Learning from Financial Markets and Misallocation

By

Jiaheng Yu

B.A. Economics and Mathematics
Tsinghua University, 2018

SUBMITTED TO THE SLOAN SCHOOL OF MANAGEMENT IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN MANAGEMENT RESEARCH

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2021

Signature of Author:____

Department of Management
May 1, 2021

Certified by: ____

Hui Chen
Professor of Finance
Thesis Supervisor

Accepted by: ____

Catherine Tucker
Sloan Distinguished Professor of Management
Professor, Marketing
Faculty Chair, MIT Sloan PhD Program

Learning from Financial Markets and Misallocation

by

Jiaheng Yu

Submitted to the Sloan School of Management
on May 1, 2021, in partial fulfillment of the
requirements for the degree of
Master of Science in Management Research

Abstract

I quantify how information frictions and learning from financial markets affect resource misallocation. I develop a dynamic model that features financial markets guiding managers in large investment decisions – mergers and acquisitions. Due to information frictions, mis-valuation of own firms and the potential gain from mergers and acquisitions prevent socially beneficial resource reallocation from happening. Compared to David et al. (2016), learning from the financial markets accumulates over time, and also occurs upon the announcement of the mergers and acquisitions. In the structural estimation, I target novel data moments including sensitivity of merger deal cancellation to announcement period returns to identify learning. The estimates suggest that a 50% decline in stock price informativeness locally would lead to 1.64% output loss for the US economy.

Thesis Supervisor: Hui Chen
Title: Professor of Finance

Acknowledgments

I am indebted to Hui Chen, Eben Lazarus, David Thesmar and Haoxiang Zhu for their continued guidance and support. I am also thankful to Debbie Lucas, Pari Sastry, Antoinette Schoar, Jingyi Su and Jian Sun for their comments and help. Part of the Fortran code used in this paper is adapted from materials in the Summer School in Structural Estimation hosted by University of Michigan.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Everyday a large amount of activities occur in secondary financial markets, in which securities are traded among investors, but without any capital flowing to firms. Do these activities affect real economic activity? A long tradition in economics, going back to Hayek (1945), would argue that prices are a useful source of information that facilitates efficient resource allocation. Indeed, financial market is a place where many speculators with different pieces of information meet to trade. Prices aggregate these diverse pieces of information and ultimately reflect an assessment of firm value. Real decision makers may learn new information from secondary market prices and use this information to guide their real decisions. Ultimately, the financial market has a real effect due to the transmission of information.

Although there are evidences on the learning channel (Luo, 2005; Kau et al., 2008; Durnev et al., 2004; Bakke and Whited, 2010), it remains unknown how large the channel's effect is on resource allocation and aggregate productivity of the economy. Existing study like David et al. (2016) estimates a model where firms rely on stock market and their own noisy information when making capital and labor decisions, and finds learning from financial markets contributes little to aggregate productivity. While dismissive on the impact of learning from financial markets, their approach identifies only learning during daily operational investments by investment-stock return correlation, and is silent on long-run, strategic firm investment decisions – like mergers and acquisitions.

I show that learning from financial market has substantial effect on aggregate productivity, once we consider such learning lowers the information friction firms face upon long-run decisions like mergers and acquisitions. The intuition is as follows. Information frictions lead to wrong perception of the long-run value of the firms. Managers over-pessimistic about their own firms and over-optimistic about merger opportunities conduct mergers even when the merger gain is negative – in other words, they redeploy their own capital too quickly. On the other hand, managers over-optimistic about their own firms and over-pessimistic about merger opportunities turn down mergers that bring positive gains – in other words, they dispose of their own capital too slowly.

The U.S. economy features an active market for corporate assets, where mergers and acquisitions are major vehicles for capital reallocation, accounting for huge flows of resources between firms. Expenditures on M&A from 1980-2009 averaged about 5% of GDP annually, reaching as high as 16% in the late 1990s, and about 44% of new business investment. The rate of capital reallocation via M&A accounts for an annual average of about two-thirds of total capital reallocation[1] among large U.S. firms, a figure that has grown to over 80%. Besides, firms rely on M&As to carry out a variety of long-run, strategic firm investment decisions that materially affect a firm's value. [2]

Mergers and acquisitions, however, seem to be accompanied by significant information frictions. The growth prospects, long-run productivity, and quality of match between management and capital, of the individual firms in an M&A talk and of the merged entity, are all hard-to-gauge objects. Unlike daily operational investments, firm managers are much less experienced with mergers and acquisitions, for which learning from other sources of information is more beneficial and should be more salient. For example, Roll (1986) argues that "the average individual bidder/manger has the opportunity to make only a few takeover offers during his career." And overconfident CEOs overpay for target companies and undertake value-destroying mergers (Malmendier and Tate, 2008).

In my model, learning from financial markets occurs in two channels. First, financial markets continuously produce signals on the long-run growth prospect of a firm. Observing such signals help the manager assess the continuation value of holding on to the existing asset, and hence facilitate socially efficient decisions on assets sales. Second, there are usually large stock price reactions upon announcements of mergers and acquisitions. Such reactions reveal the outside investor's perception of the merger gains, and are good sources of learning. For notable examples, BusinessWeek reports that Lucent stopped merger discussions with Alcatel because "investors clearly signaled their displeasure with it." After dropping the bid for the consulting arm of PricewaterhouseCoopers, Carly Fiorina, CEO of Hewlett-Packard, said to a group of analysts and institutional investors, "[A] number of you verbalized your concerns...and others simply voted with their positions in the stock... I realize you made some valid points."

The intention to quantify the effect of learning lend the project to a structural estimation. After all, parameters like the precision of market information that guides manager decisions are difficult to estimate in reduced form, and an economic model is needed to build counterfactuals. I estimate the model's parameters by applying the simulated method of moments (SMM) to data. Inspired by Luo (2005), who shows that stock market reaction to a merger and acquisition announcement predicts whether the companies later consummate the deal, I target a novel data moment – sensitivity of cancellation probability of mergers with respect to announcement period

[1] Capital reallocation takes 28% of total investment by publicly traded US firms. David (2017) estimates that gains from M&A contributes to 14% of total output.

[2] According to SDC data, the recorded purposes of M&As include: (1) acquire competitors technology/strategic assets, (2) expand presence in new geographical regions, or new/foreign markets (3) strengthen existing operations (4) offer new products and services,(5) eliminate duplicate services, (6) strengthen operations, (7) sell a loss-making/bankrupt operation, (8)concentrate on core businesses/assets, (9)pay outstanding debt, (10)raise cash through disposal, etc.

returns – to identify the manager's and market's information structure and learning. I also show empirical evidence that the slope is steeper in an environment with more informative stock prices.

The estimates suggest that when stock price informativeness improves by 50% from the current value, it will lead to a 0.57% increase in aggregate output. Meanwhile, when stock price informativeness deteriorates by 50%, will lead to a 1.64% decline in aggregate output. This means a further improvement in price informativeness of US stock market may not improve aggregate output much, but a deterioration would lead to substantial decline in aggregate output. A decomposition of the effect shows that, the main source of output gain comes from better informed merger decisions. When stock price informativeness improves by 50%, the probability of wrong merger decisions, that is, the fraction of mergers with true gains but not pursued, and mergers with negative gains but eventually conducted, decline by 1.22%. When stock price informativeness deteriorates by 50%, the probability of wrong merger decisions increases by 2.58%. Better stock price informativeness also modestly accelerates the sale of assets from low-productivity owners.

Relation to Literature. Although there is a large empirical and theoretical literature on efficiency of mergers and acquisitions [3], resource misallocations[4], and the feedback effect of stock market price discovery[5] , my contribution is to combine these strands of literature. This is one of the first attempts to quantify how learning from financial markets affect resource misallocation, using structural estimation of a dynamic learning model.

