.6, I show that the results are more pronounced for firms that are ex-ante more credit constrained. It is well known that public firms, large firms, and firms with access to the

[10] Results are unchanged if I double-cluster at the firm- and lead-bank-levels.

bond market are less constrained. Table 1.6 shows heterogeneity of the effect for interactions with these measures. The first column shows that the effect of a decrease in credit supply on net product creation is only negative for private firms. The second column shows the effect for firms with access to the bond market. Firms with access to the bond market are less likely to be credit-constrained. The sign of the interaction is consistent with financing constraints, and though it is economically significant, it is not statistically significant. The third and fourth columns of the table show heterogeneity of results based on firms' total sales and total number of products prior to Lehman's failure, respectively. Both of these measures are proxies for firm size, and are normalized to have a standard deviation of one. Both of these columns show that the credit supply effect is stronger for smaller firms.

In order to investigate the dynamics, I run regression specification 1.1 for different time intervals. More specifically, I run specification 1.2 (the same as 1.1, only for different time intervals) and plot the estimated regression coefficients and confidence intervals in Figure 1.1. The coefficient at time t is estimated from the following regression:

$$\ln(1 + N_{fg}^{t}) - \ln(1 + N_{fg}^{t-6}) = \alpha_g + \beta_t \Delta L_f + \gamma X_{fg} + \varepsilon_{fg} \tag{1.2}$$

where $\ln(1 + N_{fg}^{t})$ is the total number of products offered in the four quarters ending with quarter t, and controls are the same as regression 1.1. Figure 1.1 shows that the estimated coefficient is not statistically significant before Lehman's failure, which confirms the identification assumption of regression 1.1, i.e., there are no significant differences among firms that are differentially exposed to the change in credit supply before Lehman's failure. The coefficient becomes negative and significant in 2009Q1 and remains negative and significant for six quarters. The coefficient becomes statistically insignificant again when credit market conditions improve, which is expected if change in net product creation is related to disruptions in credit market conditions.

1.4.2 Product Creation and Destruction

Firms have two margins for reducing products. They can either destroy existing products, create fewer new products, or both. Using the regression specification 1.3, I investigate whether reduced net product creation comes from product creation or destruction:

$$\ln(1 + N_{fg}^{\text{post,k}}) - \ln(1 + N_{fg}^{\text{pre}}) = \alpha_g + \beta \Delta L_f + \gamma X_{fg} + \varepsilon_{fg} \tag{1.3}$$

where k is either creation or destruction, $N_{fg}^{\text{post,creation}}$ is the number of new products that are not available before Lehman's failure, $N_{fg}^{\text{post,destruction}}$ is the number of products whose production is discontinued and so the products are not available post-Lehman failure, and N_{fg}^{pre} is the total number of products before Lehman's failure. The controls are the same as regression 1.1.

The financing constraint only affects product creation, not product destruction. Table 1.7 shows the results. In a simple production model, firms pay a fixed cost to introduce a new product, and a fixed and variable cost for continuation of existing products each period. Theoretically speaking, financing constraints can impede a firm's ability to pay for each of these costs. However, the results found in Table 1.7 favor a model where financing constraints only limit a firm's ability to pay the fixed costs of introducing a new product. This paper is silent about what these fixed costs are. The fixed costs of producing a new product might include purchase of machinery, marketing expenses, R&D, pre-launch experimentation, etc. The reason that this paper doesn't indicate which of these costs becomes binding is lack of appropriate data. Firms that reduce creation are private, as shown in Table 1.6, and the lack of balance sheet data for these firms limits the ability of this paper to distinguish what the sources of their fixed costs are.

I investigate the dynamics of creation and destruction in Figures 1.2 and 1.3, using the dynamic version of regression specification 1.3. Consistent with the results in Table 1.7, the method by which firms respond to reduction in credit supply is reduced creation of new products. The estimated coefficients for destruction regression are never significant, and are economically minor relative to creation coefficients.

1.5 Robustness Checks

In this section, I show that the results are robust using alternative outcome variable definitions, alternative weighting methods, across different departments, and using the Retailer Scanner dataset, collected at store-level, provided by the Nielsen company.

1.5.1 Alternative Outcome Variable Definitions

In Table 1.8, I show that the main result is robust to alternative outcome variable definitions. I use the following five definitions of outcome variable, defined in equation 1.4, and repeat regression 1.1. The first definition is the one used throughout the paper and is the log number of products plus one after the crisis divided by the number of products plus one before the crisis. The second outcome variable is the same as the first one, but it doesn't include the log. The third outcome is the difference in the number of products. The fourth outcome is the difference in the number of products normalized by the total number of *firm* products. The reason that it is normalized by the number of firm products rather than the number of products in that group is some firms might be offering products in a product group that they did not offer before. Thus, normalizing by the number of products in that product group will omit those observations, which is not plausible. The last definition is the definition used to deal with zero when studying employment in the labor economics literature. Table 1.8 shows that using these different outcome variable definitions doesn't change the sign and significance of the estimated coefficient.

$$\begin{aligned}
\text{Main:} \quad & \log(1 + N_{fg}^{t}) - \log(1 + N_{fg}^{t-6}) \\
\text{Division:} \quad & \tfrac{1+N_{fg}^{t}}{1+N_{fg}^{t-6}} \\
\text{Difference:} \quad & N_{fg}^{t} - N_{fg}^{t-6} \\
\text{Normalized Difference:} \quad & \tfrac{N_{fg}^{t} - N_{fg}^{t-6}}{N_{f}^{t-6}} \\
\text{Labor Def.:} \quad & \tfrac{N_{fg}^{t} - N_{fg}^{t-6}}{N_{fg}^{t} + N_{fg}^{t-6}}
\end{aligned} \tag{1.4}$$

1.5.2 Alternative Weighting Methods

In Table 1.9, I show that the results are similar using different regressions weights. In addition to equal weights, which is used throughout the paper, I use two other weighting methods: initial number of products and initial product sales. The results are negative and significant, and quantitively similar for these alternative weighting methods, too. Running this regression has two benefits. First, using the number of products as weights gives the population average estimated effect. Second, sales and the number of products across different product groups are very different, and the distribution is skewed. Using these different weighting methods ensures that the skewness of the sample doesn't change the results of regression 1.1.

1.5.3 More Robustness Checks

In the first and second columns of Table 1.10, I repeat the main regression using a coarser definition of product and find similar results. In Table 1.5, I identify the number of products offered by a firm by counting each item with a distinct UPC as one product. However, different UPCs within a product group might be small variations of each other. For example, two canned beans offered by the same company might only differ in size but will have two different UPCs. To address this concern, I use a coarser categorization of products in the

Nielsen dataset. Each product in a product group (for example, laundry supplies) belongs to a sub-group called product module (laundry supplies product group includes product modules like bleach, fabric softener, detergent booster, etc.). Instead of counting each UPC as a distinct product, I count the number of modules and repeat regression specification 1.1, and report the result in the first column of Table 1.10. The result is negative and significant and consistent with the findings in the main table. In the second column of Table 1.10, I perform a stronger version of this test, by assuming every product in a product category is the same. So, the left-hand-side variable is one if the firm offers a product in the post-period, zero if the firm offers both in the post and pre periods, and minus one if the firm offers a product in pre and not in post. The results in the second column are also consistent with the main findings of the paper.

In the third column of Table 1.10, I repeat the main regression using a dataset that includes store-level data. The main analysis in this paper is done using the Nielsen HMS dataset. This dataset is collected by households using Nielsen provided scanners. The mean and median number of times a UPC appears in HMS over a year is 116 and 5, respectively. Since the purchase of some products happens a few times a year, it could be that there is measurement error in the left-hand-side variable. In particular, there could be products that appear in the markets but are not purchased by consumers and thus not available in the dataset. However, for this measurement error to bias the results, it should be that firms exposed to differential levels of lending face differences in purchasing behavior of at least one product differently. This is unlikely to be the case since the condition that each product being offered is weak; it needs to have at least one sales in the next eight-month. However, I address this concern using Nielsen Retailer Scanner (RMS) dataset, also provided by the Nielsen company. This dataset reports weekly purchases of *every* UPC made at about 35,000 stores across the US. The mean and median number of times a UPC appears in RMS in a year is 98,000 and 980 times. Thus it is much less likely that measurement error is an issue here. The sample of firms that are covered in RMS is smaller, but there are more product categories included in RMS. I repeat regression 1.1 with this dataset, as well. Third column of Table 1.10 shows the results. The result is negative and significant and similar in magnitude to the main table, which is reassuring.

In Table 1.11, I show that the result is robust across different departments within the Nielsen data. Each product in Nielsen dataset belongs to a department. There are several distinct departments in the data, which include: health and beauty aids, food, non-grocery food, alcohol, and general merchandise. Table 1.11 decomposes the effect into different departments. If a reduction in credit supply reduces the number of products, it should be true for all departments, which is borne out in the table, since the effect is either zero or negative and significant across all departments.

1.6 Conclusion

I combine micro-data on millions of different products offered by firms, and link them with their bank-borrowing relationship. I use US banks' exposure to the mortgage market during the 2007-2008 financial crisis as an instrument for credit supply, and show that firms that faced a one-standard-deviation decrease in credit supply created 10% fewer products over the following year. Furthermore, I find that the reduction in products derives from the limited creation of new products rather than the destruction of existing ones. This finding highlights a novel channel through which financial markets affect the real economy, and in particular, consumers.

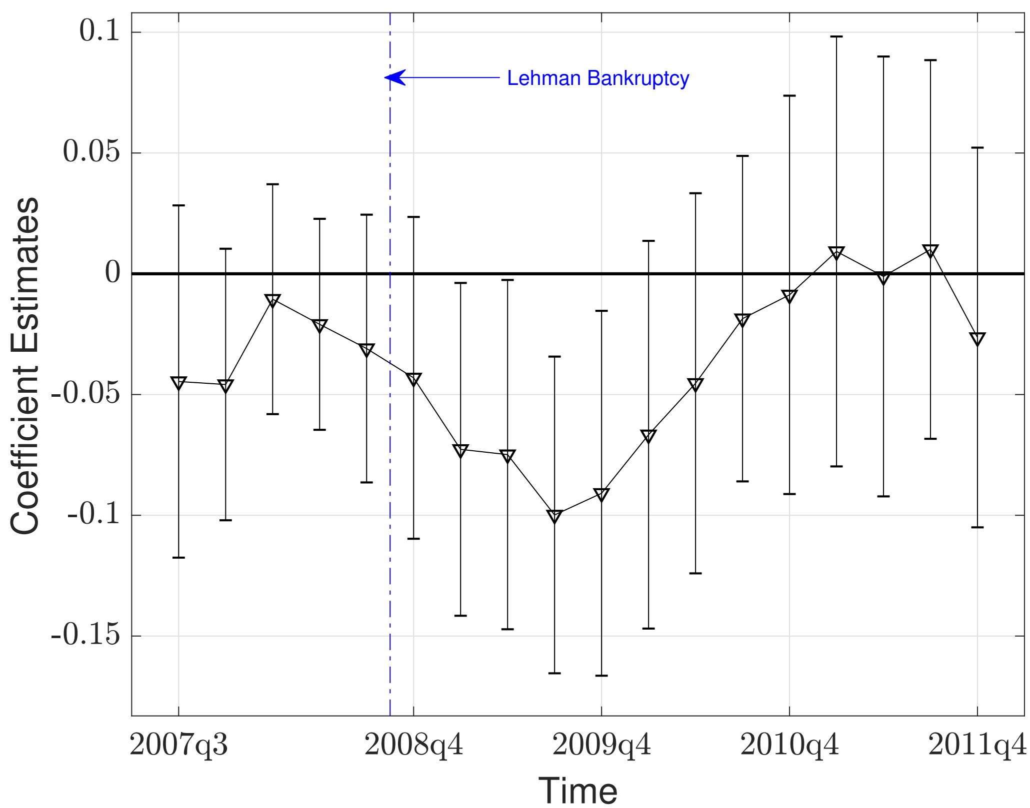

Figure 1.1: Dynamics Estimates: Total Number of Products

This figure is based on regression 1.2, $\ln(1+N^{t}_{fg}) - \ln(1+N^{t-6}_{fg}) = \alpha_g + \beta_t \Delta L_f + \gamma X_{fg} + \varepsilon_{fg}$. The estimate of β_t with the 95% confidence interval is plotted for each quarter. Standard errors are double-clustered by firm and product group code. Firm-level controls include: loan amount; maturity; log sales; log number of products; lagged value of $\Delta \ln(1+N_{fg})$; dummies for manufacturing, multiple lead lenders, term loan, bond issuance, public; and fixed effects for loan start year, pre-Lehman loan number, and 2-digit NAICS.

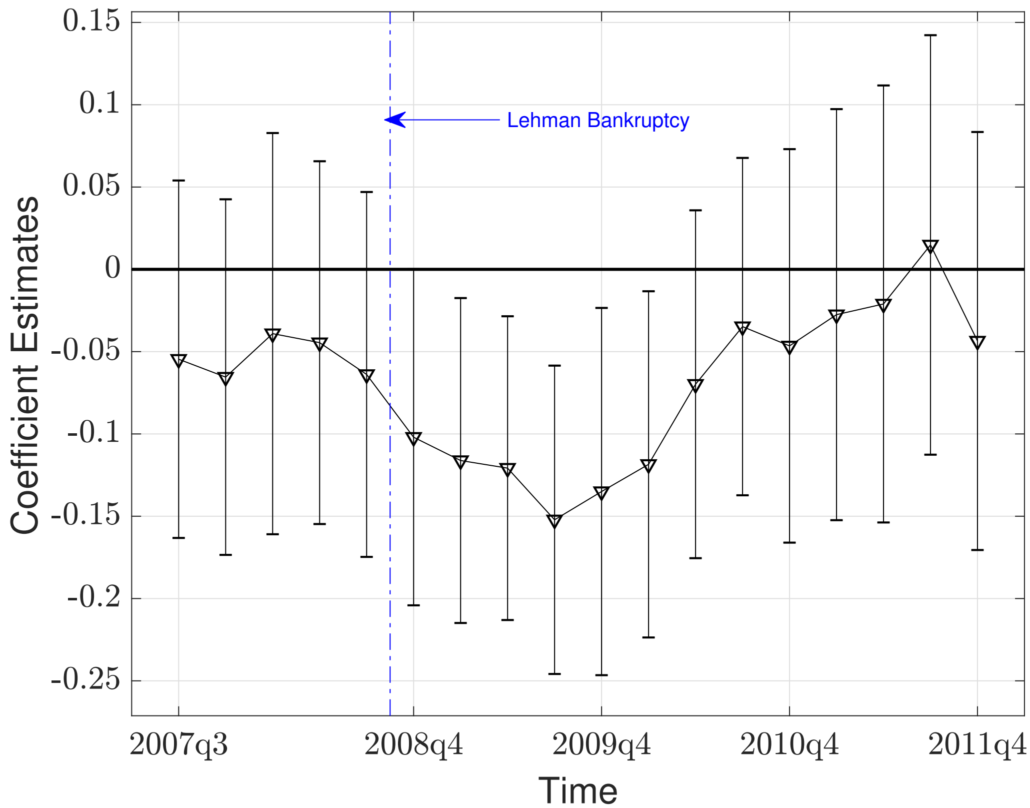

Figure 1.2: Dynamic Estimates: Product Creation Regression

This figure is based on a regression similar to 1.2, except the left-hand side represents the number of new products created, $\ln(1+N_{fg,t}^{\text{post, creation}})-\ln(1+N_{fg,t-6}^{\text{pre}}) = \alpha_g+\beta\Delta L_f+\gamma X_{fg}+\varepsilon_{fg}$. The estimate of β_t with the 95% confidence interval is plotted for each quarter. Standard errors are double-clustered by firm and product group code. Firm-level controls include: loan amount; maturity; log sales; log number of products; lagged value of $\Delta\ln(1+N_{fg})$; dummies for manufacturing, multiple lead lenders, term loan, bond issuance, public; and fixed effects for loan start year, pre-Lehman loan number, and 2-digit NAICS.

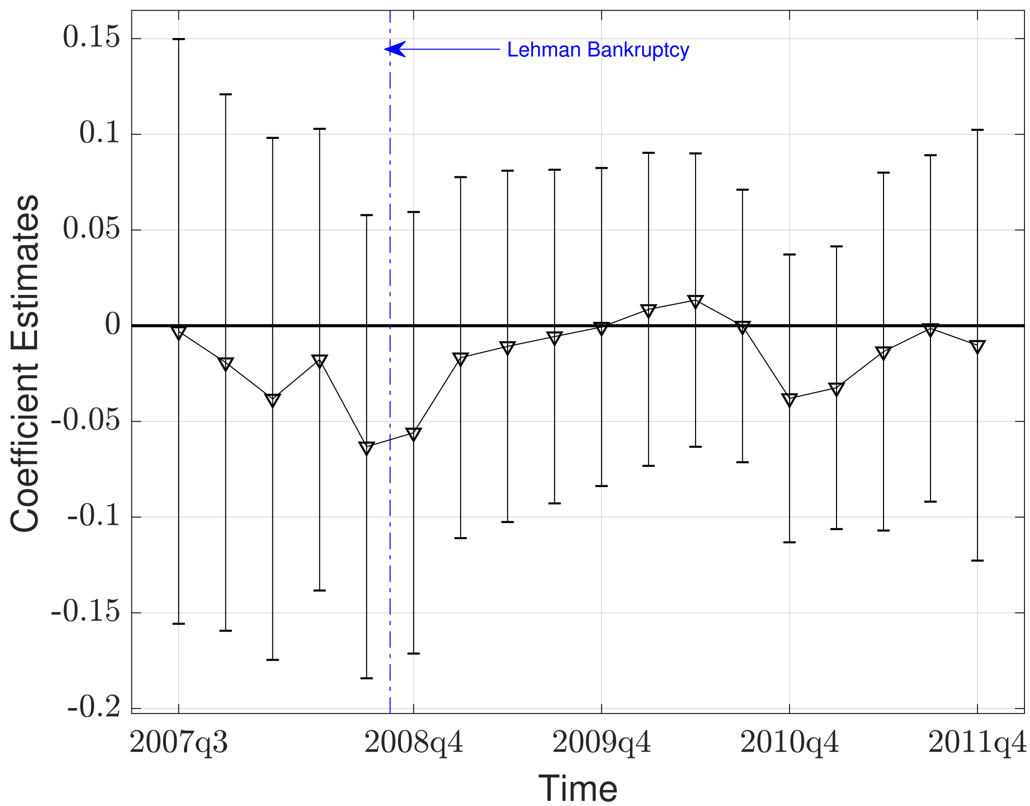

Figure 1.3: Dynamic Estimates: Product Destruction Regression

This figure is based on an regression similar to 1.2, except the left-hand side represents the number of existing products destroyed, $\ln(1 + N_{fg,t}^{\text{post, destruction}}) - \ln(1 + N_{fg,t-6}^{\text{pre}}) = \alpha_g + \beta\Delta L_f + \gamma X_{fg} + \varepsilon_{fg}$. The estimate of β_t with the 95% confidence interval is plotted for each quarter. Standard errors are double-clustered by firm and product group code. Firm-level controls include: loan amount; maturity; log sales; log number of products; lagged value of $\Delta \ln(1 + N_{fg})$; dummies for manufacturing, multiple lead lenders, term loan, bond issuance, public; and fixed effects for loan start year, pre-Lehman loan number, and 2-digit NAICS.