Closely related papers include Van Binsbergen and Opp (2019), who study the impact of informational inefficiencies of the stock market (i.e., the existence of alphas) on misallocation, featuring lumpy investment of firms. However they do not study learning. Feng (2018) and Li et al. (2020) build dynamic models where firms learn about their fundamental productivities as they age. However, firms only learn from realized output, not the financial markets. Sockin (2015) studies households and firms learning from stock prices, and shows that low price signals can distort expectations to be more pessimistic, leading to deeper recessions. But he does not study misallocation.

I bring together two types of learning documented by empirical finance papers. The first type is learning upon the arrival of a significant new project. This is documented by Luo (2005); Chen et al. (2007); Kau et al. (2008), who report that managers of acquiring firms appear to be influenced by their firms' stock price reactions at the announcement of proposed acquisitions—the more negative the stock price reaction, the greater the likelihood that a proposed transaction will be canceled.[6] The second type is learning during "daily" operations (e.g., investment at quarterly/annual frequency). This is documented by the relation between investment efficiency and price

[3]See Jain (1985); Hite et al. (1987); John and Ofek (1995); Maksimovic and Phillips (2001); Moeller et al. (2005)

[4]See Hsieh and Klenow (2009); Midrigan and Xu (2014); Bento and Restuccia (2017)

[5]See the survey article Bond et al. (2012). For international evidence, see Tan et al. (2015).

[6]The acquiror may later be acquired (Mitchell and Lehn, 1990), and the CEO lose her job (Lehn and Zhao, 2006).

discovery, e.g., (Durnev et al., 2004; Bakke and Whited, 2010; Edmans et al., 2017). Along the same line, most recently, Bennett et al. (2020) find that greater stock price informativeness (SPI) leads to higher firm-level productivity(TFP).

Throughout the paper, the notion of "learning" does not restrict to direct learning from stock prices. Even if managers do not learn from market prices, stock prices may affect manager decisions through corporate governance – managers care about stock prices which their wealth is tied to.[7] However, the corporate governance channel is ultimately similar to the learning channel, in that market prices end up having a real effect due to their informational role – the reason that managerial compensation depends on prices in the first place. And I assume away other frictions – for example, empire building, option value of delaying divestment, financial constraints, although they are important in history and well featured in literature. [8]

The rest of this paper is as follows. Chapter 2 introduces data and motivating empirical facts that support the two channels of learning featured in my model. Chapter 3 describes the model. Chapter 4 discusses estimation of the model and its implications. Chapter 5 concludes.

[7]Indeed, Datta et al. (2001) finds that, with the growth of options as a form of managerial compensation in the 1990s, making management more conscious of the impact of acquisitions on the stock price and more likely to make acquisitions that increase shareholder wealth.

[8]See Morck et al. (1990); Lambrecht and Myers (2007); Lang et al. (1995) for these frictions.

Chapter 2

Data and Empirical Facts

In this section, I present several motivating empirical evidence on the informational role of stock prices in guiding firm M&A decisions. The key findings are: (i) an M&A deal is more likely to be canceled upon negative announcement returns, and (ii) preceding voluntary sale of assets is a period with declining stock prices – suggesting stock prices facilitate the discovery of low productivity. These are also elements that feature critically in the model described in the next section.

2.1 Data

I adopt the sample of all domestic corporate transactions announced between 1990 and 2017 using data from the Thomson Reuters SDC Platinum database (SDC). SDC is a comprehensive source of data on US M&A, covering all corporate transactions involving at least 5% of the ownership of a company. Those deals with value above $1 million or undisclosed are covered before 1992. After 1992 all deals are covered. SDC covers both public and private transactions. Deal characteristics like transaction price, deal purpose, are documented in details. I also matched the SDC database to Compustat and CRSP to obtain the firm accounting data and the stock returns. I identify 14,652 M&A deals[1] that the acquiror was a public company with sufficient coverage in Compustat and CRSP, and the deal value was greater than $100 million. 4,541 unique acquirors are involved. Of these deals, 13,018 (88.8%) were completed and 1,634 (11.2%) were withdrawn.

To form a comparison with Chinese data, following Tan et al. (2015), I develop a sample of mergers and acquisitions announced between 2008 and 2013, of Chinese public firms listed in Shanghai and Shenzhen exchanges. The M&A deals data, together with the stock returns of the acquirors, come from the WIND financial database. I identify 525 M&A deals that the deal value was greater than $50 million. Of those deals, 431 (82.1%) were completed and 94 (17.9%) were withdrawn.

[1] The sample includes partial purchases less than 50% of the targets, hence a much generalized sample than those studied in literature (Luo, 2005; Kau et al., 2008; David, 2017).

2.2 Empirical Facts

I start by showing that the market reaction to a merger and acquisition announcement predicts whether the companies later complete the deal. Merging companies appear to extract information from the market reaction and later consider it in closing the deal. The extent of learning also varies across country and deal characteristics – learning seems more salient when the market has (relatively) more precise information.

First, the probability of an MA deal being canceled is higher when announcement period cumulative abnormal return is more negative. This is an already established result(Luo, 2005; Kau et al., 2008).[2] However, by extending to a much longer period and larger sample, I show this result holds very generally. I estimate the daily announcement period abnormal return as the firm's stock return less the return on the value-weighted CRSP index. For the sample of Chinese firms, the return on the Shanghai composite index is the benchmark used in the abnormal return calculation. I then estimate each bidder's cumulative abnormal return (CAR) by summing the daily abnormal returns over a three-days period after the deal's announcement.

Panel (a) of Figure 1 plots the probability of cancellation conditional on bidder's announcement period cumulative abnormal return being less than a sequence of cut-offs. Most striking is the downward-sloping curve for the US sample. The probability of cancellation is 12% if the bidder's announcement period cumulative abnormal return is negative. This probability rises to 16% when such return is less than -10%. A natural comparison to the US sample where the stock market is a benchmark in terms of informational efficiency, is an environment where stock market is less informative. China is a natural candidate for such a comparison, where with wide-spreading short sale constraints and the prevalence of retail investors, the stock market is notoriously uninformative. We see that for the Chinese sample, there is no clear monotone relationship between cancellation probability and announcement period cumulative abnormal return.

Results in Panel (a) lead to a natural data moment that identifies the firm learning effect of price informativeness – the sensitivity of deal cancellation with respect to announcement period returns. More precisely, the notion of price informativeness here measures the amount and precision of information contained in stock prices, in addition to the manager's own information, that guides real corporate decisions. We see that in the US, there is a clear downward-sloping curve of cancellation probability, but not in China. The probability upon any negative announcement return – or the level of the curve, doesn't identify the learning effect. Many other factors can affect the cancellation of an M&A deal in an economy or an informational environment, e.g. anti-trust regulations, ease of financing in M&A deals, enforcement of contracts and so on. However, the slope of the curve largely difference out these other factors and isolate the aggregate effect of learning in an economy.

To provide further evidence on the learning channel, and the validity of the sensitivity of M&A deal cancellation with respect to stock returns as a data moment,

[2] Even if there is no learning, deals that are perceived better by both the market and the merging companies at the announcement may have higher return and a better chance of consummation. Luo (2005) controls for this and shows the existence of learning effect.

I exploit variations in the cross-section of M&A deals, to show that learning seems more salient when the market has (relatively) more precise information. For example, smaller acquirors have less expertise and fewer resources to process public information on themselves. Smaller companies have less managerial talent. Anecdotally, smaller bidders can afford less for in-house M&A analysis or outside investment banking services. Their managers are likely trained with more knowledge on production than on finance. Thus, smaller bidders tend to find the market more informative than do large bidders. Panel (b) of Figure 1 plots the probability of cancellation conditional on announcement period returns, separately for small acquirors (market capitalization below 25% percentile) and large acquirors (market capitalization above 75% percentile). We see that for small acquirors, clearly downward-sloping is the curve of cancellation probability, but the downward trend is less clear and robust for large acquirors. This suggests that learning effect is indeed stronger for small acquirors.

Also, the opacity of high-tech deals makes the market's opinion less relevant. Without important raw valuation information, investors' opinions become less informative. Additionally, high-tech deals are more likely pioneers of their types. There are fewer similar deals in the past for investors to use as comparisons for deal valuation. As a result, companies are less likely to learn from the stock prices in high-tech deals than in non-high-tech deals. Panel (c) of Figure 1 plots the probability of cancellation conditional on announcement period returns, separately for high-tech deals and non-high-tech deals. [3] We see that for non-high-tech deals, the slope in the curve of cancellation probability is steeper than for high-tech deals, suggesting that learning effect is indeed stronger for non-high-tech deals.

Does the announcement period stock return provide useful information on the future growth prospect of the merged firm? To provide evidence that it does, [4] Panel (d) of Figure 1 plots the cumulative abnormal returns after the mergers became effective, of a group of acquirors with negative announcement period returns. When the announcement period cumulative abnormal return (CAR) is negative, the merged firm experiences a -5% cumulative abnormal return 18 months after the merger becomes effective. When the announcement period CAR is below -5%, the merged firm experiences a -10% CAR 18 months after the merger becomes effective. When the announcement period CAR is below -10%, the merged firm experiences a below -14% CAR 18 months after the merger becomes effective. It appears that negative announcement returns forebode the dismal future of the merger of two firms – the merger is a wrong match and the resulting negative synergies reveal themselves after the deal become effective, which is eventually reflected in the stock price of the merged firm. The result is in accordance with Moeller et al. (2005). They show that acquiring firms who experience significant drop in capitalization over the 3 days surrounding ac-

[3] If both the acquiror and the target are high-tech companies labeled by SDC, the deal is a high-tech deal. The labeled high-tech companies concentrate in the computer and IT industries, which in today's view are regular industries rather than advanced high-tech industries anymore. Here I restrict the sample to deals between 1990-2003, where the classification of high-tech is more accurate. Extending the sample period leads to similar, but less contrasting results.

[4] The result however does not imply information in announcement period stock return is not already owned by the manager.

quisition announcements perform extremely poorly afterwards. An equally weighted portfolio of firms that with significant drop in capitalization upon deal announcement will have worse than -40% cumulative abnormal returns in 58 months afterwards.

Now, I show that stock prices not only reveal information about M&A deal quality, but may also facilitate the discovery about own firm's productivity in daily operations over time. An influential theory about corporate asset sales – the efficient deployment hypothesis most explicitly advanced by Hite et al. (1987) – suggests that firms only manage assets for which they have a comparative advantage and sell assets as soon as another firm can manage them more efficiently. Going beyond this classical theory, if firm manager does not however have perfect information of his comparative advantage in operating the assets, stock price while incorporating outside investors' opinions, may provide useful information on that and guide efficient asset sales.

Figure 2 plots the cumulative abnormal returns of a sample of firms before they sell corporate assets. The sample includes voluntary asset sales between 1990 and 2003 in the SDC database, including mergers, acquisitions, and partial-firm asset sales. To exclude the firms who sell assets because they are in financial distress, asset sales whose documented purposes are to pay outstanding debt, raise cash through disposal, or sell a loss-making/bankrupt operation, are excluded from the sample. We see that the cumulative abnormal return is around -8% up to 2 months before the announcement of asset sales. The result is consistent with empirical findings in the literature. Jain (1985) finds that sell-off announcement are preceded by a period of negative returns for the sellers, and are greeted positively by the market. Lang et al. (1995) find that asset sales follow poor firm-level performance. John and Ofek (1995) find that the remaining assets of the firm improve in performance after asset sales that subsequently leave the firm more focused.

[Insert Figure 1 near here]

[Insert Figure 2 near here]

When individual firm managers have imperfect information about their capital's or investment project's productivity, they make wrong investment decisions and lead to resource misallocation. I conclude this section by a summary of empirical facts in two papers, showing that learning over the life cycle of the firm alleviates such resource misallocation.

Feng (2018) finds a consistent negative relation between marginal product of capital (MPK) dispersion, which has been interpreted as a measure of capital misallocation in the literature, and firm age. The paper uses firm-level panel data from China, Columbia, and Chile. In particular, for the sample of Chinese firms, the standard deviations of log MPK, decreases substantially by 13% from age 0 to age 5. Li et al. (2020) find that among US public listed firms, the dispersion of the marginal product of capital (MPK), monotonically declines with firms' capital age. Capital age is calculated as a weighted average of the age of each capital vintage. The intuition of both papers is that young firms have less precise information about their firm-specific productivity to facilitate resource allocations, hence have more capital misallocations.

Chapter 3

The Model

I study a discrete time, infinite-horizon economy. A cross-section of firms of measure one operate technologies with decreasing returns to scale and capital adjustment costs. Besides the capital adjustment cost, information friction is the only friction – imperfect information about their own fundamental productivity as well as the productivity of merger opportunities. I deliberately keep the household side and labor market of the economy simple because they play a limited role in the analysis. Like all models, this model presents a simplified view of the world. The simplifications render me clearer predictions from the model and make it computationally feasible to identify parameter values from the data.

3.1 Firm Technology

Each individual firm uses capital k_t to produce output, with decreasing returns to scale. Each firm has productivity a_t at time t, which follows an AR(1) process, as in most literature, for example, David et al. (2016). The output of a firm in each period is

$$y_t = exp(a_t)k_t^{\phi} \tag{3.1}$$

The long-run mean of the AR(1) process is a. Contrary to literature, I assume the long-run mean a is unknown and is being learned over time. a represents the fundamental productivity of the firm, and can also be interpreted as the quality of the match between the firm and its assets (Jovanovic, 1979), that determines the forward-looking valuation of the firm. A firm's fundamental productivity θ is constant over time, and is drawn from a normal distribution $N(m, \sigma_0^2)$ when the firm starts.

$$a_t = a_{t-1} + \rho(\theta - a_{t-1}) + \sigma_a \epsilon_{1t} \tag{3.2}$$

The shock $\sigma_a \epsilon_{1t}$ is independently and normally distributed with mean zero and variance σ_a^2. Persistence parameter $0 < \rho < 1$ is the same for all firms and commonly known. Under the specification, realized productivity and long-run mean both have long-lasting effects on outputs, hence may affect firm value and firm investment decisions.

Every period capital depreciates at a rate of δ. Firm makes an investment i_t. i_t can be either positive or negative. Disinvestment is costly because of the adjustment cost function, but it is not completely irreversible. The law of motion of capital is

$$k_{t+1} = (1-\delta)k_t + i_t \tag{3.3}$$

Every period the profit of the firm is

$$\pi_t = exp(a_t)k_t^{\phi} - i_t^{\alpha} \tag{3.4}$$

where the curvature parameter $\alpha > 1$ represents the cost of capital adjustment. Throughout the estimation I set $\alpha = 2$, corresponding to the quadratic adjustment cost widely specified in literature. Profit π_t can be either positive or negative. If positive, it represents distributions of internal cash flows to shareholders, and if negative, it represents infusions of cash from shareholders into the firm. Thus, in this model internal funds and external financing from shareholders are equally costly, and financing is trivial. For simplicity, the shareholder (household) side is not modeled.