Variable	N	Mean	Median	SD
Product Summary Statistics				
Number of UPCs (N_{fg})	1152	41.8	4	138
Average Price of a UPC	1152	3.3	1.53	9.8
Annual Firm-Group Sales (Million $)	1152	96.5	1.13	385.8
Loan Summary Statistics				
Loan Amount (Million $)	175	354	100	954
Term Loan Dummy	175	0.35	0	.48
Maturity (Months)	175	53.1	60	19.6

Table 1.1: Summary Statistics

Note: The sample includes firms that sold products in the consumer packaged goods market and obtained a loan in the syndicated loans market classified for corporate purpose or working capital purpose. Product summary statistics are based on one year period before from 2007Q2 to 2008Q1. Loan level variables are based on Oct. 2005 to June 2006, and Oct. 2006 to June 2007.

Firm Characteristics

Variable	Manufacturing	Bond Access	Public
coefficient	-0.013	0.122	0.003
p-value	0.87	0.11	0.96

Loan Characteristics

Variable	Loan Amount	Multiple Lead	Maturity	Term Loan
coefficient	0.076	0.132	-.040	0.052
p-value	0.32	0.08	0.60	0.49

Product Characteristics

Variable	Number of Products	Sale	Average Price
coefficient	-0.042	0.024	0.058
p-value	0.58	0.75	0.45

Consumer Characteristics

Variable	Household Income	Household Size
coefficient	0.021	0.023
p-value	0.78	0.76

Table 1.2: Balance of Covariates Table - Lehman Instrument

Note: This table reports p-values of regressions where the left-hand side is Lehman exposure, and the right-hand side is the variable of interest. All variables are normalized to have a standard deviation of one.

Firm Characteristics				
Variable	Manufacturing	Bond Access	Public	
coefficient	0.067	0.014	-.023	
p-value	0.38	0.86	0.76	
Loan Characteristics				
Variable	Loan Amount	Multiple Lead	Maturity	Term Loan
coefficient	.024	-.038	0.079	0.180
p-value	0.75	0.62	0.30	0.02
Product Characteristics				
Variable	Number of Products	Sale	Average Price	
coefficient	-0.031	0.025	-0.026	
p-value	0.68	0.75	0.73	
Consumer Characteristics				
Variable	Household Income	Household Size		
coefficient	-0.002	0.116		
p-value	0.82	0.12		

Table 1.3: Balance of Covariates Table - ABX Instrument

Note: This table reports p-values of regressions where the left-hand side is the ABX instrument, and the right-hand side is the variable of interest. All variables are normalized to have a standard deviation of one.

Firm Characteristics				
Variable	Manufacturing	Bond Access	Public	
coefficient	-.099	-.047	.093	
p-value	0.19	0.53	0.22	
Loan Characteristics				
Variable	Loan Amount	Multiple Lead	Maturity	Term Loan
coefficient	-.079	-.121	-.054	-.134
p-value	0.30	0.11	0.48	0.08
Product Characteristics				
Variable	Number of Products	Sale	Average Price	
coefficient	.001	-.033	.022	
p-value	0.99	0.66	0.77	
Consumer Characteristics				
Variable	Household Income	Household Size		
coefficient	.005	-.065		
p-value	0.55	0.39		

Table 1.4: Balance of Covariates Table - Bank Items Instrument

Note: This table reports p-values of regressions where the left-hand side is the bank items instrument, and the right-hand side is the variable of interest. All variables are normalized to have a standard deviation of one.

	OLS	Lehman	ABX	Bank items	All
ΔL_f	-0.05**	-0.10***	-0.08*	-0.10**	-0.10***
	(0.02)	(0.04)	(0.04)	(0.04)	(0.03)
Observations	1,152	1,152	1,152	1,152	1,152
Firm Controls	Yes	Yes	Yes	Yes	Yes
Product Group FE	Yes	Yes	Yes	Yes	Yes
First Stage F-statistics		12.28	55.18	9.22	9.86

Table 1.5: Main Result

Note: $^*p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$. This table is based on regression specification 1.1. Standard errors are double-clustered by firm and product group code. Firm-level controls include: loan amount; maturity; log sales; log number of products; lagged value of $\Delta \ln(1 + N_{fg})$; dummies for manufacturing, multiple lead lenders, term loan, bond issuance, public; and fixed effects for loan start year, pre-Lehman loan number, and 2-digit NAICS.

	(1)	(2)	(3)	(4)
ΔL_f	-0.15***	-0.13***	-0.11***	-0.10***
	(0.04)	(0.04)	(0.04)	(0.04)
$\Delta L_f \times$ Public Dummy	0.15**			
	(0.06)			
$\Delta L_f \times$ Bond Access Dummy		0.08		
		(0.08)		
$\Delta L_f \times$ Firm Sales			0.13**	
			(0.07)	
$\Delta L_f \times$ Firm Product Number				0.17**
				(0.08)
Sum of Coefficients	0.00	-0.05		
	(0.05)	(0.07)		
Observations	1,152	1,152	1,152	1,152
Firm Controls	Yes	Yes	Yes	Yes
Product Group FE	Yes	Yes	Yes	Yes

Table 1.6: Treatment Interactions

Note: $^{*}p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$. This table is based on equation 1.1 with interaction added to the right-hand side. Standard errors are double-clustered by firm and product group code. Firm-level controls include: loan amount; maturity; log sales; log number of products; lagged value of $\Delta \ln(1 + N_{fg})$; dummies for manufacturing, multiple lead lenders, term loan, bond issuance, public; and fixed effects for loan start year, pre-Lehman loan number, and 2-digit NAICS.

	Product Creation	Product Destruction
ΔL_f	-0.15***	0.00
	(0.05)	(0.04)
Observations	1,152	1,152
Firm Controls	Yes	Yes
Product Group FE	Yes	Yes

Table 1.7: Product Entry and Exit

Note: $^*p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$. This table is based on regression specification 1.3. Standard errors are double-clustered by firm and product group code. Firm-level controls include: loan amount; maturity; log sales; log number of products; lagged value of $\Delta \ln(1 + N_{fg})$; dummies for manufacturing, multiple lead lenders, term loan, bond issuance, public; and fixed effects for loan start year, pre-Lehman loan number, and 2-digit NAICS.

	Main	Division	Difference	Normalized Difference	Labor Def.
ΔL_f	-0.10***	-0.11**	-1.23**	-0.03***	-0.08***
	(0.03)	(0.05)	(0.55)	(0.01)	(0.03)
Observations	1,152	1,152	1,152	1,152	1,152
Firm Controls	Yes	Yes	Yes	Yes	Yes
Product Group FE	Yes	Yes	Yes	Yes	Yes

Table 1.8: Different Outcome Variable

Note: $^*p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$. This table repeats the main regression (regression 1.1), except the left-hand side variable represents alternative definitions defined in 1.4. Standard errors are double-clustered by firm and product group code. Firm-level controls include: loan amount; maturity; log sales; log number of products; lagged value of $\Delta \ln(1+N_{fg})$; dummies for manufacturing, multiple lead lenders, term loan, bond issuance, public; and fixed effects for loan start year, pre-Lehman loan number, and 2-digit NAICS.

	Equal Weighted	Number Weighted	Sales Weighted
ΔL_f	-0.10***	-0.08**	-0.10**
	(0.03)	(0.04)	(0.04)
Observations	1,152	1,152	1,152
Firm Controls	Yes	Yes	Yes
Product Group FE	Yes	Yes	Yes

Table 1.9: Different Weighting Methods

Note: $^*p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$. This table repeats the main regression (regression 1.1), with different weighting methods. The last two columns are weighted by the number of products and sales prior to Lehman's failure, respectively. Standard errors are double-clustered by firm and product group code. Firm-level controls include: loan amount; maturity; log sales; log number of products; lagged value of $\Delta \ln(1 + N_{fg})$; dummies for manufacturing, multiple lead lenders, term loan, bond issuance, public; and fixed effects for loan start year, pre-Lehman loan number, and 2-digit NAICS.

	Number of Modules	Product Group	RMS
ΔL_f	-0.06**	-0.10***	-0.10**
	(0.03)	(0.03)	(0.04)
Observations	1,152	1,152	1,1132
Firm Controls	Yes	Yes	Yes
Product Group FE	Yes	Yes	Yes

Table 1.10: Robustness Checks

Note: $^*p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$. This table repeats the main regression, except the left-hand side variables, are different. The left-hand side variable for the first column is the difference in the number of modules in a product group. The left-hand side variable for the second column is the difference in a dummy variable that takes the value of one if a firm is offering products in a product group, and zero otherwise. The left-hand side variable for the last column is the difference in number of products in the RMS dataset. Standard errors are double-clustered by firm and product group code. Firm-level controls include: loan amount; maturity; log sales; log number of products; lagged value of $\Delta \ln(1 + N_{fg})$; dummies for manufacturing, multiple lead lenders, term loan, bond issuance, public; and fixed effects for loan start year, pre-Lehman loan number, and 2-digit NAICS.

	Health & Beauty Aids	Food	Non-Food Grocery	Alcohol	General Merchandise
ΔL_f	-0.02**	-0.01	-0.05***	0.00	-0.03***
	(0.01)	(0.02)	(0.02)	(0.00)	(0.01)
Observations	1,152	1,152	1,152	1,152	1,152
Firm Controls	Yes	Yes	Yes	Yes	Yes
Product Group FE	Yes	Yes	Yes	Yes	Yes

Table 1.11: Decomposition to Different Departments

Note: $^{*}p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$. This table breaks down the main regression coefficient into different departments in the Nielsen HMS dataset. Standard errors are double-clustered by firm and product group code. Firm-level controls include: loan amount; maturity; log sales; log number of products; lagged value of $\Delta \ln(1+N_{fg})$; dummies for manufacturing, multiple lead lenders, term loan, bond issuance, public; and fixed effects for loan start year, pre-Lehman loan number, and 2-digit NAICS.

Chapter 2

Consumer Welfare and Product Creation: The Credit Supply Channel

Poorya Kabir

I develop a demand and supply model to quantify the equilibrium responses of consumers and firms in response to a reduction in credit supply. I estimate that a one-standard-deviation reduction in credit supply (10% reduced product creation) is responsible for a 1% drop in consumer welfare because of reduced product creation. Two types of equilibrium response are responsible for welfare loss that is smaller than a "naive" interpretation of the reduced form estimates: first, in equilibrium, consumers substitute products with other available products in the same category; and second, in equilibrium, firms' new products have lower "appeal" (quality or taste) relative to existing products.[1]

[1] I am grateful to Olivier Darmouni, Xavier Giroud, and David Weinstein for their invaluable advice during this project. I am also grateful to Charles Calomiris, Tano Santos, Patrick Bolton, Daniel Wolfenzon, Neng Wang, Kairong Xiao, Tania Babina, Gur Huberman, Giorgia Piacentino, Matt Backus, and Kent Daniel for their useful suggestions. Researcher's own analyses are calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

2.1 Introduction

There is a large body of empirical literature that documents disruptions in credit markets have "economically significant" consequences for the real economy. For instance, using reduced-form estimate techniques, Chodorow-Reich (2013) shows that a firm that faces a one-standard-deviation decrease in credit supply reduces employment by 2.4%, relative to a less affected firm. Even though the reduced form estimate is informative, it is only partially informative about the "economic significance" of the effect of credit supply on employment. The "economic significance" depends on the strength of equilibrium responses. For instance, if in equilibrium, *all* laid-off employees (and new would have been employees) of affected firms start working for unaffected firms, then the effect of a credit supply shock on employment will be *zero*, even though reduced-form estimate suggests an "economically significant" effect. In this paper, I show that abstracting away from the equilibrium responses might significantly *overstate* the "economic significance" of a reduced estimate. Thus, it is critical to understand the strength of equilibrium responses when interpreting results from a reduced form estimate.

I illustrate this point by studying the effect of reduction in credit supply on consumer welfare through product creation. I use consumer welfare to measure the "economic significance" of the effect of credit supply on product creation. I do so because, as economists, we are ultimately interested in welfare, and that is how we eventually measure and compare the "economic significance" of different economic forces[2]. To study the effect of reduction in credit supply on consumer welfare through product creation, I break it into two parts:

$$\underbrace{\frac{\partial \text{ Welfare}}{\partial \text{ Credit Supply}}}_{\text{object of interest}} = \underbrace{\frac{\partial \text{ Welfare}}{\partial \text{ Product Creation}}}_{\text{Focus of this Chapter}} \times \underbrace{\frac{\partial \text{ Product Creation}}{\partial \text{ Credit Supply}}}_{\text{Reduced-form estimate (from Chapter 1)}}$$

In Chapter 1 of this thesis, I used variation from US banks' exposure to the mortgage market

[2]Here I abstract away from measuring how much the credit supply shock affects firms due to data limitations. In a static setting, these two questions are separable and the answers are additive. Thus, this paper can be viewed as an answer to one part of the question.

to instrument for credit supply, and estimated $\frac{\partial \text{ Product Creation}}{\partial \text{ Credit Supply}}$. The estimated magnitude is substantial; a one-standard-deviation decrease in credit supply reduces product creation by 10%. Fewer products lower consumer welfare, since consumers have fewer options to choose from. The magnitude of the welfare loss, however, can be small, if $\frac{\partial \text{ Welfare}}{\partial \text{ Product Creation}}$ is small[3]. The goal of this Chapter is to estimate $\frac{\partial \text{ Welfare}}{\partial \text{ Product Creation}}$.

I study the question in the context of the consumer packaged goods market[4], using the Nielsen Home Scan (HMS) dataset. I study this question in the context of the consumer packaged goods market, because, first, it is a large component of household expenses. It accounts for 40% of consumption of goods, and 16% of total household expenditure. Second, product creation benefits both consumers. Consumers benefit from the creation of new products since it either caters to heterogeneous tastes or is a quality upgrade. Third, Nielsen HMS provides micro-level data on the price and quantity of millions of different products.

In order to estimate $\frac{\partial \text{ Welfare}}{\partial \text{ Product Creation}}$, I develop a demand and supply model where the welfare has a closed-form solution in terms of intuitive and economically interpretable parameters. Following Hottman, Redding, and Weinstein (2016a), I model demand using a nested Constant Elasticity of Substitution (CES) framework. The model takes into account two equilibrium responses. First, products within a product group are substitutes. If one product is not offered, consumers can substitute it with another product. The more substitutable the products are, the lower the loss in welfare from reduced product creation. Second, firms choose what products to create. The welfare consequences depend on how much consumers "like" created products. The more a product is liked, the higher the loss in welfare from reduced creation. I show that how much consumers "like" a product can be captured by a parameter that can be uniquely determined from the data. This parameter offers an intuitive interpretation: the marginal "benefit" of the product is relative to its marginal cost.

I model firms' supply decisions by assuming that they choose their prices to maximize

[3]Generally speaking, elements of this decomposition can not be independently and separately computed. But it is an intuitive way of illustrating the main point of this Chapter. Furthermore, it turns out that in this specific example, it can.

[4]The market for goods sold in grocery stores, drug stores, and mass merchandise stores.

profit under the Bertrand competition model. Motivated by the reduced form evidence in the Chapter 1, I assume that firms pay a fixed cost to introduce a new product, and firms' borrowing constraints limit their ability to finance this fixed cost.

I use the model to quantify the consequences of less product creation on consumer welfare and assess the importance of equilibrium responses. In the counterfactual, I assume *every* firm faced a one-standard-deviation *increase* in credit supply and thus created 10% more products. New products increase consumer welfare by 0.97%. The increase in welfare occurs via two channels: first, consumers substitute old products with new products, and this increases welfare; second, new products put competitive pressure on a firm's competitors, and they lower their prices in response, thereby further increasing consumer welfare. I document that the former channel is the dominant channel and accounts for 93% of increase in welfare.

The magnitude of the *increase* in consumer welfare is smaller relative to a "naive" interpretation of the reduced form estimate in Chapter 1, and depends on the equilibrium responses mentioned earlier: first, if products were more substitutable, the estimated effect would be greater. In particular, increasing within-firm elasticity of substitution by 10% – the parameter in the model that captures how substitutable two products within a firm are – reduces the estimated welfare to 0.92%, an approximately 5% relative change in the estimated welfare. Second, if firms offered products that consumers liked better, the welfare effect would be greater. More specifically, if the sales share of new products relative to existing products was 1.00 instead of 0.55 in the actual data, the increase in welfare would be 1.80%.

Related work: This paper contributes to several strands of literature. It contributes to the literature that studies the effects of disruptions in financial markets and reduction in credit supply on real economic outcomes. Previous papers study the effects of disruptions in financial markets on employment (Chodorow-Reich (2013), Giroud and Mueller (2017), Giroud and Mueller (2019), Greenstone, Mas, and Nguyen (2014), Huber (2018)), investment (Almeida, Campello, Laranjeira, and Weisbenner (2012), Amiti and Weinstein (2018)), pricing (Gilchrist, Schoenle, Sim, and Zakrajšek (2017), Kim (2018), Borenstein and Rose (1995), Stroebel and Vavra (2019), Chevalier and Scharfstein (1994), Chevalier and Scharfstein (1995)), likelihood of default (Khwaja and Mian (2008)), etc. I add to this literature by showing that when interpreting results from a reduced-form estimate, it is important to understand the strength of the equilibrium responses. Furthermore, I show how micro-data, in combination with reduced-form estimation, can be used to do this.

It contributes to the literature that studies how financial crises, in particular the Great Recession, affected household consumption (Mian, Rao, and Sufi (2013), Di Maggio, Kermani, Keys, Piskorski, Ramcharan, Seru, and Yao (2017), Mian, Sufi, and Trebbi (2015), Mian and Sufi (2012), Agarwal, Chomsisengphet, Mahoney, and Stroebel (2017), Benmelech, Meisenzahl, and Ramcharan (2017)). While this literature focuses on consumer demand measured by spending, this paper highlights the importance of supply-side decisions made by firms regarding the availability of products. In addition, I quantify the welfare consequences of supply-side decisions.

This paper fits into a large collection of literature in economics that uses barcode-level data to study consumer and firm behavior (for example: Faber and Fally (2017), Jaravel (2018), Hottman, Redding, and Weinstein (2016a), Redding and Weinstein (2016), etc). I also contribute to this literature by showing how price and quantity data can be combined with reduced form estimates to study welfare gains from the creation of new products.

This paper complements the financial economics literature that studies the impact of finance on growth in general (Rajan and Zingales (1998), Jayaratne and Strahan (1996)) and on innovation (Acharya and Xu (2017), Hsu, Tian, and Xu (2014) and Amore, Schneider, and Žaldokas (2013)). I contribute to this literature by demonstrating how one can use detailed product-level data to measure the value of new products for consumers. In addition, this paper contributes to the macroeconomics and international trade literature that documents and quantifies welfare gains in association with variety (Broda and Weinstein (2006a), Goldberg, Khandelwal, Pavcnik, and Topalova (2010), Handbury and Weinstein (2014), Jaravel (2018), etc), by how much credit supply affects consumer welfare through creation of new products.