3.2 The Merger Market

The structure of the merger market builds on the classical search model of Shimer and Smith (2000). Firms search with an exogenous intensity. Since empirically acquirers often themselves become targets (David, 2017), I assume firms search simultaneously on both sides of the market, i.e., as targets and acquirers.

The economy starts with a continuum of firms of mass one. Each firm faces a Poisson rate λ of arrival of an acquirer and faces the same rate of arrival of a target. Mergers reduce the mass of firms, and I assume exogenous firm entry and exit at a rate such that a steady state mass of firms, and a stationary distribution of firm characteristics maintains.

Upon entering a meeting, the firms discuss whether to carry out a merger and bargain the price of the deal. Firms, were to be merged, draws a new fundamental productivity level θ_M from the distribution $N(m, \sigma^2)$. The other characteristics of the merged firm follow

$$k_M = k_A + k_T \tag{3.5}$$

$$exp(a_{Mt})k_M^{\phi} = exp(a_{At})k_A^{\phi} + exp(a_{Tt})k_T^{\phi} \tag{3.6}$$

$$\tau_{Mt} = 0 \tag{3.7}$$

where the capital of the merged firm k_M equals the sum of of the acquirer's and the target's capital. The initial realized productivity of the merged firm a_{Mt} makes sure that the initial profit of the merged firm is the same as the total profits of the acquirer and the target.[1] The age of the merged firm τ_{Mt} is 0, meaning accumulated knowledge

[1] This assumption avoids the integration over a_{Mt} in estimation, and simplifies numerical analysis.

about the acquiror's and target's productivities becomes obsolete once they merge.

The combined gain from merger, Φ, is the value of the merged entity less the values of the two pre-merger firms.

$$\Phi(\theta_M, \theta_A, \theta_T) = V(\theta_M) - V(\theta_A) - V(\theta_T) \tag{3.8}$$

where firm values depend on their fundamental productivities. $V(\theta_A)$ is the value of the acquirer, $V(\theta_T)$ the value of the target, and $V(\theta_M)$ is the value of the merged entity. For now, I notationally suppress the dependence of the firm value function on the other state variables.

Because of information frictions, the managers only have imperfect estimates of the fundamental productivities, $\hat{\theta}_M$, $\hat{\theta}_A$ and $\hat{\theta}_T$. When the managers meet, they share their estimates with each other and rely on them to collectively assess the deal. Again suppressing the dependence on other state variables, the perceived combined gain from merger by managers is

$$\Phi(\hat{\theta}_M, \hat{\theta}_A, \hat{\theta}_T) = V(\hat{\theta}_M) - V(\hat{\theta}_A) - V(\hat{\theta}_T) \tag{3.9}$$

If the combined gain from merger is positive, managers decide to carry out the merger and announce the deal together with the deal price. Now we see the role of information frictions in resource misallocation. Information frictions lead to wrong perception of the value of the firms. Managers over-pessimistic about their own firms' productivities and over-optimistic about the merged firm's productivity conduct mergers even when the merger gain is in fact negative – in other words, they redeploy their own capital too quickly. On the other hand, managers over-optimistic about their own firms' productivities and over-pessimistic about the merged firm 's productivity turn down merger opportunities that bring positive gains – in other words, they dispose of their own capital too slowly.

How is the deal price determined? If the merger gain is perceived positive – $\Phi(\hat{\theta}_M, \hat{\theta}_A, \hat{\theta}_T) > 0$, firm managers engage in a Nash Bargaining that splits the surplus. Denote the bargaining power of the acquiror as η and $1 - \eta$ that of the target. The deal price satisfies

$$D(\hat{\theta}_M, \hat{\theta}_A, \hat{\theta}_T) = V(\hat{\theta}_T) + (1 - \eta)\Phi(\hat{\theta}_M, \hat{\theta}_A, \hat{\theta}_T) \tag{3.10}$$

Separately for the acquiror and the target, the perceived gains from a merger are

$$\Phi_A(\hat{\theta}_M, \hat{\theta}_A, \hat{\theta}_T) = \eta\Phi(\hat{\theta}_M, \hat{\theta}_A, \hat{\theta}_T); \quad \Phi_T(\hat{\theta}_M, \hat{\theta}_A, \hat{\theta}_T) = (1 - \eta)\Phi(\hat{\theta}_M, \hat{\theta}_A, \hat{\theta}_T) \tag{3.11}$$

3.3 Learning

A. Learning during Daily Operations

It's also reasonable to assume two merging firms remain their stand-alone profits right after the merge.

Now I solve the manager's learning problem, which is a Kalman-Bucy filtering problem. Let τ_t denotes the time that the firm has been in the economy (i.e. the age of the firm). The manager starts with the common prior that $\theta \sim N(m, \sigma_0^2)$. At the beginning of every period, [2] he observes the realized productivity a_t. The history of a_t provides information about fundamental productivity θ. The history of a_t is equivalent to the history of persistence-adjusted productivity

$$\frac{1}{\rho}(a_t - a_{t-1}) + a_{t-1} = \theta + \frac{1}{\rho}\sigma_a \epsilon_{1t} \tag{3.12}$$

I assume the outside investors also observe the history of the realized productivity every period. They also collectively receive a private signal every period, and the stock price however reflects that private signal. Here I abstract from spelling out a micro-structure model of the determination of stock prices. One interpretation of this assumption is that stock price should reveal the fundamental productivity if it is informationally efficient, but the presence of noise traders makes it deviate from the firm value implied by true fundamental productivity. For the manager, observing the stock price is informationally equivalent to observing a signal

$$\hat{s}_t = \theta + \sigma_z \epsilon_{2t} \tag{3.13}$$

Note that ϵ_{2t} is independent from ϵ_{1t}, meaning that signal from stock price is orthogonal to the manager's own signal. David et al. (2016) provides a micro-structure model where noise traders and imperfectly informed investors trade using limit orders, and the stock price in a rational expectation equilibrium is shown to be informationally equivalent as such a signal.

Standard results on Kalman-Bucy filters apply, and according to Chapter 6 of Oksendal (2003) and Chapter 10 of Lipster and Shiryaev (2001), the manager's posterior distribution of fundamental productivity is Gaussian. The posterior distribution is hence summarized by the conditional mean $\hat{\theta}_t = \mathbb{E}[\theta | \{a_s, \hat{s}_s\}_{s=1}^t]$, and conditional variance $\hat{v}(t) = \mathbb{E}[(\theta - \hat{\theta}_t)^2 | \{a_s, \hat{s}_s\}_{s=1}^t]$.