Finally, this paper adds to the growing financial economics literature that studies frictions in the credit market and their welfare implications by estimating demand and supply models. For example, Crawford, Pavanini, and Schivardi (2018) and Green (2018) study corporate credit, and Cox (2017), Einav, Jenkins, and Levin (2012), Nelson (2017), and Benetton (2018) study different forms of household credit.

2.2 Data

I use the Nielsen HMS dataset to identify prices and quantities of different products offered by firms in consumer packaged goods market[5]. Each year Nielsen keeps track of purchases made by a representative sample of US households numbering between 40,000 and 60,000. These purchases are made at grocery stores, drug stores, mass merchandise stores, and similar retail outlets across the United States. Each household uses a scanner to record the price, quantity, and Universal Product Code (UPC) of items purchased. The UPC is a 12-digit number, and the barcode is a machine-readable version of that number. The first 6 to 10 digits of the UPC function as an identifier assigned to a firm, and are used to merge Nielsen HMS data with the GS1 Data Hub containing the name of the firm associated with the identifier. The data span a twelve-year period from 2004 to 2015, including purchases associated with over 2 million different UPCs. Please refer to Chapter 1 of this thesis for a more detailed explanation of the dataset.

[5]The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

2.3 Model Setup

In Chapter 1, I have shown that firms that face a reduction in credit supply reduce product creation. The magnitude of this reduction is substantial – a one-standard-deviation decrease in credit supply reduces product creation by 10%. Reduced product creation lowers consumer welfare, as consumers have fewer options to choose from. Even though reduced form estimates suggest a reduction in consumer welfare, they are not fully informative about the magnitude of this reduction. In this section, I develop a demand and supply model to quantify the welfare consequences of fewer products.

There are two equilibrium responses that determine the magnitude of welfare loss. First, products in the same product group are substitutes. Consumers can substitute one product with another product offered by the same firm, or other firms in the same product group. This dampens any welfare effect that comes from a product not being created. In the extreme case where a product has a perfect substitute, the loss in welfare is zero. Second, consumers like some products more than others. This is reflected in differences in sales share, conditional on price (for example, regular Coke enjoys higher sales relative to vanilla Coke, even though both have similar prices). The loss in welfare depends on how much consumers like the products that firms cease to create. Welfare consequences are lower when firms stop creating products that consumers don't like as much.

In this section, I develop a demand and supply model to quantify the effect of lower product creation on consumer welfare by taking into account these two equilibrium responses. Following Hottman, Redding, and Weinstein (2016a), I model consumer demand as nested CES. The demand framework includes the two equilibrium responses mentioned earlier. Firm supply is modeled assuming firms maximize profit under the Bertrand competition model. On the supply side, I add a constraint to the firm optimization problem in order to incorporate the finding in the first part of the paper, i.e., financial constraints hinder a firm's ability to create new products. I assume that firms pay a fixed cost to introduce a new product. The ability of firms to pay this fixed cost depends on the degree of borrowing constraint. The model implies that the more financially constrained the firm is, the less able it is to pay this fixed cost. In subsections 2.3.1 and 2.3.2, I describe firm demand and supply in detail.

2.3.1 Demand

Consumer demand is modeled as nested CES with three levels of nesting. The lowest level aggregates utility of consumption of different UPCs within a product group (g) offered by a firm (f) (i.e., regular Coke, diet Coke, Coke zero, cherry Coke, and other UPCs in the carbonated beverage group offered by Coca Cola company[6]). The utility of consumption of UPCs offered by firm (f) in product group (g) is:

$$C_{fgt} = \left[\sum_{u \in \Omega_{fgt}} (\varphi_{ut} C_{ut})^{\frac{\sigma_g^U - 1}{\sigma_g^U}} \right]^{\frac{\sigma_g^U}{\sigma_g^U - 1}} \qquad \prod_{u \in \Omega_{fgt}} \varphi_{ut} = 1$$

where C_{ut}^U is the consumption of UPC u, Ω_{fgt} is the set of UPCs offered by firm f in product group g at time t, and σ_g^U is the elasticity of substitution of UPCs within firm f in product group g. Relative to symmetric CES (CES without φ_{ut} parameters), the term φ_{ut} appears in the utility function. I adopt the terminology used in Hottman, Redding, and Weinstein (2016a) and refer to φ_{ut} as UPC appeal. In order to illustrate the role that φ_{ut} plays in the model, consider the following example: regular Coke and vanilla Coke are both produced by Coca Cola company. Since they have similar prices, symmetric CES will predict that they have similar sales shares. However, regular Coke experiences higher sales relative to vanilla Coke. φ_{ut} is a way of adjusting for unobserved product appeal. A higher value of φ_{ut} corresponds to greater utility of consumption. Thus, through the lens of the model, regular Coke sees higher sales because it has greater appeal. One can interpret appeal as either the quality of a product or the degree to which it aligns with consumer tastes. φ_{ut}

[6]Canned Coke and bottled Coke have different UPCs, so they are considered two different products.

can be alternatively interpreted as the residual difference between symmetric-CES-predicted values of consumption of a UPC u (C^U_{ut}), and its observed value in the data. Symmetric CES predicts that there will be similar consumption of UPCs that have similar prices. Thus, since regular Coke experiences higher sales, it has greater residual difference and higher associated appeal parameters. This latter way of interpreting φ_{ut} becomes more clear in the appendix, where I describe how product appeal can be determined by price and quantity data.

Adding product appeal to the CES framework offers a big advantage. For each UPC at each time, a unique value of φ_{ut} can be determined by using only quantity and price data. In other words, the demand function is invertible in φ_{ut}. What makes φ_{ut} useful is that one can think of it as constant over time, i.e., $\varphi_{ut} = \varphi_{us}$ for all s, t. Evidence in support of this interpretation can be found in a regression run by Hottman, Redding, and Weinstein (2016a). The authors compute φ_{ut} for each UPC over a period of 32 quarters. They run a regression where φ_{ut} is on the left-hand side, and UPC fixed effects are on the right-hand side. The R^2 of the regression is 0.84, which is very high and consistent with the interpretation that φ_{ut} is constant over time. Thus, I assume that these parameters are also constant over time.

The middle nest aggregates utility of consumption of products from product group g across all firms using another CES:

$$C_{gt} = \left[\sum_{f \in \Omega_{gt}} (\varphi_{fgt} C_{fgt})^{\frac{\sigma_g^F - 1}{\sigma_g^F}} \right]^{\frac{\sigma_g^F}{\sigma_g^F - 1}} \qquad \prod \varphi_{fgt} = 1$$

Ω_{gt} is the set of firms in product group g at time t, and σ_g^F is the elasticity of substitution for product group g among different firms. φ_{fgt} has a similar role to φ_{ut} and adjusts for the appeal of a firm within a product group. I will refer to φ_{fgt} as firm appeal. Similarly, Hottman, Redding, and Weinstein (2016a) run a regression where φ_{fgt} is on the left-hand side, and firm product group fixed effects are on the right-hand side. The R^2 of the regression is 0.84, which is very high and consistent with the interpretation that φ_{fgt} is constant over time. Thus, I assume that appeal is also constant over time.

The final nest aggregates utility of consumption at product group (C_{gt}), and uses the Cobb-Douglas functional form:

$$U_t = \sum_{g \in \Omega_t} \psi_{gt} \ln C_{gt} \qquad \qquad \sum \psi_{gt} = 1$$

where Ω_t is the set of all product groups, and ψ_{gt} is sales share of product group g.

To summarize, the nested CES structure has three levels of nests. The lowest level aggregates utility of UPC consumption at the level of product group. The middle level aggregates utility of firm consumption at the level of product group, and the highest level aggregates utility of consumption across product groups.

2.3.2 Supply

The firm optimization problem is a standard one. Firms have two decisions to make: first, firm f chooses the set of UPCs u to offer within product group g at time t $u \in \Omega_{fgt}$; second, firms choose prices P^U_{kt} to maximize profits under the Bertrand competition model. A firm's introduction of a new product is limited by the following assumptions. First, motivated by the reduced form results, I assume that firms pay a fixed cost to create a new product and their ability to pay this fixed cost depends on their financing. Second, I assume that the firm's profit from creating a new product is large enough to cover the fixed cost. Thus, the firm's optimization problem is:

$$\max_{P_{kt}, \Omega_{fgt}} \Pi_{fgt} = \textstyle\sum_{u \in \Omega_{fgt}} P_{kt} Y_{kt} - A_{kt}(Y_{kt}) \tag{2.1}$$

$$\textstyle\sum_{\text{New Product}} O_{kt} \leq \chi(\Delta L_f)$$

$$\forall k \quad O_{kt} \leq P_{kt} Y_{kt} - A_{kt}(Y_{kt})$$

where Y_{kt} denotes the quantity of UPC k, P_{kt} price, and $A_{kt}(Y_{kt})$ the cost of producing Y_{kt} units. O_{kt}[7] is the fixed cost of introducing a new UPC, and the sum of the fixed costs for new UPCs should be less than or equal to $\chi(\Delta L_f)$, where higher ΔL_f corresponds to lower available credit, and χ is a decreasing function of ΔL_f. The first inequality constraint is consistent with the findings in the first part of the paper, i.e., financially constrained firms produce fewer new UPCs. The second inequality constraint insures that creating a new UPC is profitable for the firm. In the next subsection, I demonstrate how the model can be used for calculating welfare.

2.4 Model's Solution and the Welfare Index

The model's solution is similar to Hottman, Redding, and Weinstein (2016a). For derivations, please refer to the paper. Here, I only show the relevant equations. First, I summarize the results, and show how the available data can be used to back out unique values for φ_{ut} and φ_{fgt}. I denote the fraction of sales of UPC u in product group g by firm f at time t by S_{ut} (for example original Coke's sale as a fraction of sales of Coca Cola company in the carbonated beverages product group). S_{ut} is equal to:

$$S_{ut} = \frac{(P_{ut}/\varphi_{ut})^{1-\sigma_g^U}}{\sum_{k\in\Omega_{fgt}}(P_{kt}/\varphi_{kt})^{1-\sigma_g^U}} \qquad \prod_{u\in\Omega_{fgt}} \varphi_{ut} = 1$$

The number of equations and unknowns is the same, and there are unique values of φ_{ut} that satisfies all equations simultaneously. I denote the fraction of sales of firm f in product group g at time t by S_{fgt} (for example Coca Cola company's sale as a fraction of sales in carbonated beverages). S_{fgt} is equal to:

$$S_{fgt} = \frac{(P_{fgt}/\varphi_{fgt})^{1-\sigma_g^F}}{\sum_{k\in\Omega_{gt}}(P_{kgt}/\varphi_{kgt})^{1-\sigma_g^F}} \qquad \prod_{u\in\Omega_{gt}} \varphi_{fgt} = 1$$

[7] The inequality constraint included here is the only difference between my firm optimization problem and that of Hottman, Redding, and Weinstein (2016a).

where P_{fgt} is firm expenditure function, and is computed as follows:

$$P_{fgt} = \Big(\sum_{u \in \Omega_{fgt}} (P_{ut}/\varphi_{ut})^{1-\sigma_g^U} \Big)^{\frac{1}{1-\sigma_g^U}}$$

The number of equations and unknowns is the same, and there are unique values of φ_{fgt} that satisfies all equations simultaneously. Furthermore, markup is the same across all UPCs offered by a firm a product group and is equal to:

$$\mu_{fgt} = \frac{\sigma^F(1 - S_{fgt}) + S_{fgt}}{\sigma^F(1 - S_{fgt}) + S_{fgt} - 1}$$

and thus marginal cost of UPC u is equal to:

$$\gamma_{ut} = \frac{P_{ut}}{\mu_{fgt}}$$

Estimation of elasticities of substitution within firm's products (σ_g^U) and across firms within a product group (σ_g^F) follows Hottman, Redding, and Weinstein (2016a).

The firm-level expenditure function, which is the minimum amount of spending required to gain one unit of utility through consuming products offered by firm f within product group g, is equal to:

$$P_{fgt}^F = \Big(\sum_{u \in \Omega_{fgt}} ({\varphi_{ut}}^{-1} P_{ut})^{1-\sigma_g^U} \Big)^{\frac{1}{1-\sigma_g^U}} \tag{2.2}$$

and, the group-level expenditure function, or the minimum amount of spending required within product group g to gain one unit of utility, is:

$$P_{gt} = \Big(\sum_{f \in \Omega_{gt}} (\varphi_{fgt}^{-1} P_{fgt})^{1-\sigma_g^F} \Big)^{\frac{1}{1-\sigma_g^F}} \tag{2.3}$$

and, the degree of consumer utility from spending \$1 on any available products:

$$W_t = (\sum_g \psi_{gt} P_{gt})^{-1} \tag{2.4}$$

where ψ_{gt} is the share of expenditure on product group g. Since consumer preferences is homothetic, the ratio of two different UPCs in the optimal consumption choice is the same for all levels of income. Thus, in order to understand welfare, I need to understand the utility of consumption provided by \$1. Thus, W_t measures welfare, and I refer to W_t as the welfare index. W_t possesses intuitive properties. W_t is a decreasing function of UPC prices (P_{ut}), while it is an increasing function of both UPC (φ_{ut}) and firm appeal (φ_{fgt}). W_t also depends on how substitutable products are within a firm, as captured by σ_g^U, and how substitutable products are across firms, as captured by σ_g^F. Finally, the greater the share of expenditure on a product group, the more important it is in the welfare index.

2.5 Estimation and Results

2.5.1 Parameter Estimation

Nielsen data include the prices (P_{ut}) and quantities of all UPCs (C_{ut}). Additionally, sales share (ψ_{gt}) can be computed from the data. Thus, in order to compute W_t, the parameters that are absent from the data are φ_{ut}, φ_{fgt}, σ_g^U and σ_g^F. As shown in the online appendix, if elasticities are known (σ_g^U and σ_g^F), φ_{ut} and φ_{fgt} can be uniquely determined by using price and quantity data available in the Nielsen dataset.

In order to estimate the elasticities of substitution across products within a firm (σ_g^U) and across firms within a product group (σ_g^F), I follow Hottman, Redding, and Weinstein (2016a)[8]. σ_g^U is estimated using GMM and variation in double-differenced price and log product expenditure shares of UPCs over time and relative to the largest product of the firm.

[8]Please refer to the paper for a detailed discussion of how elasticities are estimated. I thank Hottman, Redding, and Weinstein (2016a) for sharing their elasticity estimates with me.

σ_g^F is estimated using an instrumental variable strategy and variation in double-differenced price and log product expenditure shares of firms over time and relative to the largest firm associated with the product group.

Figures 2.1 and 2.2 shows histogram of the estimated elasticities of substitution. The median σ^U (within-firm elasticity of substitution) is 6.9, and the median σ^F (across-firm elasticity of substitution) is 3.9. Not only median across firm elasticity of substitution is smaller compared to the median with firm elasticity of substitution, but also, this pattern holds for every single product group. This finding suggests that products within firms are more substitutable than products across firms, which seems intuitive.

The goal of the rest of this subsection is to understand how the framework developed by Hottman, Redding, and Weinstein (2016a) can be adapted to measure the welfare derived from creation of new products. First, I show what parameters are needed to solve the model. Second, I show that a subset of those parameters are needed to calculate welfare. Third, I discuss how model parameters can be estimated from the data.

The next two propositions demonstrate sufficient statistics for computing welfare.

Proposition 1: For any values of φ_{fgt} (firm appeal f within product group g), φ_{ut} (appeal of UPC u) and γ_{ut} (marginal cost of UPC u) there exists a unique equilibrium, i.e., there exists a unique price and quantity for every UPC that clears the market.

Proof: for proof refer to appendix.

Proposition 2: In order to compute the welfare index (W_t), using $\frac{\varphi_{ut}\varphi_{fgt}}{\gamma_{ut}}$ for all UPCs is sufficient. In other words, separate knowledge of φ_{ut}, φ_{fgt} and γ_{ut} isn't necessary to calculate the welfare index.

Proof: for proof refer to appendix.

The above propositions establish the existence and uniqueness of equilibrium, and that $\frac{\varphi_{ut}\varphi_{fgt}}{\gamma_{ut}}$ is key for measuring the welfare index. Define Appeal divided by Marginal Cost

(AMC) of UPC u by:

$$\text{AMC of UPC u} = \frac{\varphi_{ut}\varphi_{fgt}}{\gamma_{ut}}$$

Thus previous propositions establish that having AMCs for all UPCs is sufficient for measuring welfare. In order to understand how a change in number of products affects consumer welfare, I need to be able to model a firm's choice in new products. This includes: first, the number of new products; second, the AMC of new products. Using reduced form results from the first part of the paper, I am able to compute the number of new products. Since the AMC of existing products is known, I need only know the AMC of new products. I assume the AMC of new products can be estimated with the following model:

Assumption: The AMC of a new UPC u produced by firm f from product group g, follows the following model:

$$\log(\text{AMC}_{ut}) = \alpha_g + \theta E(\log(\text{AMC}_{kt})|\text{exist}) + \varepsilon_{ut} \tag{2.5}$$

where $\log(\text{AMC}_{ut})$ is the log of the AMC of new UPC u at time t, $E(\log(\text{AMC}_{kt})|\text{exist})$ denotes the average log of AMCs of existing UPCs of firm f in product group g, α_g is the product group fixed effect, and ε_{ut} is the error term.

I use data from the period 2004 to 2006 to estimate the coefficients of regression 2.5. I split the data into two equal intervals, and define a product as "existing" if the product is in the first half of the sample, and "new" if it is not in the first half, but exists in the second half. In the next subsection, I provide two pieces of evidence that this is a plausible way of thinking about the AMC of a new product.

2.5.2 Evidence on Model's Plausibility

I show that the assumption in equation 2.5 is a plausible way of estimating AMCs for new products by: 1) running regression 2.5 in the data, 2) doing a counterfactual exercise. I use data from 2004 to 2006 and run regression 2.5. I split the data into two equal intervals, and define a product as existing if the product is in the first half of the sample, and new if it is not in the first half of the sample, but it is in the second half. Table 2.1 shows some statistics of regression 2.5. The R^2 is 0.97 which is very high [9] and suggests that the average AMC of existing products can be used to simulate AMCs of new products. In addition, $\theta < 1$, and 87% of estimated α_gs are negative, which is intuitive and consistent with the fact that new products have lower sales share relative to existing products. Thus, Table 2.1 provides in-sample evidence consistent with the assumption in equation 2.5.

Even though Table 2.1 provides in-sample evidence consistent with the assumption in the equation 2.5, it is not clear whether the model is able to generate out-of-sample counterfactuals that approximate the data well enough. I use data from 2004-2006, as described before and estimate the coefficients of equation 2.5, and the distribution of error term. I use these coefficients and distribution of error term to test out-of-sample properties of the equation 2.5. I perform the following exercise: for 2007, I identify a product as new if it is available in 2007, and was not available before. I replace the AMC of new products with the AMC simulated from equation 2.5. I then use the simulated AMCs to solve firm optimization problem, and compute sales share of new products and the welfare index.I do this exercise 20 times, and report the ratio of sales share of new products and welfare index ($W_{\text{ratio}} = \frac{\sum_g \psi_g P_{g,\text{real data}}}{\sum_g \psi_g P_{g,\text{counterfactual}}}$) in Table 2.2.

The mean estimated price ratio is 1.0003 and is very close to one. The standard deviation of estimated W_{ratio} is also tiny. So the model can replicate the welfare index very well. Also, the model does an excellent job of approximating the sales share of new products on average, which is a desirable feature of the model, even though not of direct interest.

[9] Part of the reason that there is such a high R^2 is that both left-hand side and right-hand side include $\log(\varphi_{fgt})$.

2.5.3 Selection of Most Appealing Products

In this subsection, I show that under certain conditions it is optimal for firms to offer products with the highest AMC, and I test the implications of the model for the data.

Proposition 3: Assume that firms pay a fixed cost to introduce a new product, and their ability to pay this fixed cost depends on availability of credit. If different products have different fixed costs, then a reduction in credit supply could decrease or increase the number of new products. However, if the cost of introducing every product is the same, then firms will decrease the number of their products when faced with stricter borrowing constraints.

Proof: for proof refer to appendix.

Proposition 4: Assume that a firm has to choose K UPC out of $N \geq K$ potential UPCs with the same fixed cost[10]. Denote UPC i's appeal and marginal cost by φ_i and γ_i, respectively. Without loss of generality assume that $\frac{\varphi_N}{\gamma_N} \leq \cdots \leq \frac{\varphi_1}{\gamma_1}$. If a firm observes $\frac{\varphi_i}{\gamma_i}$ before choosing which products to produce, then the firm's optimal decision is to produce UPCs 1 to K. In other words, the most profitable UPCs are the ones with the highest AMCs[11].

Proof: for proof refer to appendix.

My previous proposition establishes that firms produce products that have higher AMCs first, and then move to products with lower AMCs. I test this in the data. In order to test this prediction, I compute the residual AMC of a new product using equation 2.5. Since the distribution of the error term is available for each product group, I compute the percentile of the error term in the distribution. Figure 2.3 plots these percentiles. If firms stop introducing lowest-AMC products earlier, the distribution of percentiles should have higher mass on the right and lower mass on the left, which is the case here.

The previous graph represents a weak implication of selecting higher-AMC products. A deeper implication is that firms that reduce the number of new products will choose to

[10]With the level of data that is available it is impossible to identify the fixed cost of producing products. Thus, I will assume that a firm's cost of producing a UPC is the same across all products within a product group. In the appendix, I show why making this assumption is important.

[11]Here, since I am focusing on a single firm, multiplying all firm UPCs by firm product group appeal will not change the ranking of the products.

produce higher-AMC products. Consequently, on average, more constrained firms should have higher AMCs. I exchange the left-hand variable of regression 1.1 with the average percentile of new products. The results are reported in Table 2.3. The table suggests that there is no selection of AMCs[12].

What is the implication of lack of selection in association with appeal? It is useful for distinguishing between two models of firm. In the first model, a firm observes the AMC of its products before paying the fixed cost of introduction. The implication of this model is that the average AMC of products for firms with tighter borrowing constraints should be higher. In the second model, a firm doesn't observe the AMC of products before paying the fixed cost. The evidence in Table 2.3 is consistent with the latter model[13].

2.5.4 Decomposing Welfare Increase from Product Creation

Introducing a new product increases consumer welfare through two channels. First, consumers that purchase the product are better off, since, by revealed preference, they prefer consuming that UPC to other UPCs. I refer to this as a "direct" increase in welfare from a new UPC. Second, when a firm introduces a new UPC, it will draw customers of UPCs produced by other firms. Consequently, other firms lower their prices to compete. Thus, introducing a new UPC lowers the price of other firms' UPCs through competition, and raises consumer welfare as well. I refer to this as an "indirect" increase in welfare[14]. I break down the effect of adding a new UPC to these two components. More formally, consider the

[12]With similar reasoning, it can be shown that the lowest AMC of a constrained firm should be higher than the lowest AMC of a non-constrained firm. I test this using regression framework 1.1 by changing the left-hand side to the percentile of the lowest AMC of new products. In Table 2.4 in the online appendix, I show the estimated coefficients. The coefficients are not significant, either.

[13]Note that in both of these models firms make pricing decisions on the basis of knowing the AMCs of products. Thus, evidence in this subsection supports the assumption that firms don't observe the AMCs of new products before introducing them.

[14]When a firm introduces a new product, it will have more market power and thus will increase its prices. Other firms, as discussed above, will decrease their prices. If the increase in a firm's prices exceeds the decrease in other firms' prices, the price of existing UPCs can go up. However, this doesn't happen in the data most of the time, so I ignore it.

following decomposition:

$$\log\left(\frac{W_{\mathrm{c}}}{W_{\mathrm{e}}}\right) = \log\left(\frac{W_{\mathrm{c}}}{W_{\mathrm{en}}}\right) + \log\left(\frac{W_{\mathrm{en}}}{W_{\mathrm{e}}}\right)$$

W_{c} is the welfare index of the counterfactual, where I add new UPCs, W_{e} is the welfare index before adding new UPCs, and, W_{en} is the welfare index of existing products computed with new prices, i.e., prices after adding new UPCs. Thus $\log\left(\frac{W_{\mathrm{en}}}{W_{\mathrm{e}}}\right)$ captures the indirect welfare effect coming from changes in price of existing UPCs, and $\log\left(\frac{W_{\mathrm{c}}}{W_{\mathrm{en}}}\right)$ captures the direct effect of introducing a new product after accounting for changes in price of existing UPCs.

2.5.5 Counterfactuals

Illustrative example: Before moving to counterfactual scenarios, I illustrate the mechanisms of the model by considering what happens when one firm introduces new UPCs. In this example, I take a firm that earns about 6.7% of its sales from a product group in 2007. The number of UPCs that this firm offers is 1,095, within a product group with 12,287 UPCs[15]. I increase the number of UPCs produced by the firm by 10% to 1,205, and leave other firms' products as is. The AMC of the 110 new products is constructed using 2.5, and all firms choose prices to maximize profit according to 2.1. I draw 20 random sets of new AMCs.

Table 2.5 includes some statistics from this exercise. New products increase the firm's sales by 7.2%. Sales of new UPCs comes from customers substituting existing UPCs with new ones. Some of these existing UPCs belong to the same firm that introduced the new UPCs. Thus, sales of the firm's original product go down. In this example, sales of existing products go down by 2.8%. Thus, overall firm sales increase by 4.4%. This increase in overall firm sales comes at the expense of other firms. In a CES framework, other firms lose the same amount of market share, i.e., other firms lose 4.4% of their market share. New UPCs add to welfare. A consumer who spends all their funds on the firm that offers the new products

[15]To make sure that the firm's information is not identifiable, I don't mention the product group's name.

will experience a 0.89% increase in welfare. However, this increase in welfare lessens when considering there are other firms in the product group. A consumer that spends all their funds on the product group that includes the firm with added products (thus some of the expenditure goes to the firm with new products and the rest to other firms) will experience a 0.07% increase in welfare, which is lower than 0.89% since other firms' products didn't change. I also decompose the welfare index into direct and indirect components. The direct effect dominates and accounts for 82% of the increase in welfare index, and the indirect effect accounts for 18% of the change in welfare index. Next, I describe the counterfactual scenarios.

10% **increase in products created**: In this counterfactual, I assume all firms in the Nielsen HMS sample experience a one-standard-deviation increase in credit supply. Reduced form estimates from Table 1.5 imply that firms would have created 10% more new products[16]. I generate the AMCs of new products using 2.5, and solve for firms' optimal pricing decisions to maximize profits according to equation 2.1. I repeat this exercise 20 times.

A 10% increase in newly created products increases consumer welfare by 0.97%. The first row of Table 2.6 shows that a 10% increase in newly created products will make up 5.71% of sales, and will increase consumer welfare by 0.97%. These results are aggregated across different product groups using sales shares as weights. Furthermore, welfare decomposition shows that 93% of the increase in welfare comes from the direct effect of adding new UPCs to the market, while only 7% of the increase comes from the indirect effect. Both the sales share of new UPCs and the welfare index are aggregated across product-group-level variables. In figure 2.4, I plot the histograms of change in both the product group expenditure function and the new product sales share. The median decrease in the product group expenditure function is 1.10%, and the median new product sales share is 5.78%.

Changing σ_U and σ_F: In this exercise, I investigate the extent to which welfare results

[16] I assume that the only constraint that is binding in the firm optimization problem in 2.1 is the credit supply condition. I assume that it is profitable for firms to create new products. In other words, the second inequality in 2.1 is non-binding. Since, on average, 20% of the products in this market are new each year, it is likely that a 10% increase in products created will still be profitable.

depend on the substitution pattern captured by across-UPC within-firm elasticity of substitution (σ_U) and across-firm within-product-group elasticity of substitution (σ_F). In order to answer this question, I increase and decrease the elasticities across all product groups by 10%, and repeat the previous exercise.

Across-UPC elasticity of substitution σ_U controls how substitutable two products within a firm can be. As σ_U increases, products become more substitutable, and the importance of a new product decreases. Thus, welfare gains from creating a new product decrease as σ_U increases. Rows four and five of Table 2.6 show these results. The increase in the welfare index is 0.92% and 1.05%, for a 10% increase and a 10% decrease in σ^U, respectively. Across-firm within-product-group elasticity controls how substitutable products can be across firms. The higher these elasticities, the more substitutable products become. This changes firms' pricing incentives. In particular, firms decide on lower prices as products across firms become more substitutable. This is reflected in firm markup equation A.2 in the appendix. Markups charged by firms lower consumer welfare; thus an increase in σ_F increases consumer welfare by reducing markups through changes in pricing incentives. This is reflected in rows two and three of Table 2.6. The increase in the welfare index is 1.01% and 0.94%, for a 10% increase and a 10% decrease in σ^F, respectively. The sales share of new UPCs doesn't change. The reason for this is that sales shares of new UPCs depend on the relative AMC of new UPCs, and the relative AMC of new UPCs remains the same. This is further evidence that the model for the AMC of new UPCs captures the AMCs of new UPCs well.

Selection of more appealing products: In this exercise, I document how the change in welfare index depends on how much consumers like new products. I assume that the average AMC for new UPCs is shifted by 18.1% [17]. The results are reported in the last row of Table 2.6. Sales shares of new UPCs increase to 10.01%, and the increase in the welfare index is 1.80%. This result highlights the importance of offering more appealing products.

[17] I chose 18.1% so that the sales shares of a new product relative to an existing product would be 1.00.

2.6 Conclusion

I develop a demand and supply framework to quantify the impact of reduced product creation on consumer welfare. My estimates suggest that 10% less product creation lowers consumer welfare by 0.97%. Three equilibrium responses are responsible for welfare loss that is smaller than a "naive" interpretation of the reduced form estimates. First, consumers can substitute products. The easier it is to substitute a product, the lower the loss in welfare. Second, firms choose what products to offer. Welfare loss depends on how much consumers like a product, as measured by the AMC. Third, welfare loss depends on the markup charged by the firm. The higher the markup, the lower the loss in consumer welfare.

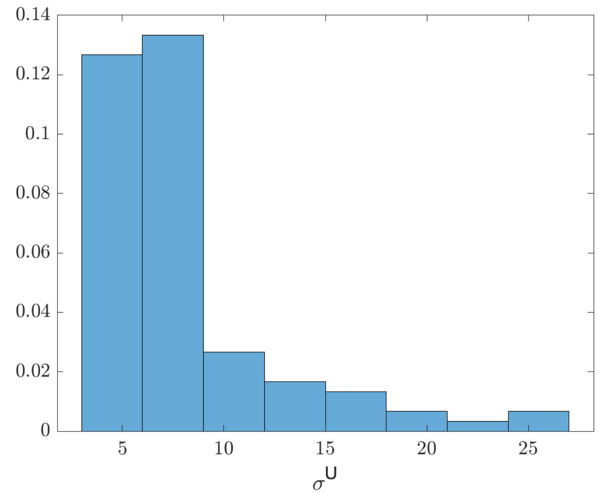

Figure 2.1: Histogram of estimated σ^U (within firm elasticity of substitution)

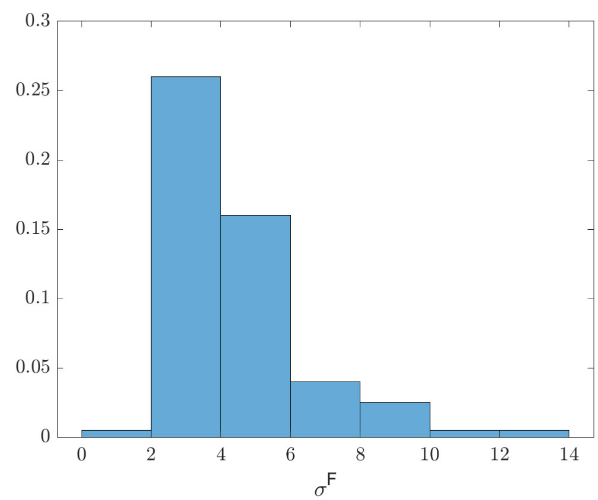

Figure 2.2: Histogram of estimated σ^F (across firm elasticity of substitution)

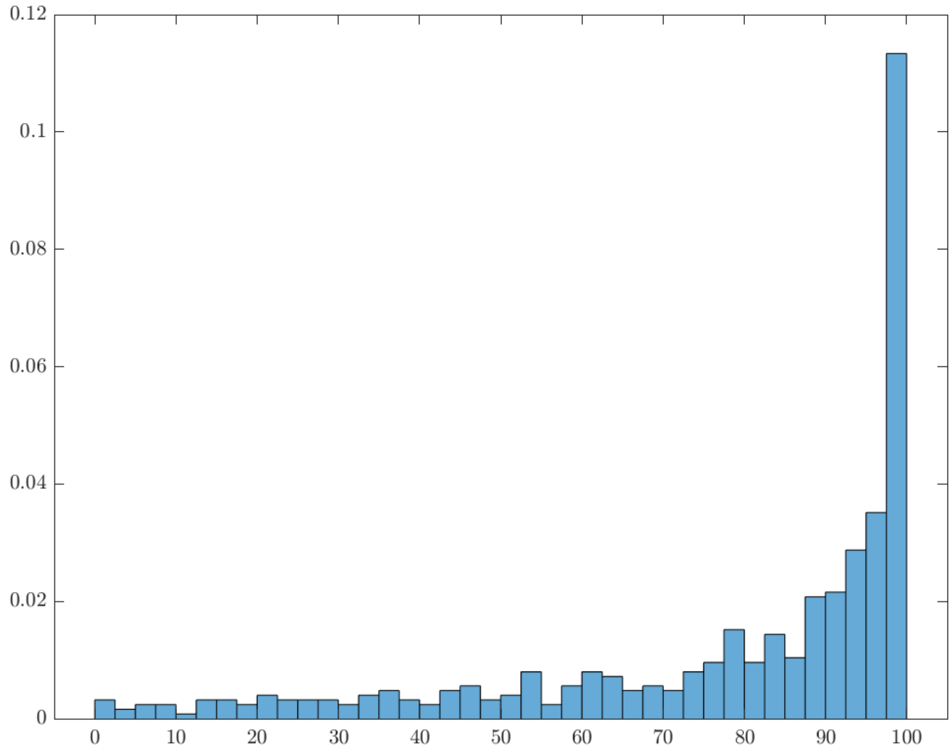

Figure 2.3: Histogram of AMC Percentile for New Products

Note: This figure plots the histogram of the AMC (AMC is defined in equation 2.5, and can be interpreted as marginal "benefit" divided by marginal cost) percentile for the new products created after Lehman's failure. AMC percentiles are calculated based on estimation of equation 2.5 using data from 2004-2006.

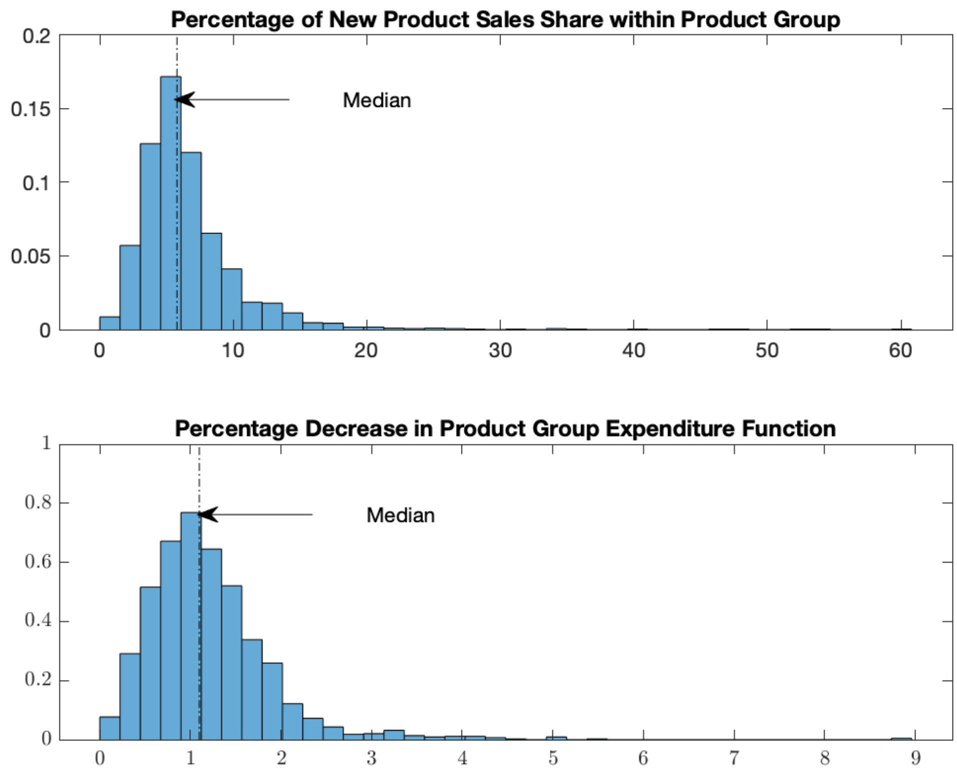

Figure 2.4: Histograms of Percentage of New Product Sales Share and Percentage Decrease in Product Group Expenditure Function

The top figure plots the histogram of the percentage of sales share associated with newly created products within various product groups (for example, carbonated beverages) when *every* firm faces a one-standard-deviation increase in credit supply. The bottom figure plots the percentage decrease in product group expenditure function (expenditure function is computed using equation 2.3) when *every* firm faces a one-standard-deviation increase in credit supply.

	R^2	Adjusted R^2	θ	Mean(α_g)	StD(α_g)	Fraction of α_g Negative
Statistics Value	0.97	0.97	0.92	-0.11	0.12	0.87

Table 2.1: New and Existing Product AMC Relation

This table reports statistics of regression 2.5. Regression 2.5 is estimated using data from 2004-2006.