I use the notation $\kappa_a = \sigma_a^2/(\phi^2\sigma_0^2)$, $\kappa_z = \sigma_z^2/\sigma_0^2$ to denote the relative precisions of signals to the prior. Suppose the posterior distribution is $N(\hat{\theta}_{t-1}, \hat{v}_{t-1})$, and the surprises in the signals equal

$$\delta_{at} = \frac{1}{\rho}(a_t - a_{t-1}) + a_{t-1} - \hat{\theta}_{t-1} \tag{3.14}$$

$$\delta_{zt} = \hat{s}_t - \hat{\theta}_{t-1} \tag{3.15}$$

Standard results on Bayesian learning imply that $\hat{v}(\tau)$, the posterior variance of fundamental productivity a after τ periods of learning, decays monotonically and

[2] The manager observes realized productivity before making investment decision, as assumed in Midrigan and Xu (2014). This simplifies the estimation.

deterministically with firm age according to

$$\hat{v}(\tau) = \sigma_0^2[1 + \tau(\kappa_a^{-1} + \kappa_z^{-1})]^{-1} \tag{3.16}$$

The posterior mean evolves according to

$$\hat{\theta}_\tau = \hat{\theta}_{\tau-1} + \theta_a(\tau)\delta_{a\tau} + \theta_z(\tau)\delta_{z\tau} \tag{3.17}$$

$$\theta_a(\tau) = \kappa_a^{-1}[1 + \tau(\kappa_a^{-1} + \kappa_z^{-1})]^{-1} \tag{3.18}$$

$$\theta_z(\tau) = \kappa_z^{-1}[1 + \tau(\kappa_a^{-1} + \kappa_z^{-1})]^{-1} \tag{3.19}$$

B. Learning upon Mergers

The acquiror and the target, were to be merged, would draw a new fundamental productivity level θ_M from the distribution $N(m, \sigma_0^2)$. The managers of participating firms receive a private signal about the θ_M,

$$s_M = \theta_M + \sigma_a e_t \tag{3.20}$$

The shock $\sigma_a e_t$ is independently and normally distributed with mean zero and variance σ_a^2. I assume the precision of this signal is the same as that implied by the realized productivity. This implies the manager adopts the same information technology in daily operations and for mergers. Clearly a courageous assumption, but this greatly simplifies the model and estimation.

When the firms announce the merger, the outside investors also receive a private signal about the fundamental productivity of the merged entity.

$$\hat{s}_M = \theta_M + \sigma_z z_t \tag{3.21}$$

where the shock $\sigma_z z_t$ is independently and normally distributed with mean zero and variance σ_z^2. We see another critical assumption here, that is the precision of this signal is the same as the signal the market receives about existing firms' fundamental productivity during daily operations. On the one hand this is a simplifying assumption, although it is not entirely unreasonable to assume the distribution of traders in the stock market is stationary, hence generating a stationary signal. And relying on this assumption, I can identify the precision of market signals during daily operations using the abundant mergers and acquisitions data.

On the other hand, the key channel my model would highlight is that, the extent that stock price informativeness affects resource misallocation, depends how strongly stock price can inform the fundamental productivity of stand-alone firms and the merged entity. What matters is the relative precision of market signal to manager's signal both during daily operations and upon mergers, rather than difference in precisions of either the manager's or the market's signal between daily operation and mergers. Hence I assume the manager's signal precision is constant between daily operation and merger, and the same for market's signal. And in counterfactural analysis, I focus on the effect of changes in σ_z while keeping σ_a fixed.

One can alternatively model the precision of signals in daily operations to be different from that in mergers. However, the additional parameters may be difficult to identify. I intentionally make the model simple, and leave the subtlety of the whole information structure to future work.

Observing the acquiror's stock price reaction after announcement is equivalent to observing the market signal $\hat{s}_M$. The managers then combine the market signal with their private signal, and re-evaluate the merger gain. The managers' updated belief about θ_M, by Bayes rule, is

$$\hat{\theta}'_M = \frac{\mathbb{V}}{\sigma^2}m + \frac{\mathbb{V}}{\sigma_a^2}s_M + \frac{\mathbb{V}}{\sigma_z^2}\hat{s}_M \tag{3.22}$$

where $\mathbb{V}$ is the posterior variance given by

$$\mathbb{V} = \left(\frac{1}{\sigma_0^2} + \frac{1}{\sigma_a^2} + \frac{1}{\sigma_z^2}\right)^{-1} \tag{3.23}$$

Managers renegotiate the deal price after observing the stock price reaction after the announcement of the merger. If under the managers' updated belief, the gain from the merger is still positive, the same Nash Bargaining determines the updated deal price, and the deal is consummated at the new price. Otherwise, the deal is canceled. That is, if $\Phi(\hat{\theta}'_M, \hat{\theta}_A, \hat{\theta}_T) > 0$, then the deal is consummated. Otherwise, the deal is canceled.

3.4 Firm Objective

Managers maximize the value of their firms by undertaking investment decision i_t during daily operations and carrying out mergers with positive gains, under their imperfect information. The subjective value of a stand-alone firm V_t may depend on state variables $(\hat{\theta}_t, \hat{v}_t, k_t, a_t)$, which summarize the manager's information set. $\hat{\theta}_t$ is the manager's conditional expectation of the fundamental productivity, $\hat{v}_t$ is the conditional variance of the fundamental productivity, k_t is capital, and a_t is realized productivity. As shown in (3.16), conditional variance $\hat{v}_t$ is only a function of firm age τ_t, the state variables are effectively $(\hat{\theta}_t, \tau_t, k_t, a_t)$.

For the value function, it is simpler to use the recursive formulation.

$$\begin{aligned} V(\hat{\theta}_t, \tau_t, k_t, a_t) = \max_i \Big\{ &\pi_t + \beta\lambda\mathbb{E}_t^*\big[\eta \max\{\Phi(\hat{\theta}'_M, \hat{\theta}_{t+1}, \hat{\theta}_T, \Gamma_{t+1}), 0\} \\ &+ V(\hat{\theta}_{t+1}, \tau_{t+1}, k_{t+1}, a_{t+1})\big] \\ &+ \beta\lambda\mathbb{E}_t^*\big[(1-\eta)\max\{\Phi(\hat{\theta}'_M, \hat{\theta}_A, \hat{\theta}_{t+1}, \Gamma_{t+1}), 0\} \\ &+ V(\hat{\theta}_{t+1}, \tau_{t+1}, k_{t+1}, a_{t+1})\big] \\ &+ \beta(1-2\lambda)\mathbb{E}_t^* V(\hat{\theta}_{t+1}, \tau_{t+1}, k_{t+1}, a_{t+1}) \Big\} \end{aligned} \tag{3.24}$$

The stand-alone firm value consists of several parts. π_t is the flow profit, β is the effective discount rate of the firm, $\lambda\mathbb{E}_t^*\eta\max\{\Phi(\hat{\theta}'_M, \hat{\theta}_{t+1}, \hat{\theta}_T), 0\}$ is the expected flow gain from acquiring a firm and $\lambda \mathbb{E}_t^*(1-\eta)\max\{\Phi(\hat{\theta}'_M, \hat{\theta}_A, \hat{\theta}_{t+1}), 0\}$ the expected flow gain from being acquired. The last term is the continuation value of the stand-alone firm, given optimal policies will be chosen in the future. $\mathbb{E}_t^*$ denotes the expectation conditional on the manager's information set. The continuation value depends on next period's posterior mean of fundamental productivity, implying that managers' preference discount subjective valuation in a time-consistent manner, consistent as in Jovanovic (1982).

Boundary conditions and smooth pasting conditions of this model are simple. When two stand-alone firms merge, the merged entity appears as a new firm, and the stand-alone firms disappear from the economy. When a firm disappears, its terminating value is still the firm value as if the stand-alone firm keeps operating.

The Bellman equation (3.24) of value function can be simplified given symmetry of the merger gains. The combined merger gains when a firm arrives at a meeting as an acquiror and as a target are respectively,

$$\begin{aligned}\Phi(\hat{\theta}'_M, \hat{\theta}_{t+1}, \hat{\theta}_T, \Gamma_{t+1}) =& V(\hat{\theta}'_M, 0, k_{t+1}+k_T, a_M) - V(\hat{\theta}_{t+1}, \tau_{t+1}, k_{t+1}, a_{t+1}) \\ & - V(\hat{\theta}_T, \tau_T, k_T, a_T) \end{aligned} \tag{3.25}$$

$$\begin{aligned}\Phi(\hat{\theta}'_M, \hat{\theta}_A, \hat{\theta}_{t+1}, \Gamma_{t+1}) =& V(\hat{\theta}'_M, 0, k_{t+1}+k_A, a_M) - V(\hat{\theta}_{t+1}, \tau_{t+1}, k_{t+1}, a_{t+1}) \\ & - V(\hat{\theta}_A, \tau_A, k_A, a_A) \end{aligned} \tag{3.26}$$

Since the characteristics of the counterpart firm in the meeting, either being the target or the acquiror, is drawn from the same distribution. Hence the expected combined gain from a merger perceived by a manager is the same, either being a target or an acquiror, that is, $\mathbb{E}_t^* \max\{\Phi(\hat{\theta}'_M, \hat{\theta}_{t+1}, \hat{\theta}_T, \Gamma_{t+1}), 0\}$
$= \mathbb{E}_t^* \max\{\Phi(\hat{\theta}'_M, \hat{\theta}_A, \hat{\theta}_{t+1}, \Gamma_{t+1}), 0\}$. The Bellman equation simplifies to

$$\begin{aligned} V(\hat{\theta}_t, \tau_t, k_t, a_t) = \max_i \Big\{ & exp(a_t) k_t^{\phi} - i_t^{\alpha} + \beta\lambda \mathbb{E}_t^* \max\{\Phi(\hat{\theta}'_M, \hat{\theta}_{t+1}, \hat{\theta}_T, \Gamma_{t+1}), 0\} \\ & + \beta \mathbb{E}_t^* V(\hat{\theta}_{t+1}, \tau_{t+1}, k_{t+1}, a_{t+1}) \Big\} \end{aligned} \tag{3.27}$$

The transition probabilities $(\hat{\theta}_{t+1}, a_{t+1} | \hat{\theta}_t, a_t)$ that are useful in calculating manager's expectations are described in the following

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \theta_a(\tau_{t+1})\delta_a + \theta_e(\tau_{t+1})\delta_e + \theta_z(\tau_{t+1})\delta_z \tag{3.28}$$

$$\begin{pmatrix}\delta_a \\ \delta_e \\ \delta_z\end{pmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \sigma_a^2/\rho^2 + \hat{v}(\tau) & 0 & 0 \\ 0 & \sigma_e^2 + \hat{v}(\tau) & 0 \\ 0 & 0 & \sigma_z^2 + \hat{v}(\tau) \end{bmatrix}\right) \tag{3.29}$$

$$a_{t+1} \sim \mathcal{N}\Big((1-\rho)a_t + \rho\hat{\theta}_t, \sigma_a^2 + \rho^2 \hat{v}(\tau)\Big) \tag{3.30}$$

$$Cov(\hat{\theta}_{t+1}, a_{t+1}) = \theta_a(\tau_{t+1}) Var(a_{t+1})/\rho \tag{3.31}$$

3.5 Stock Prices

In the beginning of every period t, outside investors receive the signal $\hat{s}_t$ about fundamental productivity of the firm. That is, the outside investors' posterior distribution of the fundamental productivity is $N(\hat{s}_t, \sigma_z^2)$. I assume the stock price of a firm equals its valuation perceived by the outside investors. More precisely, I interpret the stock price as the value of the firm if outside investors replace the manager, access the same information technology of the manager, and operate the firm optimally. Under this assumption, I can use the same value function derived before to model stock price. This is a strong assumption, but intended to simplify the analysis. The stock price at the beginning of period t is

$$S_t = V(\hat{s}_t, \sigma_z^2, k_t, a_t) \tag{3.32}$$

As will be shown by the numerical solution of Bellman equation, the value function is monotone in the first argument. Hence for the manager, observing stock price is equivalent to observing $\hat{s}_t$. Denoting $\hat{v}(\tau_z) = \sigma_z^2$, we can replace posterior variance σ_z^2 in the state variable to τ_z.

Before the announcement of a merger deal, the stock price of the acquiror is $S_A^- = V(\hat{s}_A, \tau_z, k, a)$, and the stock price of the target is $S_T^- = V(\hat{s}_T, \tau_z, k, a)$. Right after the announcement, the stock price of the target is the deal price $S_T^+ = D = V(\hat{\theta}_T) + (1-\eta)\Phi(\hat{\theta}_M, \hat{\theta}_A, \hat{\theta}_T)$. And the stock price of the acquiror is $S_A^+ = V(\hat{s}_M) - D$. Since D is announced by the manager, observing the stock price S_A^+ is equivalent to observing the market's signal $\hat{s}_M$ about the fundamental productivity of the merged entity.

3.6 Equilibrium

Most structural estimation literature on resource misallocation or mergers, solves for a stationary industry equilibrium suggested by Hopenhayn (1992). They usually assume entry, exit and interactions among firms change the distribution of firm characteristics in the economy. Firm's optimal decisions both depend on the distribution of firm characteristics, and shape the distribution. They iterate to find a stationary distribution and firm decision rule that together is a fixed point.

Computing such an equilibrium involves iterations within iterations when solving value function and decision rules, and again iterations over parameter space for SMM estimation. With four state variables like in my model, the task soon becomes computationally infeasible. On the other hand, a stationary distribution takes a long time of evolving to achieve – when the majority of firms in the economy are old, they would have precise enough knowledge about productivity such that learning is unnecessary. Given my focus on the post-war history of US enterprises, the majority of firms do not appear long-lived enough such that this is the case.

Thus, I adopt an equilibrium definition that is easier to compute and offers a better approximation to the economy. The equilibrium contains (1) A initial joint cross-section distribution $\Omega(\hat{\theta}_{it}, \tau_{it}, k_{it}, a_{it})$. (2) Manager's investment decisions, optimal with respect to the cross-section distribution $\Omega(\hat{\theta}_{it}, \tau_{it}, k_{it}, a_{it})$ as if the distribution is stationary. (3) Firms flow out through being acquired. Firms flow in as the continuing entity from merger, and through entry by new firms with characteristics $\tau = 0, k = 0$ and $\hat{\theta}_0 \sim N(m, \sigma_0^2)$ and $a_0 \sim N(m, \rho^2\sigma_0^2 + \sigma_a^2)$ at a rate so that the mass of firms remain stationary.

3.7 Numerical Solution of Bellman Equation

I numerically solve the Bellman equation to find the manager's optimal investment decision. In the numerical solution, I obtain an approximate solution for value function by discretizing the state space and iterating on the Bellman equation (See Appendix C for more details). Since I have 4 state variables, to avoid the curse of dimensionality, I use sparse grid that reduces the total number of grid points value function is to be evaluated, following Brumm and Scheidegger (2017). (See Appendix D for details). When calculating the manager's expectations in the value function, I utilize Gauss-Hermite quadrature methods. (See Appendix E for details).

Figure 3 shows the value function and investment decision at parameter values close to SMM estimates. We see that value function is mostly affected by fundamental productivity, less so by firm age and capital. The realized productivity has limited impact on firm value. Optimal investment is decreasing in existing capital stock.

[Insert Figure 3 near here]

Chapter 4

Estimation and Quantitative Implications

4.1 Parameterization

Table 1 below summarizes the parameterization of the model. First, some parameters are directly assigned, based on existing practices in the literature. The time period for the estimation is one year. Real output in US manufacturing grew at an annual real rate of 3 percent in the years I study. I consequently set $r = 1.03$ and choose a value of β equal to $0.92r$, following Midrigan and Xu (2014). Capital depreciates at a rate $\delta = 0.06$, following Midrigan and Xu (2014). Curvature of production function $\phi = 0.63$ following Hennessy and Whited (2007). I assume quadratic adjustment cost of capital, that is, $\alpha = 2$. I set the bargaining power of acquirors in mergers $\eta = 0.51$ following David (2017).