	Mean	StD	5th Quantile	Median	95th Quantile
W_{ratio}	1.0003	0.0011	0.9983	1.0002	1.0020
New Product Sales Ratio	1.016	0.016	0.985	1.017	1.045

Table 2.2: Ratio of the Welfare Index Ratio (W_{ratio}) and Sales for Counterfactual and Actual Data

W_{ratio} is the welfare index using simulated AMCs (AMC is defined in equation 2.5, and can be interpreted as marginal "benefit" divided by marginal cost) from equation 2.5 for all new products in 2007 divided by the welfare index using actual data. New Product Sales Ratio is the sales share of new products using simulated AMCs from equation 2.5 for all new products in 2007 divided by the sales share of new products in actual data. Equation 2.5 is estimated using data from 2004-2006.

	Lehman	ABX	Bank items	All
ΔL_f	-0.03	0.04	0.03	-0.01
	(0.02)	(0.04)	(0.04)	(0.02)
Observations	485	485	485	485
Firm Controls	Yes	Yes	Yes	Yes
Product Group FE	Yes	Yes	Yes	Yes

Table 2.3: Selection of More Appealing Products

Note: $^{*}p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$. This table is based on a regression similar to 1.1, except the left-hand side variable represents the average percentile of the AMC (AMC is defined in equation 2.5, and can be interpreted as marginal "benefit" divided by marginal cost) of new products after Lehman's failure. Standard errors are double-clustered by firm and product group code. Firm-level controls include: loan amount; maturity; log sales; log number of products; lagged value of $\Delta \ln(1+N_{fg})$; dummies for manufacturing, multiple lead lenders, term loan, bond issuance, public; and fixed effects for loan start year, pre-Lehman loan number, and 2-digit NAICS.

	Lehman	ABX	Bank items	All
ΔL_f	-1.80	5.07	-0.90	-0.28
	(2.55)	(4.39)	(4.47)	(2.50)
Observations	486	486	486	486
Firm Controls	Yes	Yes	Yes	Yes
Product Group FE	Yes	Yes	Yes	Yes

Table 2.4: Selection of More Appealing Products

Note: $^{*}p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$. This table is based on a regression similar to 1.1, except the left-hand side variable, represents the percentile of the lowest AMC (AMC is defined in equation 2.5, and can be interpreted as marginal "benefit" divided by marginal cost) of new products after Lehman's failure. Standard errors are double-clustered by firm and product group code. Firm-level controls include: loan amount; maturity; log sales; log number of products; lagged value of $\Delta \ln(1+N_{fg})$; dummies for manufacturing, multiple lead lenders, term loan, bond issuance, public; and fixed effects for loan start year, pre-Lehman loan number, and 2-digit NAICS.

	Percentage Increase in Firm's Sale	Share of New Sales Coming from Other Firms	Percentage Decrease in Firm's Expenditure Function	Percentage Decrease in Product Group Expenditure Function
Mean	4.4	0.61	0.89	0.07
(StD)	(1.3)	(0.01)	(0.23)	(0.02)

Table 2.5: Summary Statistics of Increasing a Firm's UPCs by 10%

Note: This table reports summary statistics of an illustrative example, where I increase the products of one firm by 10%. AMCs (AMC is defined in equation 2.5, and can be interpreted as marginal "benefit" divided by marginal cost) of new products are simulated using equation 2.5. Firm and product group expenditure function are calculated using equations 2.2 and 2.3, respectively.

		Sales Share of New UPCs	Percentage Increase in the Welfare Index
10% increase in products created		5.71 (0.10)	0.97 (0.16)
Changing σ_U and σ_F	decrease σ_F by 10%	5.72 (0.10)	0.94 (0.18)
	increase σ_F by 10%	5.71 (0.11)	1.01 (0.15)
	decrease σ_U by 10%	5.64 (0.10)	1.05 (0.16)
	increase σ_U by 10%	5.78 (0.10)	0.92 (0.17)
Selection of more appealing products		10.01 (0.14)	1.80 (0.21)

Table 2.6: New Product Sales Share & Increase in the Welfare Index

Note: This table reports the sales shares of new products and the increase in welfare from the creation of new products. I assume *every* firm faces a one-standard-deviation increase in credit supply. In rows 2-5, I change the elasticities for every product group. In the last row, I increase the mean AMC (AMC is defined in equation 2.5, and can be interpreted as marginal "benefit" divided by marginal cost) of new products by 18.1% . This increase in average AMC increases the sales share of a new product relative to an existing product from 0.55 (in the actual data) to 1.00.

Chapter 3

Unpacking the Effect of Financial Constraints on Investment

Poorya Kabir & Seyed Mohammad Mansouri

This paper studies how financially constrained firms reduce investment. We utilize a novel dataset on firms' investment projects that, for narrowly defined product categories, includes information on the project cost and added capacity to the firm. We break down total investment project cost into two components: investment's production capacity (e.g., number of machinery), defined as the additional capacity generated by the investment, and unit investment cost (e.g., price of machinery), defined as the total cost divided by the additional capacity. We document substantial variation in unit investment cost across firms within narrowly defined product categories. Next, we hypothesize that higher unit investment cost is associated with lower depreciation rates (or higher repair and maintenance cost) or higher "quality" output, and provide evidence consistent with both hypotheses. Finally, we show that firms that become financially constrained reduce both unit investment cost and investment's production capacity.[1]

[1]We grateful to Olivier Darmouni, Xavier Giroud, and David Weinstein, Charles Calomiris, Neng Wang, Daniel Wolfenzon, and Gur Huberman for their valuable advice and suggestions.

3.1 Introduction

Investment decisions are arguably one of the most critical decisions made by firms. Consequently, substantial empirical and theoretical literature in financial economics has focused on investments. The primary and probably yet unsettled debate in the empirical literature is how to measure financial constraints. The consensus, however, is that financially constrained firms reduce investment. However, even if we agree on the "true" measure of financial constraints, we won't be able to say if financially constrained firms are indeed *affected* if we don't have data on the details of investment projects. Why? Because financially constrained firms might be able to purchase cheaper investment options that, for example, depreciate faster and compensate later on – when their constraints are potentially relaxed – by investing more. By purchasing cheaper investment options, they might fully alleviate the effects of financial constraints on output. But, do financially constrained firms, indeed, purchase cheaper investment options? If so, what are the potential losses associated with cheaper investment options? In this paper, we seek to answer these questions using a unique dataset that includes the details of firm-level investment projects.

In Table 3.1, we provide examples of three potential scenarios to illustrate *how* financially constrained firms can reduce investment, and potentially mitigate the adverse effects of financial constraints by choosing a cheaper investment option. Before explaining the examples, it is useful to define:

$$\text{Unit Investment Cost (UIC)} = \frac{\text{Total Investment Cost}}{\text{Production Capacity Added to the Firm}} \tag{3.1}$$

Firms invest to create production capacity, and UIC captures how costly it is for a firm to create the capacity to produce one unit of output. In theory, for producing the same output, firms might be able to choose from a menu of investment options with different UICs. In this paper, we argue that UIC is an important margin of adjustment for firm investment.

How? In Example 1 of Table 3.1, the constrained firm lowers investment by only adjusting the additional capacity. In Examples 2 and 3, the constrained firm lowers investment by only adjusting the UIC.

In Example 1, the constrained firm lowers investment by only adjusting the additional capacity. Both the unconstrained and constrained firm choose investment options with the same UIC. Thus all the fall in investment is reflected in the firm's additional capacity. Thus, a financially constrained firm's ability to produce future output is impaired. In Examples 2 and 3, the constrained firm lowers investment by only adjusting the UIC. Thus, the firm's ability to produce future output is *not affected*. The difference between Examples 2 and 3 comes from *why* UIC is different for investment options chosen by the constrained and unconstrained firm. In Example 2, the lower UIC option depreciates faster (or it needs higher future maintenance cost). In Example 3, the higher UIC option produces higher "quality" output. In Examples 2 and 3, the constrained firm fully mitigates the effects of financial constraints on future output "quantity" by choosing a cheaper investment option in the short term.

Data availability on investment projects is an important impediment in making progress in answering these questions. To address this challenge, we use a novel and unique dataset that includes information about firm-level *investment projects* in the context of India. In addition to the total investment cost, the dataset includes information about the additional capacity added to the firm and the product of the investment project. We combine this dataset by firm-level financial variables and data on firm-level prices and quantities of output.

We start by establishing that indeed UIC varies significantly even within narrowly defined product groups. If UIC is the same for every project for producing in the same product group, then total investment cost is a perfect indicator of additional capacity added to the firm. If so, firms are unable to choose cheaper investment options to mitigate the effects of financial constraints on output. Thus, as a first step, we document variation in UIC within narrowly defined product groups. In particular, we show that, on average, the ratio of the

75th percentile to the 25th percentile of UIC is approximately 3, which is quite substantial.

Next, we show that UIC has intuitive properties. To do this, we suggest two hypotheses that can potentially explain the differences in UIC, and provide empirical evidence consistent with both of these hypotheses. In the first hypothesis, a more expensive investment (investment with higher UIC) depreciates at a slower rate or needs less future maintenance cost (Example 2 in Table 3.1). In the second hypothesis, a more expensive investment option produces higher "quality" output (Example 3 in Table 3.1).

First, we provide empirical evidence consistent with the interpretation that a higher UIC option depreciates at a slower rate, or needs less future maintenance cost. We show that firm-level depreciation rates are negatively correlated with UIC within narrowly defined product categories – moving from 25th percentile to 75th percentile of UIC is associated with a .15 standard deviation decrease in depreciation rate. Furthermore, we show that firm-level maintenance cost is negatively correlated with UIC within narrowly defined product categories, as well – moving from 25th percentile to 75th percentile of UIC is associated with a .18 standard deviation decrease in repair and maintenance rate. These findings are intuitive and suggest that a higher (lower) upfront investment cost decreases (increases) future expected costs.

Second, we provide empirical evidence that a higher UIC option produces higher "quality" output. Prior literature has used different proxies to measure output "quality". We use two of those proxies. The first proxy to measure output "quality" is output unit *price*[2]. We show that within narrowly defined product categories, output price is positively correlated with UIC – a 10% increase in UIC is associated with a 0.6% increase in "quality". Even though output prices could potentially be a measure of output "quality", but they can also capture differences in price-setting behaviors of firms. Thus, we use a second proxy for output "quality". The second proxy measures "quality" as sales conditional on prices. Intuitively, when comparing two products with similar prices, the product that has higher sales should

[2]Interpreting higher unit price product as higher quality product is common in macroeconomics and trade literature. See for instance,

be of higher quality. Using this second proxy for "quality", we find that within narrowly defined product categories, output "quality" is positively correlated with UIC, as well – a 10% increase in UIC is associated with a .53% increase in "quality". These findings are intuitive and suggest that a higher (lower) upfront cost increases (decreases) future output "quality".

Finally, we provide suggestive evidence that financially constrained firms adjust investments by both cutting down on UIC and additional capacity. We offer both cross-sectional and time-series evidence. In the cross-section, we use Earnings Before Interest, Taxes, Depreciation, and Amortization (EBITDA), and firm size (proxied by total assets) as two proxies for the extent of financial constraints faced by firms. Using these two proxies, we find that within narrowly defined product categories, financially constrained firms reduce investment by cutting down on UIC – – a one standard deviation increase in EBITDA and size is associated with 8.2% and 16% increase in UIC. For time-series evidence, we use business cycles as a source tightness of financial constraints over time and find that firms choose lower UIC options in recessions. Finally, we document that Overall, these findings suggest that firms partially mitigate the effects of financial constraints on future output by choosing lower UIC options (cheaper investment options), with about half of the reduction coming from purchase of lower UIC options.

Related work: This paper contributes to several strands of literature. Perhaps the closest papers to ours are papers that study the effects of financial constraints on the "type" of capital. In Eisfeldt and Rampini (2007), capital "type" is either new or used (in Rampini (2019) capital "type" refers to the depreciation rate of capital). Eisfeldt and Rampini (2007) provide evidence that financially constrained firms employ used instead of new capital to alleviate financial constraints. Since used capital is cheaper upfront and needs future repair and maintenance, later on, the substitution helps alleviate financial constraints. First, while the distinction between used and new capital is useful, UIC is a more direct measure for capturing differences in "types" of capital. To the best of our knowledge, we are the first to highlight the importance of UIC in investment decisions. Second, the difference between new and used capital is depreciation rate in Eisfeldt and Rampini (2007) (or repair and maintenance rate in Rampini (2019)). We provide micro-level empirical evidence that supports their assumptions. Furthermore, we provide suggestive evidence that cheaper investment options produce lower "quality" output. To the best of our knowledge, we are the first to provide empirical evidence that highlight this dimension of capital.

This paper contributes to the large literature in financial economics, documenting financially constrained firms reduce investment. While the main debate in the literature is how to measure financial constraints (see for instance Fazzari, Hubbard, and Petersen (1988), Kaplan and Zingales (1997), Rauh (2006), Whited and Wu (2006), Hadlock and Pierce (2010), Campello, Graham, and Harvey (2010), Almeida and Campello (2007), Almeida, Campello, Laranjeira, and Weisbenner (2009), and Hennessy and Whited (2007)) we highlight the importance of going beyond the total investment cost to measure investment behaviors of firms. We show in the short term, financially constrained firms partially mitigate the effects of financial constraints on future output by choosing lower UIC options. By doing this, however, they are likely to be affected in the long term, since, their investment depreciates faster (needs higher future maintenance cost), or produces lower "quality" output or both.

This paper also contributes to the microeconomics literature on production functions and investment decisions. In this literature, a firm 's capital decision is single-dimensional, i.e., total investment in capital. We contribute to this literature by providing evidence consistent with a two-dimensional model of capital choice. Considering this dimension is essential in dynamic models of firm investment since firms might be able to reverse the short-term negative impacts of shocks partially, but might harm themselves in the long term. Furthermore, we provide suggestive evidence of the potential harms of choosing a cheaper investment option.

Furthermore, it contributes to the literature in international trade and macroeconomics that document the relationship between firm size, input, and output quality. This literature suggests that larger firms purchase higher quality variables input and produce higher quality output. We contribute to this literature by showing larger firms purchase higher UIC options, which depreciates at a slower rate, needs lower future maintenance cost, and produces higher "quality" output.

Furthermore, this paper contributes to the literature that investigates differences in output "quality" across firms (for instance, see Kugler and Verhoogen (2012), Goldberg, Khandelwal, Pavcnik, and Topalova (2010)). We suggest a novel channel that emphasizes the role of capital in the output "quality".

The rest of this article is organized as follows. Section 3.2 describes the data, and provides summary statistics. Section 3.3 documents variation in UIC. Section 3.4 provides evidence consistent with two potential hypothesis that can explain the variation in UIC. Section 3.5 shows that financially constrained firms choose lower UIC options. Section 3.6 decomposes total investment cost into two components, and Section 3.7 concludes.

3.2 Data

3.2.1 CapEx dataset

We use CapEx dataset to obtain detailed data on firm-level investment projects. This dataset is provided by the Center for Monitoring the Indian Economy (CMIE)[3], and provides data on planned capital expenditure at the project level. Investment projects covered in the dataset are real investments that involve capacity expansion, as opposed to financial investments. The dataset includes information about projects that cost more than 10 million Indian rupees. The dataset covers projects announced from 1996 to 2018, and has an overall of 66,000 different projects. The dataset covers a wide range of projects in different industries and across India. These projects are large and a significant part of firm's total expenditure. Depending on year these projects constitute of about 45% to 75% of firm-level

[3]CMIE is a private company that collects information about Indian firms. The CMIE also collects the Prowess dataset widely used in economics and finance.

annual expenditure.

The dataset is collected from publicly available data sources such as newspaper, company websites, promoter's website, and project contractors. Furthermore, data is also collected through direct contact with companies engaged in executing the projects. Data is collected throughout the course of the project until the project is completed or abandoned.

The dataset includes detailed information about firm-level investment projects. Most importantly, in addition to total project cost, we can observe additional capacity added to the firm and the product that the firm can produce as a result of the investment. This level of information is unique and we are not aware of any other dataset that includes such information. Each project creates additional capacity to produce products. 99% of projects create capacity to produce in a single product category[4].

CapEx also includes data on the status of projects over time. Each project has an announcement date. After the initial announcement the dataset keeps track of project status. Projects are discontinued for various reasons. Over the sample period 58% of projects become complete. The rest are either abandoned (7.1%), stalled (5.4%) or have on going status (23.2%). Not all projects are undertaken by firms. The dataset also includes information about investment projects undertaken by states or central government.

Definition of a project can be somewhat arbitrary. An investment project can consist of several investment projects done in different locations or at different stages[5]. As a rule of thumb, CMIE records components of "one" project as separate projects if it can obtain separate cost data for each component, or if the company itself, treats them as separate projects. We use the definition of CMIE to identify a project[6].

Table 3.2 is a sample project in the CapEx dataset. The project "Haldwani Dry Grinding Talc Project" is undertaken by company "20 Microns Ltd.". This projects is announced on March 2008 and completed on March 2009, and costs 48.6 Million Indian rupees. The project

[4]We drop the other 1% since it is impossible to calculate UIC for a single product category.

[5]About 1.5% of the projects in CapEx have more than one phase.

[6]The main results would remain the same if we treat all projects undertaken by firm to produce one product in a single year as one project.

is located in Gujarat, and adds a total capacity of 7,000 Tonnes for production of Talc to the company.

Capacity added to the firm might be measured in different units even for one product category. For instance, projects for production of sunflower oil might record capacity in "Tonnes Per Day" or "Kg Per Day". We manually convert measurement units to a standard format for each product category, whenever possible. For some products measurement units this is not possible. For instance, a project to build a hotel project might be recorded in "Number of Rooms" or "Square Meters"[7]. We consider products in the same product category with different measurement units as distinct products.

3.2.2 Prowess dataset

We gather firm-level financial data from the Prowess dataset, also maintained by the CMIE[8]. Prowess has annual financial data for listed and unlisted Indian firms starting from 1989 for a wide range of firms in different industries. While it covers the universe of listed firms, its coverage for unlisted firms is less comprehensive. The firms in the dataset cover a significant portion of the economic activity in the organized industrial sector, between 60% to 70%[9].

What makes Prowess unique relative to other firm-level financial datasets is the presence of product-level data. In particular, for narrowly defined product categories, it includes information about sales and output quantity[10]. This data is publicly available since Indian firms are required by the 1956 Companies Act to disclose product-level information on capacities, production and sales in their annual reports. CMIE uses an internal product classification that is based on the National Industry Classification (NIC) schedules. Overall, there are a

[7]For less than 2% of the projects we can't convert measurement units to get a single measurement unit for one product category.

[8]The Prowess dataset has been used previously. See for instance Vig (2013), Gopalan, Mukherjee, and Singh (2016), and Goldberg, Khandelwal, Pavcnik, and Topalova (2010).

[9]Furthermore, firms in the Prowess sample account for 75 percent of corporate taxes and 95 percent of excise duty collected by the Government of India.

[10]The Prowess dataset is the only dataset in India that records detailed annual information on firms' product-mix in large scale

total of 1,712 products linked to 137 four-digit NIC industries across the 22 manufacturing sectors (two-digit NIC codes).

We take the following steps to clean the data. From the overall Prowess sample of 1989 to 2019, we exclude all financial firms (NIC code: 641-663), firms owned by central and state governments, firms with less than three years of data with positive values of total assets, firms with leverage more than unity, and observations with ratio of investment to lagged total assets greater than 1. Furthermore, we restrict the sample to firms with a book value of total assets greater than INR 10 million, since project total cost is . To mitigate outliers, we require that the firm's capital and sales be at least INR 1 million (around US$ 0.02 million) in the previous year. The financial year in India starts from April 1 to March 31. We make necessary adjustments to reflect this in our analysis. We limit our CapEx sample to projects with known announcement date, total cost and positive new capacity. We only focus on subsample of projects completed before or in 2018. Furthermore, we assume that firms invest the total investment amount equally over the life of a project. For instance, if a project takes two years to complete, we assume that firm invests half of the amount each year. We use common company identifier shared by CapEx and Prowess to merge these datasets.

The final sample includes about 2,500 firms, and 7,500 firm-year observations. These firms completed about 11,000 projects in approximately 1,000 different product categories over the entire sample. Tables 3.3 and Table 3.4 include description of the companies and projects employed in the study are provided in, respectively.

In Table 3.3, we present summary statistics of some financial variables used in the analysis. We also report separate statistics for subsample of listed and unlisted firms. The mean and median firm size, measured by total assets, is 1,023 million INR (22.7 million USD in 2010), and 145 million INR (3.2 million USD in 2010), respectively. The average and median EBITDA are 323 million INR (7.2 million USD in 2010) and 17 million INR (.38 million USD in 2010). Average firm-level leverage, measured by total debt divided by total assets, is

35%. Average asset tangibility, measured by Property, Plant and Equipment (PPE) divided by total assets, is around 32%.

In Table 3.4, we provide the descriptive statistics for sample of projects. We also report separate statistics for subsample of listed and unlisted firms. On average a project costs about 6 billion INR (133 million USD). The average duration of a project, the period from project announcement date to project completion date, is 2.2 years. About 64% of projects are completed. Also, nearly 61% of all projects are projects with less than a year.

3.3 Variation in UIC

Firms invest to create capacity to produce output. They purchase machine and build factories to create and expand their ability produce products in future. In doing so, they might have different options. These differences could be differences physical attributes of fixed capital or technologies. UIC defined in 3.1 aggregates these different dimensions into a single variable. It captures how costly it is for a firm to create the capacity for production of one unit of output. The heterogeneities in firm capital are likely to have consequences for firm growth, as well. UICs can also be interpreted as "unit price" of capital for producing output.

Even though, in theory, firms can choose between investment options with differing UICs, it is not clear whether, there is indeed significant variation in UIC in the real worl. Thus, as a first step, we document the extent of variation in UIC within narrowly defined product categories. To investigate the degree of variation in UICs, let's define:

$$r_x^y(l,t) = \frac{y\text{th percentile of }\{\text{UIC of investment for product l at time t}\}}{x\text{th percentile of }\{\text{UIC of investment for product l at time t}\}}$$

Consider the set of investment projects for producing product l at time t, then, $r_x^y(l,t)$ is the ratio of yth percentile of UIC to the ratio of xth of UIC for these projects. Intuitively, $r_x^y(l,t)$ measures the extent of variation in UIC for product l at time t.

Table 3.5 shows significant variation in $r_x^y(l,t)$. The Table reports summary statistics for variable $r_x^y(l,t)$. Each statistic is computed using observations across all different products and over the entire period. For calculating the statistics, we limit the sample to product-year pairs (l,t), where are at least ten firms invest in product l at time t[11]. On average, the 75th percentile of UIC is three times larger than the $25th$ percentile. In other words, a financially constrained firm can invest one-third of an unconstrained firm without affecting the quantity of future output in the short term. Besides, the standard deviation is not trivial and shows that there substantial variation in this ratio, as well. A similar picture emerges by looking at the 90th relative to 50th percentile, as well.

3.4 What Explains Variation in UIC?

We suggest two hypotheses that can potentially explain the differences in UIC, and provide empirical evidence consistent with both of these hypotheses. In the first hypothesis, a more expensive investment (investment with higher UIC) depreciates at a slower rate or needs less future maintenance costs. In the second hypothesis, a more expensive investment produces higher "quality" output.

3.4.1 Depreciation Rate and Repair & Maintenance Cost

We begin by providing empirical evidence consistent with the interpretation that a higher UIC option depreciates at a slower rate. To conduct our analysis, we first define firm-level depreciation rate of capital as:

$$\text{Depreciation Rate}_{ft} = \frac{\text{Depreciation Cost}_{ft}}{\text{PPE}_{ft}}$$

Depreciation Cost_{ft} measures the average annual depreciation rate of *firm* f at time t, and $\text{PPE}_{f}t$ stands for Plant, Property, and Equipment, and measures the value of property, plant

[11]We restrict the sample to product-year observations with at least ten observations to have variation for the definition of $r_x^y(l,t)$.

and equipment owned by firm f at time t. We run the following regression to investigate whether higher UIC options depreciate at a slower rate:

$$\text{Depreciation Rate}_{ft} = \alpha_l + \alpha_t + \beta \times \ln(\text{UIC}_{lft}) + \lambda X_{ft} + \varepsilon_{lft} \tag{3.2}$$

where f, l and t index firm, product, and year, respectively; Depreciation Rate_{ft} is the depreciation rate of firm f at time t; α_l and α_t are product category and year fixed effects, respectively; and ε_{lft} is the error term, and β is the coefficient of interest. X_{ft} includes time-varying firm-level controls consisting of log(PPE) to capture differences in firm's initial property, plant and equipment, log(total assets) to capture firm size, leverage to capture differences in capital structure, EBITDA/total assets to capture differences in tangible assets, total employee compensation/PPE to capture differences in capital to labor ratio, total employee compensation/sales to proxy for relative ratio of skilled labor[12].

Having granular product category data is helpful in ruling out some potential alternative hypotheses that can explain the findings of regression 3.2. For instance, differences in depreciation rates can not be attributed to differences in production technology that comes from production of different products. Since the identification of β comes from comparing two firms offering products in the *same* product category, the prior story can not explain finding a significant β.

Regression specification 3.2 is not ideal. First, for a more direct test, we would have needed the depreciation rate of each investment project to see if higher UIC options indeed depreciate at a slower rate. However, in the absence of individual project-level depreciation rates, we use firm-level depreciation rates to proxy for the individual investment depreciation rate. Using firm-level depreciation rates should make it harder to find a significant coefficient. Thus, our finding is likely to be an underestimate of the magnitude of the coefficient. Second, since we can only observe firm-level depreciation rates, it is not clear *when* the depreciation

[12]The last two control variables are used in the prior literature, as well. For instance, look at Eisfeldt and Rampini (2007). Similar results holds if we remove these controls, as well.

rate of new investment enters a firm's balance sheet. We repeat the results with lead and lagged value of $\ln(\text{UIC}_{lft})$ and find similar results.

Table 3.6 shows UIC is negatively correlated with firm-level depreciation rate using regression specification 3.2. Columns 1 and 2 show the estimates of the regression coefficient with and without region[13] fixed effects. Both columns suggest similar findings. The point estimate suggests that a 10% increase in UIC is associated with a 0.15% lower depreciation rate. Since the average depreciation rate is 10.2, a 10% increase in UIC lowers depreciation rate by 1.5% relative to the average. Thus, a 10% increase in UIC is associated with 1.5% longer asset life, which is economically significant. Another way to interpret the magnitude is that moving from 25th percentile to 75th percentile of UIC is associated with a .15 standard deviation decrease in depreciation rate.

Next, we show that expensive investment (investment with higher UIC) needs less future repair and maintenance cost[14]. To do so, we take steps similar to the depreciation argument. We begin by defining:

$$\text{Repair \& Maintenance Rate}_{ft} = \frac{\text{Repair \& Maintenance Cost}_{ft}}{\text{PPE}_{ft}}$$

Repair & Maintenance Cost Rate_{ft} measures the average repair and maintenance cost of *firm* f at time t. In order to study whether higher UIC options have lower future maintenance cost, we run the following regression:

$$\text{Repair \& Maintenance Rate}_{ft} = \alpha_l + \alpha_t + \beta \times \ln(\text{UIC}_{lft}) + \lambda X_{ft} + \varepsilon_{lft} \tag{3.3}$$

where Repair & Maintenance Rate_{ft} is the rate of repair and maintenance for firm f at time t, and the rest of the variables are defined similar to regression specification 3.2.

[13] Region refers to thirty six geographical areas in India, including twenty eight states, and eight union territories.

[14] Repair and maintenance expenditures are expenditures made to keep the stock of fixed assets or product capacity in good working order during the life intended initially. These include costs incurred to forestall breakdowns of equipment and structures (maintenance) and costs incurred to restore fixed assets to a state of good working condition after malfunctioning (repair).

Table 3.6 shows UIC is negatively correlated with firm-level repair and maintenance rate using regression specification 3.3. Columns 1 and 2 show the estimates of the regression coefficient with and without region fixed effects. Both estimates are negative and statistically significant, and have similar magnitudes. The point estimate in column 1, suggests that a 10% increase in UIC is associated with a 0.28% lower repair and maintenance rate. Since the average repair and maintenance rate is 5.3%, a 10% increase in UIC lowers repair and maintenance rate by 5.2% on average. Thus, a 10% increase in UIC is associated with 5.2% less repair and maintenance cost, which is economically significant. Another way to interpret the magnitude is that moving from 25th percentile to 75th percentile of UIC is associated with a .18 standard deviation decrease in repair and maintenance rate.

Even though depreciation and repair and maintenance rate are conceptually similar, but these are slightly different dimensions of firm cost. Most importantly, they are reported on different balance sheet items, and thus, we can do separate tests for each variable. Furthermore, prior literature (Eisfeldt and Rampini (2007) and Rampini (2019)) also considers them as two different mechanisms.

3.4.2 Output Quality

We provide empirical evidence consistent with the interpretation that a higher UIC option produces higher "quality" output. We show that within narrowly defined product categories, output "quality" is positively correlated with UIC. To do so, we use two different proxies to measure output "quality". The first proxy is output price, and the second proxy is sales conditional on prices.

In order to investigate the relationship between UIC and output "quality", we use the following regression:

$$\text{Output Quality}_{lft} = \alpha_l + \alpha_t + \beta \times \ln(\text{UIC}_{lft}) + \lambda X_{ft} + \varepsilon_{lft} \tag{3.4}$$

where Output Quality$_{lft}$ is the quality of product l sold by firm f at time t, and the rest of the variables are defined similar to regression specification 3.2. Note that both the left-hand side variable and the right-hand side variable are at the product-level. Furthermore, by including product fixed effects, α_l, we are comparing output "qualities" of firms that invested in producing in the *same product category*. This is possible because of the level of granularity of data in both the CapEx and Prowess datasets.

Table 3.8 shows the results of regression specification in 3.4, where we use unit price as a measure of "quality". Columns 1 and 2 show the estimates of the regression coefficient with and without region fixed effects. Both columns suggest similar findings. The point estimate suggests that a 10% increase in UIC is associated with a 0.6% increase in "quality" (here, unit price).

Our findings are consistent with prior literature. As mentioned earlier, an alternative interpretation of UIC is "unit price" of fixed assets. Thus, our finding suggests that "unit price" of fixed assets is positively correlated with unit output price. Prior literature finds that unit input price is positively correlated with unit output price. For instance, Kugler and Verhoogen (2012)'s findings suggest that the unit output price is positively correlated with unit input price of *variable* inputs. Showing the positive correlation between unit output price and unit input price is challenging since, in datasets used in the literature, we don't know which input is used for the production of output. Prior literature does this indirectly by showing a positive correlation of unit input and output price with a third variable, for instance, firm size in Kugler and Verhoogen (2012). An advantage of our setup relative to other studies is having observable data on the *same product category* fixed asset that is used to produce an output.

Prices might measure output "quality", but they can also capture the price-setting behaviors of firms. Thus, we use a second proxy for output "quality". The richness of our data makes it possible to proxy for "quality" more directly. The proxy for "quality" that we will use is the residual of sales conditional on prices with a Constant Elasticity of Substitution

(CES) demand system. The higher the residual, the higher the sales conditional on prices, and thus the higher the "quality". Intuitively speaking, when comparing two products with similar prices, the product that has higher sales should be of higher quality. This intuition is used to back out "quality" from unit price and sales data available in the Prowess dataset. Measures with similar intuition have been used in the trade and macroeconomics literature, for instance, Hallak and Schott (2011), and Hottman, Redding, and Weinstein (2016b).

To define the second proxy for output "quality"[15], we use a CES demand system. We assume that firms producing products in product category g face a CES demand, with an elasticity of substitution of σ. Thus, the representative consumer's utility maximization problem is:

$$\max_{C_f} \left(\sum_{f\in\Omega_g} (Q_f C_f)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \quad \Pi_{f\in\Omega_g} Q_f = 1$$
$$\sum_{f\in\Omega_g} P_f C_f \leq K$$

where the representative consumer maximizes utility by optimally choosing consumption level of products offered by firm f, C_f, subject to the budget constraint $\Sigma_{f\in\Omega_g} P_f C_f \leq K$. Ω_g is the set of *all* firms offering products in product category g, P_f is the unit price of product f, and Q_f is the "quality" of the product offered by firm f and σ is the elasticity of substitution and measures the degree of substitution between products offered by different firms. A one-percent decrease in the unit price of firm f increases its sale by $\sigma - 1$. Thus, the higher the σ, the more sensitive the representative consumer is to unit prices.

Solving the optimization problem to find the optimal C_f, and rearranging:

[15] "Quality" here can have two interpretations. It can capture differences in consumer tastes. For instance, more people might prefer black hard drives rather than green hard drives. It can, however, capture differences in ranking common across all consumers. For instance, a faster hard drive has higher quality relative to a slower one.

$$\ln(Q_f) = \frac{\sigma}{\sigma - 1}\ln(P_f) + \frac{1}{\sigma - 1}\ln\left(\frac{P_f C_f^*}{\Sigma_g P_f C_f^*}\right) + \frac{1}{\sigma - 1}\ln\left(\Sigma_f P_f^{-\sigma} Q_f^{\sigma-1}\right) \tag{3.5}$$

where C_f^* is the optimal consumption of firm f's product in product category g. The equation has an intuitive interpretation: if two products have the same unit prices (P_f), then the product with higher "quality" (Q_f) will have higher sales share $\left(\ln\left(\frac{P_f C_f^*}{\Sigma_g P_f C_f^*}\right)\right)$. Furthermore, the more substitutable the products are (higher σ), the higher the sales share of a product with higher "quality"[16]. Another way of interpreting the above equation is that output "quality" is the residual of sales share of product f, conditional on unit price (P_f).

The set of equations in 3.5 is used to back out output "quality" from unit price and sales data. If we assume a value for σ, then the only unknown for each firm is product quality, Q_f. Since there is one equation for each unknown, there exists a unique solution that satisfies Equation 3.5 for every firm[17].

Table 3.9 shows the results of regression specification 3.4, where output "quality" is measured using $\ln(Q_f)$ defined in Equation 3.5. To do the exercise, we choose three different values for $\sigma \in \{3, 5, 10\}$[18]. Across all different values of σ, the coefficient of the regression coefficient is positive and statistically significant, which suggests firms with higher UIC, produce higher quality products in the same *product category.* Interestingly, the magnitude of the estimated coefficient doesn't depend on σ. The point estimate suggests that a 10% increase in UIC is associated with a .53% increase in "quality" for $\sigma = 5$.

[16]The term $\frac{1}{\sigma-1}\ln\left(\Sigma_{f\in\Omega_g} P_f^{-\sigma} Q_f^{\sigma-1}\right)$ is constant for all products offered in the same product category. The constant is such that $\sum_{f\in\Omega_g}\ln(Q_f) = 0$, since we have assumed $\Pi_{f\in\Omega_g} Q_f = 1$.

[17]To be very precise since we have assumed $\Pi_{f\in\Omega_g} Q_f = 1$, the number of unknowns is equal to the number of equations minus one. However, the equations are not independent either. In particular, since $\sum_{f\in\Omega_g}\ln(Q_f) = 0$, the number of independent observations is the number of equations minus one, as well. Thus, the number of unknowns is equal to the number of independent equations, and thus a unique solution exists.

[18]Our choice of these values for σ is motivated by the macroeconomics and international trade literature, where the CES demand system is widely used. Our product-level data is very granular, and the estimated elasticities in trade literature for this level of granularity tend to be high. See for example, see Broda and Weinstein (2006b).

3.5 Financial Constraints and UIC

In this section, we provide suggestive evidence that financially constrained firms choose lower UIC options. We offer both cross-sectional and time-series evidence. In the cross-section, we use EBITDA, and firm size (proxied by total assets) as two proxies for the extent of financial constraints faced by firms. Using these two measures, we find that within narrowly defined product categories, financially constrained firms choose lower UIC options. For time-series evidence, we use business cycles as a source tightness of financial constraints over time and find that firms choose lower UIC options in recessions. Moreover, we show that financially constrained firms choose lower UIC options in recessions relative to booms. Overall, these findings suggest that firms partially mitigate the effects of financial constraints on future output by choosing lower UIC options (cheaper investment options).

In this section, we provide suggestive evidence, both in the cross-section and time-series, that financially constrained firms choose lower UIC options.

3.5.1 Cross-Sectional Evidence

To investigate how financially constrained firms adjust UIC, we use the following cross-sectional regression specification[19]:

$$\ln(\text{UIC}_{lft}) = \alpha_l + \alpha_t + \beta \times \text{Measure of Financial Constraint}_{ft} + \lambda X_{ft} + \varepsilon_{lft} \tag{3.6}$$

where we use size and EBITDA for the Measure of Financial Constraint. Other variables are the same as regression specification 3.2.

Table 3.10 shows the results of regression specification 3.6. Columns 1, 2, and 3 shows the estimated coefficient using EBITDA, size, and both EBITDA and size together. All the estimated coefficients are negative and statistically significant. The point estimate suggests

[19] We estimate the regression over time; however, since the primary source of variation is the cross-sectional variation of the financial constraints measure, we refer to this regression as cross-sectional.

that a one standard deviation increase in EBITDA is associated with a 8.2% increase in UIC. Similarly, a 10% increase in firm size is associated with 0.67% increase in UIC.

3.5.2 Time-Series Evidence

To provide time-series evidence, we use recessions as a source of time-series variation in financial constraints. We interact a recession dummy with firm-level EBITDA. If financial constraints are tightened during recessions, we would expect the UIC to be more sensitive to cash-flows in recessions. We use the following regression specification:

$$\ln(\text{UIC}_{lft}) = \alpha_l + \alpha_t + \beta_1 \text{EBITDA}_{ft} + \beta_2 \text{EBITDA}_{ft} \times \text{Recession Dummy}_t + \lambda X_{ft} + \varepsilon_{lft} \quad (3.7)$$

where Recession Dummy_t is a dummy variable that takes the value of one for recessions and zero, otherwise. Other variables are the same as regression specification 3.2.

We get data on recession dates from the Federal Reserve Bank of St. Louis website[20]. During our sample, there are five recessions: (1) June 1996 - October 1997; (2) January 2000 - January 2003; (3) October 2007 - May 2009; and (4) January 2011 - June 2013; (5) and May 2018 - end of the sample, where the second and third recessions were more severe than others. Thus, we classify the years 1996 - 1997, 2000 - 2002, 2008 - 2009, and 2011 - 2012 and 2018 as recession years. We identify recessions annually since our firm-level data is available at an annual frequency. Doing so reduces the power of our tests since some pre and post-recession months are classified as recessions, and some recession months are not classified as a recession.