Three parameter can be directly identified by the data, which are the persistence of productivity ρ, prior mean and variance of fundamental productivity (m, s^2). Realized productivity can be recovered by by $\log(output) - \phi\log(capital)$, according to the firm production function. Since I abstract from intermediate inputs in the model, the measure of output that most closely relates to that in my model is value added. However, I currently do not find data on payments to intermediate inputs. Instead, following Chen and Song (2013) and Li et al. (2020) in their calculation of MPK, I use operating income before depreciation (oibdpq) to measure firm output, and the one-year-lag of net property, plant, and equipment (ppentq) as firm capital. For robustness, I also tried replacing the operating income before depreciation (oibdpq) with sales (saleq).[1] All the quantities are expressed in 2012 constant dollars using the implicit price deflator for non-residential fixed investment.

Then following David et al. (2016), I conduct AR(1) regression of realized productivities. I first regress realized productivity against its lag and firm fixed effect, to get an estimate for $\rho = 0.48$. Then for each firm, the mean of $a_{t+1} - (1-\rho)a_t$ divided by

[1] Using sales (saleq) to proxy a firm's output alleviates any missing data concerns, given that the coverage of sales (saleq) is higher than that of operating income before depreciation (oibdpq) in Compustat.

ρ will recover the fundamental productivity a. It turns out that the distribution of a is well approximated by a normal distribution, with mean $m = 0.74$ and standard deviation $\sigma_0 = 1.27$. Pooling the error terms $\sigma_a \epsilon_{it}$ of all firms together and take the variance as estimate for σ_a, I have $\sigma_a = 0.53$.

The rest 2 parameters, (λ, σ_z) , are estimated by targeting data moments. λ is identified by targeting the frequency of M&A announcements. In my sample, I find that about 3.7% of Compustat firms are acquired in announcements annually over the sample period. σ_z is identified by targeting the sensitivity of deal cancellation with respect to announcement period return. As shown in Figure 1, the cancellation probability when announcement period return is less than -8%, is 4% higher than when announcement period return is less than 0%. To show that data moments are good at identifying the parameters. Figure 4 plots parameter values against the simulated moments. We see there is a quite sharp relationship between parameter values and moments, suggesting that the data moments can identify the parameters.

[Insert Table 1 near here]

[Insert Figure 4 near here]

4.2 Estimation Method

I estimate the parameters $\theta = \{\lambda, \sigma_z\}$ using SMM.[2] SMM estiamtes parameter values by matching certain data moments and model-implied moments as closely as possible. The SMM estimator $\hat{\theta}$ is

$$\hat{\theta} = \arg\min_{\theta} \left(\hat{M} - \frac{1}{S} \sum_{s=1}^{S} \hat{m}^s(\theta) \right)' W \left(\hat{M} - \frac{1}{S} \sum_{s=1}^{S} \hat{m}^s(\theta) \right) \tag{4.1}$$

where $\hat{M}$ is a vector of moments estimated from the empirical data, $\hat{m}^s(\theta)$ is the corresponding vector of moments estiamted from the s-th sample simulated using parameters θ, and W denotes my choice of weighting matrix which I discuss in more details below. Michaelides and Ng (2000) find that using a simulated sample 10 times as large as the empirical sample generates good small-sample performance. I use S = 10 simulated samples.

For each simulated sample, I simulate at annual frequency. Each simulation has 1,500 firms and 60 years. I firstly draw each firm's fundamental productivity a from the prior distribution. Then I generate realized productivity a_t and stock market's signals using the fundamental productivities and simulated shocks, and I update the manager's beliefs according to the learning rule in Equation (3.18). Managers make investment decision according to the optimal rule from the Bellman equation. And they make merger announcement based on their beliefs upon the meeting and adjust the decision after observing market reactions. Once a target is acquired, a new firm enters the economy with randomly drawn fundamental productivity. I construct

[2] See, for example, Hennessy and Whited (2007); Midrigan and Xu (2014).

moments $\hat{m}$ the exactly same way in simulate data as in real data. Following Hennessy and Whited (2007), I use simulated annealing optimization algorithm to find global minimum of (4.1).

I use the optimal weighting matrix

$$W = (\hat{\Sigma} + \frac{1}{S}\Omega(\hat{\theta}))^{-1} \tag{4.2}$$

where $\hat{\Sigma}$ is an estimate of the variance-covariance matrix of the data moments $\hat{M}$ following Erickson and Whited (2012), and

$$\Omega(\hat{\theta}) = \frac{1}{S}\left(\hat{m}^s(\hat{\theta}) - \frac{1}{S}\sum_{s=1}^{S}\hat{m}^s(\hat{\theta})\right)\left(\hat{m}^s(\hat{\theta}) - \frac{1}{S}\sum_{s=1}^{S}\hat{m}^s(\hat{\theta})\right)' \tag{4.3}$$

is the estimate of sampling errors of model moments $\hat{m}$. I estimate Ω using a first stage estimate of the model that uses identity weighting matrix.

Given the choice of the optimal weighting matrix W in (4.2), the standard errors of the parameter estimates $\hat{\theta}$ is

$$V(\hat{\theta}) = \left(\hat{G}'W\hat{G}\right)^{-1} \tag{4.4}$$

where $\hat{G}$ is an estimate of the gradient of the moment conditions evaluated at the parameter estimates $\hat{\theta}$.

4.3 Estimation Results and Quantitative Implications

Table 2 reports the parameter estimates for (λ, σ_z), along with their standard errors. The estimated value of λ suggests that every year 21% of the firms "meet" other firms as targets – so 42% firms "think about" merger opportunities every year. This may not be an overly exaggerating number, given that US has an active corporate assets market, and anecdotal news abounds that managers discuss with each other potential merger opportunities. The estimated value of σ_z suggests the precision of market signal is similar to that owned by the managers. Besides, top panel of the table shows that model simulated moments closely match the data moments. Bottom panel shows the estimated gradient of the moment conditions evaluated at the parameter estimates, suggesting the local sensitivity measure of Andrews et al. (2017) is large.

[Insert Table 2 near here]

Table 3 reports the quantitative implications of the model on resource misallocation. It reports the effect of a counterfactural change in σ_z governing the precision of market signal, on aggregate output, the probability of wrong MA decisions, and the average time to sale of firms with bottom 10% fundamental productivities. Aggregate output is yearly-average of the sum of all firms' output in the 60 years of simulated data. A 50% decrease in σ_z from the current value, that is, an improvement in price informativeness, will lead to 0.57% increase in aggregate output. Meanwhile, a 50% increase σ_z from the current value, will lead to a 1.64% decline in aggregate output. This means a further improvement in price informativeness of US stock market may not improve aggregate output that much, but a deterioration would lead to a much larger decline in aggregate output.

Of course, whether the effect on aggregate output is large ultimately depends on how difficult one thinks it is to increase or decrease price informativeness. My results are however in contrast to David et al. (2016). Based on their preferred model where only capital is chosen under imperfect information, while labor can adjust perfectly to contemporaneous conditions, they find the total loss of TFP due to information friction is 4%, however, the total effect of market information is only associated with 0.2% TFP gains. That is, learning from the financial market contributes little. My results however show that aggregate output could decline as much as 1.64% if the US stock market is 50% less informationally efficient than it is. My estimates for the informational role of financial markets, however, are likely to be sensitive to model assumptions.