Table 3.11 shows the result of running regression specification3.7[21]. The interaction term is positive and statistically significant, suggesting that investments are more sensitive to cashflows during recessions.

[20]https://fred.stlouisfed.org/series/INDREC

[21]The results are robust if we use $\overline{\text{EBITDA}}_{ft} = \frac{1}{2}\left(\text{EBITDA}_{ft} + \text{EBITDA}_{ft-1}\right)/\text{Asset}_{ft-1}$, instead of EBITDA_t.

3.6 Investment Decomposition

In this section, we investigate how financially constrained firms respond to financial constraints. To do so, let's define:

$$\text{Investment Production Capacity (IPC)} = \frac{\text{Production Capacity Added to the Firm}}{\text{Duration of the Project}} \tag{3.8}$$

IPC measures the capacity added to the firm in *one* year, assuming that the project is divided into equal pieces over the implementation duration. The reason that we make this assumption is to make this comparable to annual firm-level investment regressions. Even though most projects in the sample take less than a year to complete, some are long term projects that take several years to complete. Thus, dividing by the duration of the project is a way to normalize the project cost to annual frequency.

To see how financially constrained firms trade-off between UIC and Investment Production Capacity (IPC), we run the following regression:

$$y_{lft} = \alpha_l + \alpha_t + \beta \times \text{Measure of Financial Constraint}_{ft} + \lambda X_{ft} + \varepsilon_{lft} \tag{3.9}$$

where y_{lft} is the outcome variable of interest. The outcome variable is either $\ln(\text{UIC}_{lft})$, $\ln(\text{IPC}_{lft})$, or $\ln(\text{Total Investment Cost}) = \ln(\text{UIC}_{lft})+\ln(\text{IPC}_{lft})$. Measure of Financial Constraint$_{ft}$ is firm-level financial constraints faced by firm f at time t. The rest of the variables are the same as regression specification 3.2.

Table 3.12 shows the result of regression specification 3.9. The first two columns show how financially constrained firms reduce investment by adjusting UIC and IPC. Both columns are statistically significant[22]. The estimates suggest that financially constrained firms reduce

[22] We can't calculate Tobin's Q for the whole sample, since, unlisted firms are also included in our sample. However, limiting the sample to listed firms, and using Tobin's Q as a control variable, doesn't change the result.

investment by both cutting down on UIC and IPC. 53% of investment adjustment comes from purchasing lower UIC options. Thus, UIC is a critical margin of adjustment for financially constrained firms.

3.7 Conclusion

This paper shows that for studying firm-level investment decisions, it is critical to have data on the details on investment projects. In particular, financially constrained firms choose cheaper investment options (lower UIC) to mitigate the effects of financial constraints on future output partially. Investing in cheaper investment options, however, is costly. The investment depreciates faster, needs higher future repair and maintenance cost, and produces lower "quality" output.

		Cost	Output	Unit Investment Cost	Depreciation	Quality
Example 1	Unconstrained Firm	\$100	10 widgets/year	\$10 per widget-year		
	Constrained Firm	\$50	5 widgets/year	\$10 per widget-year		
Example 2	Unconstrained Firm	\$100	10 widgets/year	\$10 per widget-year	2 years	
	Constrained Firm	\$50	10 widgets/year	\$5 per widget-year	1 year	
Example 3	Unconstrained Firm	\$100	10 widgets/year	\$10 per widget-year	1 year	High
	Constrained Firm	\$50	10 widgets/year	\$5 per widget-year	1 year	Low

Table 3.1: Three Potential Scenarios for Investment Reduction

This table illustrates three potential scenarios where firms that become financially constrained can reduce investment.

Company	Product	Product code	Announcement	Completion	Cost
20 Microns Ltd.	Talc	50280416160000000000	March 2008	March 2009	48.6 Million INR

State of Completion	Location	New Capacity	Unit	Type	Industry
Completed	Gujarat	7	'000 Tonnes	New Unit	Cosmetics & Detergents

Table 3.2: A Sample Project

This table presents part of the information available in the dataset for a sample project ("Haldwani Dry Grinding Talc Project") undertaken by "20 Microns Ltd." company.

Variables	All Firms	Listed Firms	Unlisted Firms
Total Assets (Million INR)	1,023	1,863	414.8
	(4,487)	(6,574)	(1,592)
Observations	7,537	3,166	4,371
$\text{EBITDA}_t/\text{Asset}_{t-1}$	0.161	0.187	0.139
	(0.191)	(0.132)	(0.228)
Observations	6,575	3,123	3,452
Leverage	0.351	0.329	0.372
	(0.202)	(0.180)	(0.218)
Observations	6,468	3,053	3,415
Log(EBITDA) (Million INR)	2.545	3.654	1.541
	(2.378)	(1.870)	(2.344)
Observations	6,575	3,123	3,452
$\text{Depreciation}_t/\text{PPE}_t$	0.102	0.0961	0.107
	(0.088)	(0.648)	(0.104)
Observations	6,782	3,154	3,628
Repair & Maintenance $\text{Cost}_t/\text{PPE}_t$	0.053	0.038	0.067
	(0.139)	(0.417)	(0.188)
Observations	6,478	3,128	3,350
$\text{PPE}_t/\text{Asset}_t$	0.324	0.344	0.309
	(0.209)	(0.178)	(0.230)
Observations	7,295	3,161	4,134
# Firms	2,456	1,108	1,348
# Firm-Year	7,537	3,166	4,371
# Projects	11,051	5,220	5,851
# Products	1,039	736	747

Table 3.3: Firm-Level Summary Statistics

This table reports summary statistics of financial variables for the sample of firms used in this paper (in 2010, 1 USD = 45 Indian Rupees (INR)).

Variables	All Projects	Listed Firms	Unlisted Firms
Total Project Cost (Million INR)	5,935 (9,635)	5,412 (9,123)	6,315 (10,171)
Duration (Years)	2.217 (2.512)	1.987 (2.123)	2.516 (2.871)
Total Project Cost/Duration (Million INR/Years)	2,812 (4,712)	2,973 (5,314)	2,513 (5,435)
Median $\{$(Total Project Cost/Duration)/Total Asset$\}$	0.13	0.11	0.16
Median $\{$(Cost of Project/Duration)/PPE $\}$	0.43	0.37	0.49
% Project Completed	64.3%	66.7%	59.3%
% Projects with Completed in Less than One Year	60.7%	62.4%	57.7%

Table 3.4: Project-Level Summary Statistics

This table reports summary statistics of projects completed by the sample of firms used in the paper. Duration refers to the length of project from announcement to completion date. The sample period is 1997 to 2018.

Variables	Mean	Median	StD
$r_{25}^{75}(.,.)$	3.20	2.30	1.74
$\log\left(r_{25}^{75}(.,.)\right)$	.87	.71	.60
$r_{50}^{90}(.,.)$	3.25	1.98	1.79
$\log\left(r_{50}^{90}(.,.)\right)$	.87	.68	.54

Table 3.5: Variation in UIC

This table reports summary statistics of $r_x^y(l,t)$. Each statistic is computed using observations across all different products and over the entire time period.

	ln(Depreciation Rate)	
	(1)	(2)
ln(UIC)	$-.015^{**}$	$-.015^{**}$
	(.006)	(.006)
Controls	✓	✓
Year FE	✓	✓
Product FE	✓	✓
Region FE	×	✓
Observations	10,275	10,275
R^2	0.26	0.31

Table 3.6: UIC and Depreciation Rate

*, **, and *** indicate significance at the 0.1, 0.05 and 0.01 level. This table presents estimates of regression 3.2 where depreciation rate is at firm-level, and UIC is at product category level. Controls include: log(PPE), log(total assets), leverage, EBITDA/total assets, total employee compensation/PPE, and total employee compensation/sales. Standard errors are double-clustered by product category and year.

	ln(Repair & Maintenance Rate)	
	(1)	(2)
ln(UIC)	$-.024^{**}$	$-.028^{**}$
	(.009)	(.011)
Controls	✓	✓
Year FE	✓	✓
Product FE	✓	✓
Region FE	×	✓
Observations	10,055	10,055
R^2	0.35	0.38

Table 3.7: UIC and Repair & Maintenance Cost

*, **, and *** indicate significance at the 0.1, 0.05 and 0.01 level. This table presents estimates of regression 3.3 where repair & maintenance rate is at firm-level, and UIC is at product category level. Controls include: log(PPE), log(total assets), leverage, EBITDA/total assets, total employee compensation/PPE, and total employee compensation/sales. Standard errors are double-clustered by product category and year.

	ln(Output Quality)	
	(1)	(2)
ln(UIC)	.061***	.059***
	(.016)	(.014)
Controls	✓	✓
Year FE	✓	✓
Product FE	✓	✓
Region FE	×	✓
Observations	2,371	2,371
R^2	0.97	0.97

Table 3.8: UIC and Output Quality Proxied by Output Price

*, **, and *** indicate significance at the 0.1, 0.05 and 0.01 level. This table presents estimates of regression 3.4 where output "quality" is proxied by output price. Controls include: log(PPE), log(total assets), leverage, EBITDA/total assets, total employee compensation/PPE, and total employee compensation/sales. Standard errors are double-clustered by product category and year.

	ln(Output Quality)		
	$\sigma = 3$	$\sigma = 5$	$\sigma = 10$
ln(UIC)	.056**	.053**	.052**
	(.019)	(.017)	(.016)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Product FE	✓	✓	✓
Region FE	×	×	×
Observations	2,371	2,371	2,371
R^2	0.93	0.95	0.96

Table 3.9: UIC and Output Quality Calculated from Equation 3.5

*, **, and *** indicate significance at the 0.1, 0.05 and 0.01 level. This table presents estimates of regression 3.4 where output "quality" is calculated using Equation 3.5. Controls include: log(PPE), log(total assets), leverage, EBITDA/total assets, total employee compensation/PPE, and total employee compensation/sales. Standard errors are double-clustered by product category and year.

	ln(UIC)		
$\text{EBITDA}_{ft}/\text{Asset}_{ft-1}$	.435***		.467***
	(.150)		(.156)
Size_{ft}		.067***	.068***
		(.010)	(.010)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Product FE	✓	✓	✓
Region FE	✓	✓	✓
Observations	10,179	10,179	10,179
R^2	0.82	0.82	0.83

Table 3.10: UIC and Financial constraint

*, **, and *** indicate significance at the 0.1, 0.05 and 0.01 level. This table shows the result of regression specification 3.6. Controls include: log(PPE), log(total assets), leverage, EBITDA/total assets, total employee compensation/PPE, and total employee compensation/sales. Standard errors are double-clustered by product category and year.

	ln(UIC)	
	(1)	(2)
EBITDA_{ft}	.435***	.501***
	(.150)	(.161)
$\text{EBITDA}_{ft} \times \text{Recession Dummy}_t$		.510**
		(.230)
Controls	✓	✓
Year FE	✓	✓
Project FE	✓	✓
Region FE	✓	✓
Observations	10,179	10,179
R^2	0.83	0.84

Table 3.11: UIC and Cash-Flow Sensitivity During Recessions

*, **, and *** indicate significance at the 0.1, 0.05 and 0.01 level. This table shows the result of regression specification 3.7. Controls include: total assets, EBITDA, leverage and PPE. Standard errors are double-clustered by product category and year.

	ln(UIC)	ln(IPC)	ln(Investment)
EBITDA_{ft}	$.435^{***}$	$.341^{**}$	$.802^{***}$
	(.150)	(.178)	(.255)
Controls	✓	✓	✓
Year FE	✓	✓	✓
Product FE	✓	✓	✓
Region FE	✓	✓	✓
Observations	10,179	10,179	10,179
R^2	0.83	0.73	0.41

Table 3.12: Financial Constraints and Investment Decomposition

*, **, and *** indicate significance at the 0.1, 0.05 and 0.01 level. This table shows the result of regression specification 3.9. Controls include: log(PPE), log(total assets), leverage, EBITDA/total assets, total employee compensation/PPE, and total employee compensation/sales. Standard errors are double-clustered by product category and year.

References

[1] G. Chodorow-Reich, “The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis,” *The Quarterly Journal of Economics*, vol. 129, no. 1, pp. 1–59, 2013.

[2] A. Mian, K. Rao, and A. Sufi, “Household balance sheets, consumption, and the economic slump,” *The Quarterly Journal of Economics*, vol. 128, no. 4, pp. 1687–1726, 2013.

[3] M. Di Maggio, A. Kermani, B. J. Keys, T. Piskorski, R. Ramcharan, A. Seru, and V. Yao, “Interest rate pass-through: Mortgage rates, household consumption, and voluntary deleveraging,” *American Economic Review*, vol. 107, no. 11, pp. 3550–88, 2017.

[4] A. Mian, A. Sufi, and F. Trebbi, “Foreclosures, house prices, and the real economy,” *The Journal of Finance*, vol. 70, no. 6, pp. 2587–2634, 2015.

[5] A. Mian and A. Sufi, “The effects of fiscal stimulus: Evidence from the 2009 cash for clunkers program,” *The Quarterly journal of economics*, vol. 127, no. 3, pp. 1107–1142, 2012.

[6] S. Agarwal, S. Chomsisengphet, N. Mahoney, and J. Stroebel, “Do banks pass through credit expansions to consumers who want to borrow?” *The Quarterly Journal of Economics*, vol. 133, no. 1, pp. 129–190, 2017.

[7] E. Benmelech, R. R. Meisenzahl, and R. Ramcharan, “The real effects of liquidity during the financial crisis: Evidence from automobiles,” *The Quarterly Journal of Economics*, vol. 132, no. 1, pp. 317–365, 2017.

[8] X. Giroud and H. M. Mueller, “Firm leverage, consumer demand, and employment losses during the great recession,” *The Quarterly Journal of Economics*, vol. 132, no. 1, pp. 271–316, 2017.

[9] ——, “Firms’ internal networks and local economic shocks,” *American Economic Review*, vol. 109, no. 10, pp. 3617–49, 2019.

[10] M. Greenstone, A. Mas, and H.-L. Nguyen, “Do credit market shocks affect the real economy? quasi-experimental evidence from the great recession and ‘normal’economic times,” National Bureau of Economic Research, Tech. Rep., 2014.

[11] K. Huber, “Disentangling the effects of a banking crisis: Evidence from german firms and counties,” *American Economic Review*, vol. 108, no. 3, pp. 868–98, 2018.

[12] H. Almeida, M. Campello, B. Laranjeira, and S. Weisbenner, "Corporate debt maturity and the real effects of the 2007 credit crisis," *Critical Finance Review*, vol. 1, no. 1, pp. 3–58, 2012.

[13] M. Amiti and D. E. Weinstein, "How much do idiosyncratic bank shocks affect investment? evidence from matched bank-firm loan data," *Journal of Political Economy*, vol. 126, no. 2, pp. 525–587, 2018.

[14] S. Gilchrist, R. Schoenle, J. Sim, and E. Zakrajšek, "Inflation dynamics during the financial crisis," *American Economic Review*, vol. 107, no. 3, pp. 785–823, 2017.

[15] R. Kim, "The effect of the credit crunch on output price dynamics: The corporate inventory and liquidity management channel," *Available at SSRN 3163872*, 2018.

[16] S. Borenstein and N. L. Rose, "Bankruptcy and pricing behavior in us airline markets," *The American Economic Review*, vol. 85, no. 2, pp. 397–402, 1995.

[17] J. Stroebel and J. Vavra, "House prices, local demand, and retail prices," *Journal of Political Economy*, vol. 127, no. 3, pp. 1391–1436, 2019.

[18] J. A. Chevalier and D. S. Scharfstein, "Capital market imperfections and countercyclical markups: Theory and evidence," National Bureau of Economic Research, Tech. Rep., 1994.

[19] ——, "Liquidity constraints and the cyclical behavior of markups," *The American Economic Review*, vol. 85, no. 2, pp. 390–396, 1995.

[20] A. I. Khwaja and A. Mian, "Tracing the impact of bank liquidity shocks: Evidence from an emerging market," *American Economic Review*, vol. 98, no. 4, pp. 1413–42, 2008.

[21] B. Faber and T. Fally, "Firm heterogeneity in consumption baskets: Evidence from home and store scanner data," National Bureau of Economic Research, Tech. Rep., 2017.

[22] X. Jaravel, "The unequal gains from product innovations: Evidence from the us retail sector," *The Quarterly Journal of Economics*, vol. 134, no. 2, pp. 715–783, 2018.

[23] C. J. Hottman, S. J. Redding, and D. E. Weinstein, "Quantifying the sources of firm heterogeneity," *The Quarterly Journal of Economics*, vol. 131, no. 3, pp. 1291–1364, 2016.

[24] S. J. Redding and D. E. Weinstein, "Measuring aggregate price indexes with demand shocks: Theory and evidence for ces preferences," National Bureau of Economic Research, Tech. Rep., 2016.

[25] C. Broda and D. E. Weinstein, "Product creation and destruction: Evidence and price implications," *American Economic Review*, vol. 100, no. 3, pp. 691–723, 2010.

[26] J. Schmookler, "Economic sources of inventive activity," *The Journal of Economic History*, vol. 22, no. 1, pp. 1–20, 1962.

[27] A. Shleifer, "Implementation cycles," *Journal of Political Economy*, vol. 94, no. 6, pp. 1163–1190, 1986.

[28] R. G. Rajan and L. Zingales, "Financial dependence and growth," *The American Economic Review*, vol. 88, no. 3, pp. 559–586, 1998.

[29] J. Jayaratne and P. E. Strahan, "The finance-growth nexus: Evidence from bank branch deregulation," *The Quarterly Journal of Economics*, vol. 111, no. 3, pp. 639–670, 1996.

[30] V. Acharya and Z. Xu, "Financial dependence and innovation: The case of public versus private firms," *Journal of Financial Economics*, vol. 124, no. 2, pp. 223–243, 2017.

[31] P.-H. Hsu, X. Tian, and Y. Xu, "Financial development and innovation: Cross-country evidence," *Journal of Financial Economics*, vol. 112, no. 1, pp. 116–135, 2014.

[32] M. D. Amore, C. Schneider, and A. Žaldokas, "Credit supply and corporate innovation," *Journal of Financial Economics*, vol. 109, no. 3, pp. 835–855, 2013.

[33] V. Ivashina and D. Scharfstein, "Bank lending during the financial crisis of 2008," *Journal of Financial Economics*, vol. 97, no. 3, pp. 319–338, 2010.

[34] O. Darmouni, "Informational frictions and the credit crunch," *Journal of Finance Forthcoming*, 2019.

[35] G. Chodorow-Reich and A. Falato, "The loan covenant channel: How bank health transmits to the real economy," National Bureau of Economic Research, Tech. Rep., 2017.

[36] C. Broda and D. E. Weinstein, "Globalization and the gains from variety," *The Quarterly Journal of Economics*, vol. 121, no. 2, pp. 541–585, 2006.

[37] P. K. Goldberg, A. K. Khandelwal, N. Pavcnik, and P. Topalova, "Imported intermediate inputs and domestic product growth: Evidence from india," *The Quarterly journal of economics*, vol. 125, no. 4, pp. 1727–1767, 2010.

[38] J. Handbury and D. E. Weinstein, "Goods prices and availability in cities," *The Review of Economic Studies*, vol. 82, no. 1, pp. 258–296, 2014.

[39] G. S. Crawford, N. Pavanini, and F. Schivardi, “Asymmetric information and imperfect competition in lending markets,” *American Economic Review*, vol. 108, no. 7, pp. 1659–1701, 2018.

[40] D. Green, “Corporate refinancing, covenants, and the agency cost of debt,” Working Paper, Tech. Rep., 2018.

[41] N. Cox, “The impact of risk-based pricing in the student loan market: Evidence from borrower repayment decisions,” SIEPR Working Paper, Tech. Rep., 2017.

[42] L. Einav, M. Jenkins, and J. Levin, “Contract pricing in consumer credit markets,” *Econometrica*, vol. 80, no. 4, pp. 1387–1432, 2012.

[43] S. T. Nelson, “Private information and price regulation in the us credit card market,” Ph.D. dissertation, Massachusetts Institute of Technology, 2017.

[44] M. Benetton, “Leverage regulation and market structure: An empirical model of the uk mortgage market,” *Available at SSRN 3247956*, 2018.

[45] A. L. Eisfeldt and A. A. Rampini, “New or used? investment with credit constraints,” *Journal of Monetary Economics*, vol. 54, no. 8, pp. 2656–2681, 2007.

[46] A. A. Rampini, “Financing durable assets,” *American Economic Review*, vol. 109, no. 2, pp. 664–701, 2019.

[47] S. M. Fazzari, R. G. Hubbard, and B. C. Petersen, “Financing constraints and corporate investment,” *Brookings Papers on Economic Activity*, vol. 1988, no. 1, pp. 141–206, 1988.

[48] S. N. Kaplan and L. Zingales, “Do investment-cash flow sensitivities provide useful measures of financing constraints?” *The Quarterly Journal of Economics*, vol. 112, no. 1, pp. 169–215, 1997.

[49] J. D. Rauh, “Investment and financing constraints: Evidence from the funding of corporate pension plans,” *The Journal of Finance*, vol. 61, no. 1, pp. 33–71, 2006. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2006.00829.x`.

[50] T. M. Whited and G. Wu, “Financial constraints risk,” *The Review of Financial Studies*, vol. 19, no. 2, pp. 531–559, 2006.

[51] C. J. Hadlock and J. R. Pierce, “New evidence on measuring financial constraints: Moving beyond the kz index,” *The Review of Financial Studies*, vol. 23, no. 5, pp. 1909–1940, 2010.

[52] M. Campello, J. R. Graham, and C. R. Harvey, "The real effects of financial constraints: Evidence from a financial crisis," *Journal of financial Economics*, vol. 97, no. 3, pp. 470–487, 2010.

[53] H. Almeida and M. Campello, "Financial constraints, asset tangibility, and corporate investment," *The Review of Financial Studies*, vol. 20, no. 5, pp. 1429–1460, 2007.

[54] H. Almeida, M. Campello, B. Laranjeira, and S. Weisbenner, "Corporate debt maturity and the real effects of the 2007 credit crisis," National Bureau of Economic Research, Tech. Rep., 2009.

[55] C. A. Hennessy and T. M. Whited, "How costly is external financing? evidence from a structural estimation," *The Journal of Finance*, vol. 62, no. 4, pp. 1705–1745, 2007.

[56] M. Kugler and E. Verhoogen, "Prices, Plant Size, and Product Quality," *Review of Economic Studies*, vol. 79, no. 1, pp. 307–339, 2012.

[57] V. Vig, "Access to collateral and corporate debt structure: Evidence from a natural experiment," *The Journal of Finance*, vol. 68, no. 3, pp. 881–928, 2013.

[58] R. Gopalan, A. Mukherjee, and M. Singh, "Do debt contract enforcement costs affect financing and asset structure?" *The Review of Financial Studies*, vol. 29, no. 10, pp. 2774–2813, 2016.

[59] J. C. Hallak and P. K. Schott, "Estimating Cross-Country Differences in Product Quality," *The Quarterly Journal of Economics*, vol. 126, no. 1, pp. 417–474, 2011.

[60] C. J. Hottman, S. J. Redding, and D. E. Weinstein, "Quantifying the sources of firm heterogeneity," *The Quarterly Journal of Economics*, vol. 131, no. 3, pp. 1291–1364, 2016.

[61] C. Broda and D. E. Weinstein, "Globalization and the gains from variety," *The Quarterly journal of economics*, vol. 121, no. 2, pp. 541–585, 2006.

Appendix A: Chapter 2

A.0.1 Proof of Proposition 1

First, I establish uniqueness. Since consumer demand is Cobb-Douglas across product groups, it is enough to prove the claim for one product group. Note that since markups (μ_{fgt}) is the same for all products offered by a firm in a product group, S_{ut} are unique and they can be computed by:

$$S_{ut} = \frac{(\gamma_{ut}/\varphi_{ut})^{1-\sigma_g^U}}{\sum_{k\in\Omega_{fgt}}(\gamma_{ut}/\varphi_{ut})^{1-\sigma_g^U}}$$

Thus, I only need to show that there exist unique values of S_{fgt} and μ_{fgt} that satisfy the set of equations A.1 and A.2 for all firms:

$$S_{fgt} = \frac{(P_{fgt}/\varphi_{fgt})^{1-\sigma_g^F}}{\sum_{k\in\Omega_{gt}}(P_{kgt}/\varphi_{kgt})^{1-\sigma_g^F}} \tag{A.1}$$

$$\mu_{fgt} = \frac{\sigma^F(1-S_{fgt}) + S_{fgt}}{\sigma^F(1-S_{fgt}) + S_{fgt} - 1} \tag{A.2}$$

Note that $P_{fgt} = \left(\sum_{u\in\Omega_{fgt}}(P_{ut}/\varphi_{ut})^{1-\sigma_g^U}\right)^{\frac{1}{1-\sigma_g^U}} = \mu_{fgt}\left(\sum_{u\in\Omega_{fgt}}(\gamma_{ut}/\varphi_{ut})^{1-\sigma_g^U}\right)^{\frac{1}{1-\sigma_g^U}}$. Let's denote $\xi_{fgt} = \left(\left(\sum_{u\in\Omega_{fgt}}(\gamma_{ut}/\varphi_{ut})^{1-\sigma_g^U}\right)^{\frac{1}{1-\sigma_g^U}}/\varphi_{fgt}\right)^{1-\sigma^F}$, then:

$$S_{fgt} = \frac{\xi_{fgt}\mu_{fgt}^{1-\sigma_g^F}}{\sum_{k\in\Omega_{gt}}\xi_{kgt}\mu_{kgt}^{1-\sigma_g^F}} \tag{A.3}$$

ξ_{fgt} is constant, and equations A.1 and A.3 are equivalent. Thus, it is enough to show that the set of equations A.2 and A.3 have unique solutions for μ_{fgt} and S_{fgt}. I break the uniqueness proof in two parts. First, I fix K defined as $K = \sum_{k\in\Omega_{gt}} \xi_{kgt}\mu_{kgt}^{1-\sigma_g^F}$, and show that for each K, if a solution exists, it is unique. Second, I show that there can not exist $\bar{K} \neq K$ with a solution. Combining equations A.2 and A.3, I solve for μ_{fgt} in terms of K:

$$\frac{1}{\mu_{fgt}} = 1 - \frac{1}{\frac{(1-\sigma_g^F)\mu_{fgt}^{1-\sigma_g^F}\xi_{fgt}}{K} + \sigma_g^F} \tag{A.4}$$

The left hand side is decreasing and and right side is increasing in μ_{fgt}, respectively. Thus there exists a unique μ_{fgt} that solves the above equation for each K. This completes the first part of the proof. Let's assume that there exists $\bar{K} \neq K$, and both are solutions. Denote the first set by μ_{fgt} and $K = \sum_{k\in\Omega_{gt}^F} \xi_{kgt}\mu_{kgt}^{1-\sigma_g^F}$ and the second set of solutions by $\bar{\mu}_{fgt}$, and $\bar{K} = \sum_{k\in\Omega_{gt}^F} \xi_{kgt}\bar{\mu}_{kgt}^{1-\sigma_g^F}$, and without loss of generality assume that $K < \bar{K}$. Then using equation A.4 we see $\mu_{fgt} < \bar{\mu}_{fgt}$. Thus, $K = \sum_{k\in\Omega_{gt}^F} \xi_{kgt}\mu_{kgt}^{1-\sigma_g^F} > \bar{K} = \sum_{k\in\Omega_{gt}^F} \xi_{kgt}\bar{\mu}_{kgt}^{1-\sigma_g^F}$ which contradicts the initial assumption that $K < \bar{K}$. This concludes the uniqueness part. Thus, if a solution exists it is unique. It remains to show that there exists a solution to the set of above equations. It is enough to show that the set of equations $A.4$ has an answer. S_{fgt} can then be derived from equation A.3. After manipulation of equation A.4 and substituting for K, set of equations A.4 is equivalent to:

$$\sum \mu_{fgt}^{1-\sigma_g^F}\xi_{fgt} - \frac{(1-\mu_{fgt})(1-\sigma_g^F)}{1-\sigma_g^F+\sigma_g^F\mu_{fgt}}\mu_{fgt}^{1-\sigma_g^F}\xi_{fgt} = 0$$

Note that when $\mu_{fgt} = 1+\delta$ for very small δ, then the left hand side is positive and when $\mu_{fgt} = \frac{1}{1--\frac{1}{\sigma_g^F}} - \delta$, left hand side is negative. Thus by intermediate value theorem, there exists a solution where left hand side is exactly equal to zero. This concludes the existence proof.

A.0.2 Proof of Proposition 2

This is not hard to see from the definition of the welfare index. Note that the welfare index depends on product group expenditure function, and product group expenditure function is aggregate of firm level expenditure function scaled by φ_{fgt}. Firm level expenditure function scaled by φ_{fgt} is equal to $\frac{P_{fgt}}{\varphi_{fgt}} = \left(\sum_{u\in\Omega_{fgt}}(\frac{P_{ut}}{\varphi_{fgt}\varphi_{ut}})^{1-\sigma_g^U}\right)^{\frac{1}{1-\sigma_g^U}}$, which is only a function of $\frac{P_{ut}}{\varphi_{ut}\varphi_{fgt}}$. However note that $P_{ut} = \gamma_{ut}\mu_{fgt}$, thus price is a function of $\frac{\gamma_{ut}\mu_{fgt}}{\varphi_{ut}\varphi_{fgt}}$, and from previous proposition, it can be seen that μ_{fgt} can be solved by having $\frac{\mu_{fgt}}{\varphi_{ut}\varphi_{fgt}}$. Thus the welfare index is a function of $\frac{\varphi_{ut}\varphi_{fgt}}{\gamma_{ut}}$.

A.0.3 Proof of Proposition 3

If the fixed cost of introducing different products are different then firm's optimal decision might be to increase, not change or decrease products when faced with tighter borrowing constraints. Firms decreasing or not changing products is more clear. For increase, let's consider the following hypothetical example. Assume that a firm has two choices: 1) one excellent quality product with high fixed cost of introducing, 2) two lower quality products with low fixed cost, such that the sum of the fixed cost is less than the excellent product. In this case as long as firm is able to borrow it will offer the excellent quality product, however, as soon as firm can not pay for the excellent quality product, it produces the two low quality products, and thus in this example firm products increases. When the fixed cost of introducing every product is the same, then tighter borrowing constraint leads to fewer products.

A.0.4 Proof of Proposition 4

Assume that the firm observes $\frac{\varphi_i}{\mu_i}$ and the fixed cost of introducing new products is the same. I need to show that choosing products with highest AMC is the best decision. In other words, it is most profitable for firm to offer highest AMC products. A higher AMC product benefits firm in two ways. First, its relative sales share is higher. Second, a new product increases firm's market share and thus firm charges higher markups on other products. Both of these mechanisms make it more profitable for firm to offer highest AMC products.

A.0.5 Error Term in Equation 2.5

In simulating AMCs of new products using equation 2.5, I include the error term, but why? Since the R^2 of regression 2.5 is very high it might seem plausible that the out-of-sample properties of the model does not depend on the error term. In order to answer this question, I use the simulated AMCs with and without error term and repeat the previous exercise, in which I simulated AMCs of new products in 2007 using equation 2.5. I plot the distribution of $P_{\text{g,ratio}} = \frac{P_{g,\text{real data}}}{P_{g,\text{counterfactual}}}$ with and without error term in Figure 3.1. P_{ratio} captures the ratio of the product group expenditure function in each product group. If both models are good fits, the distribution of P_{ratio} should be centered around one. Figure 3.1 shows that adding the error term increases the accuracy of the model in matching the product group expenditure function. The histogram with non-zero error is centered around one. However, the histogram with zero error term is generally less than one and is skewed to the left. The reason is as follows: the median standard deviation of error term across product group is 0.27. With the median σ_U of 8.13, a UPC that has an error term that is 0.27 larger than another UPC, sells 6.85 times more. This heterogeneity in new product error terms matters for product group expenditure function, since large gains from new UPCs happen when firms produce exceptionally good products, products with high value of the error term, and thus limiting the error term to zero will underestimate change in product group expenditure function and the welfare index.

A.0.6 Change in Welfare Gains in Response to Change in σ_U and σ_F

How does changing σ^U and σ^F change the estimated welfare index? I illustrate how changing values of σ^U and σ^F changes welfare index by adding one new product to one firm, while keeping products of other firms constant. Adding more products has a similar intuition. First, I illustrate how the firm level expenditure function changes when adding a new product. Let's denote firm expenditure function before adding a new product by $P_{fg} = (\sum_u (\frac{P_{ut}}{\varphi_{ut}})^{1-\sigma^U})^{\frac{1}{1-\sigma^U}}$, and assume that the price divided by appeal of the new product is p. Addition of new product lowers new price index to $P_{fg} - \epsilon = (P_{fg}^{1-\sigma^U} + p^{1-\sigma^U})^{\frac{1}{1-\sigma^U}}$ (here, I assume that pricing other firm products doesn't change as a result of adding a new product. In the model, firms re-optimize pricing of all products when a new product is added. However, as shown in the paper, the effect of change in prices of other products is smaller relative to the direct effect of adding a new product. Thus for simplicity, I assume that firms don't change their prices when a new product is added). If we assume that the new product has high price divided by appeal, then $\epsilon << P_{fg}$ and:

$$P_{fg}^{1-\sigma^U} + p^{1-\sigma^U} = (P_{fg} - \epsilon)^{1-\sigma^U} \approx P_{fg}^{1-\sigma^U} - (1-\sigma^U)\epsilon P_{fg}^{-\sigma^U} \quad \Rightarrow \quad \epsilon \approx \frac{p^{1-\sigma^U} P_{fg}^{\sigma^U}}{\sigma^U - 1} \tag{A.5}$$

Now, let's understand how a change in firm expenditure function changes the product group expenditure function. Denote the product group expenditure function before adding a new product by $P_g = (\sum_{k\in\Omega} (\frac{P_{kg}}{\varphi_{kg}})^{1-\sigma^F})^{\frac{1}{1-\sigma^F}}$, and by $P_{g,new}$ after adding a new product, then change in group expenditure function is

$$1 - \frac{P_{g,new}}{P_g} = \frac{(P_{fg}/\varphi_{fg})^{1-\sigma_g^F}}{\sum_{k\in\Omega_{gt}} (P_{kg}/\varphi_{kg})^{1-\sigma_g^F}} \times \frac{1}{\sigma^U - 1} (\frac{P_{fg}}{p})^{1-\sigma^U}$$

Thus an increase in σ^U, decreases the change in product group expenditure function, and thus lowers welfare gains of adding a new product.

In order to understand the role of σ^F, remember that markups is

$$\mu_{fgt} = \frac{\sigma^F(1 - S_{fgt}) + S_{fgt}}{\sigma^F(1 - S_{fgt}) + S_{fgt} - 1}$$

Thus, an increase in σ^F decreases markups. So for a given product a decrease in markup charged by firms increases the welfare gains for consumers. Thus an increase in σ^F increases consumer welfare from creation of new products.

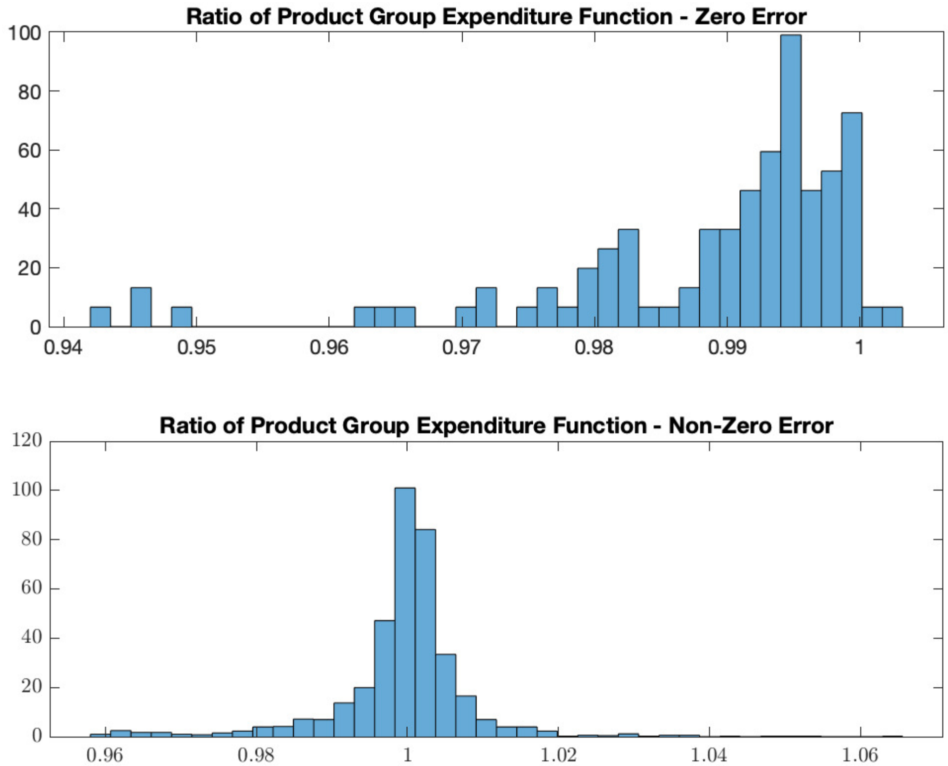

Figure 3.1: Histograms of Product Group Expenditure Function with Zero and Non-zero Error Term

Top panel plots the histogram of product group expenditure function of simulated AMCs (AMC is defined in equation 2.5, and can be interpreted as marginal "benefit" divided by marginal cost) from equation 2.5 divided by the welfare index using actual data (for all new products in 2007), assuming $\varepsilon_{ut} = 0$ in equation 2.5. Bottom panel plots the histogram of the same variable assuming ε_{ut} comes from the data. Equation 2.5 is estimated using data from 2004-2006.