A decomposition of the effect shows that, the main source of output gain comes from the better informed merger decisions given the higher stock price informativeness. When σ_z decrease 50% from the current value, the probability of wrong merger decisions declines by 1.22%. That is, when stock price informativeness is higher, the fraction of mergers with true gains but not pursued, and mergers with negative gains but eventually conducted, decline by 1.22%. When σ_z increases by 50% from the current value, the probability of wrong merger decisions increases by 2.58%. Better stock price informativeness also facilitates the sale of assets from low-productivity owners. A 50% decrease in σ_z shortened the average time to asset sale from the owners with lowest 10% fundamental productivity, by 0.03%. A 50% increase would rather lengthen the average time to sale by 0.2%.

[Insert Table 3 near here]

Chapter 5

Conclusion

How much does learning from financial markets affect resource misallocation? I develop a dynamic model that features financial markets guiding managers in large investment decisions – mergers and acquisitions. Due to information frictions, misvaluation of own firms and the potential gain from mergers and acquisitions deters socially beneficial resource reallocation. Learning from the financial markets accumulates over time, and also occurs upon the announcement of the mergers and acquisitions. My structural estimation targets novel data moments including sensitivity of M&A cancellation with respect to announcement period returns to identify learning. The estimates suggest that a 50% decrease in stock price informativeness would lead to 1.64% output loss for the US economy.

Appendix A

Tables

Table 1: Parameterization: Summary

Parameter	Description	Target/Value
Panel A. Assigned parameters		
β	Discount rate	0.95 (Midrigan and Xu, 2014)
ϕ	Curvature of production function	0.63 (Hennessy and Whited, 2007)
δ	Capital depreciation	0.06 (Midrigan and Xu, 2014)
α	Adjustment cost of capital	2, Quadratic adjustment cost
η	Bargainning power	0.51 (David, 2017)
Panel B. Calibrated		
ρ	Persistence of productivity	From regression (David et al., 2016):
σ_a^2	Variance of productivity shock	
(m, σ_0^2)	Mean and variance of fundamental productivity	$a_{it} = \rho\theta + (1-\rho)a_{i,t-1} + u_{it}$
Panel C. Estimated		
λ	Intensity of meeting	Frequency of M&A deal announcements
σ_z	Std. error of market private signal	Sensitivity of M&A cancellation probability

Table 2: Target moments and parameter estimates

Moments	**Target**	**Model**
MA announcement frequency	0.037	0.036
Sensitivity of MA cancellation	0.04	0.04
Parameters	**Estimates**	**Standard Error**
λ	0.21	(0.12)
σ_z	0.52	(0.19)
Gradient Matrix		
λ	0.039	0.006
σ_z	0.012	-0.15

Notes. This table reports the estimated parameters of the model and assess the model fit and identification. See the main text for details on the estimation of the model.

Table 3: Consequences of learning on aggregate output

	TFP	Output	Pr(wrong MA decisions)	Average time to sale
σ_z dec. by 50%	1.65%	0.57%	-1.22%	-0.03%
σ_z inc. by 50%	-2.49%	-1.64%	2.58%	0.20%

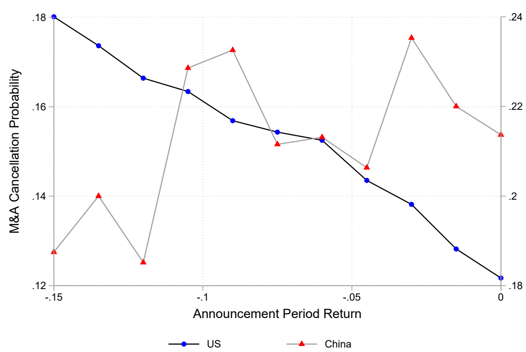
Notes. This table reports the effect of a counterfactural change in σ_z governing the precision of market signal, on TFP, aggregate output, the probability of wrong MA decisions, and the average time to sale of firms with bottom 10% fundamental productivities. See the main text for details.

Appendix B

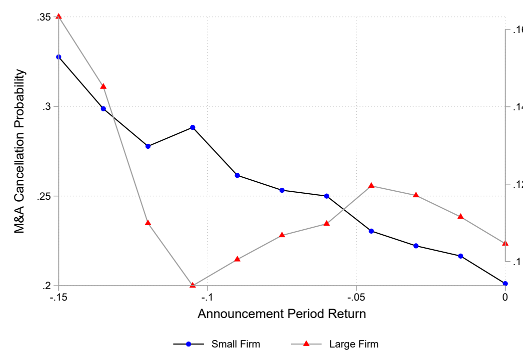
Figures

Figure 1: M&A cancellation, announcement returns and ex-post returns – evidence of learning in M&A deals

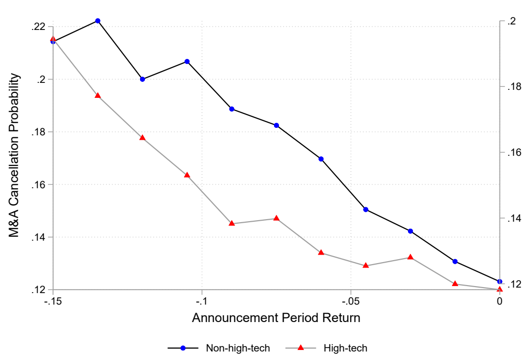
(a) Cancellation probability conditional on returns



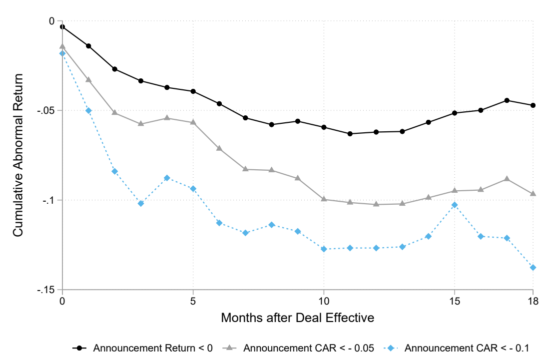
(b) Small acquirors vs large acquirors



(c) High-tech deals vs non-high-tech deals

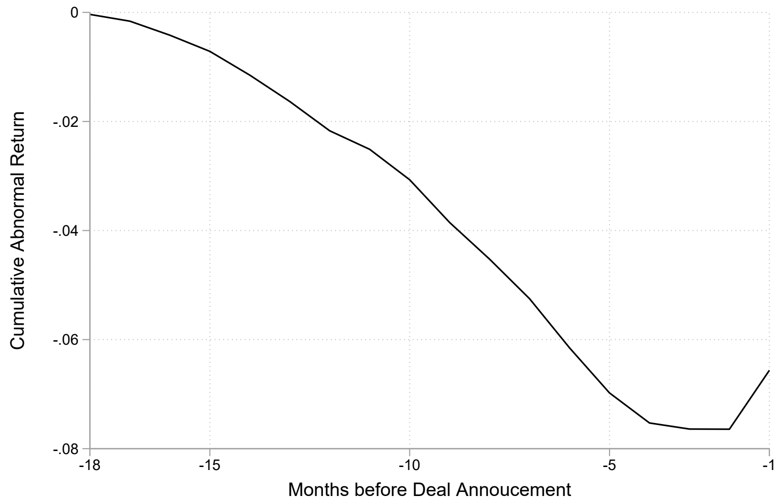


(d) Stock returns after loss-creating deals effective



Notes. Panel (a) plots the probability of M&A deal cancellation conditional on announcement period cumulative abnormal returns (CAR) less than a sequence of cutoffs, for US and Chinese firms. Panel (b) plots the cancellation probability separately for small acquirors (market capitalization below 25% percentile) and large acquirors (market capitalization above 75% percentile). Panel (c) plots the cancellation probability separately for high-tech deals and non-high-tech deals. If both the acquiror and the target are high-tech companies, the deal is a high-tech deal. Otherwise, it is a non-high-tech deal. Panel (d) plots the cumulative abnormal returns after the deal was effective, for deals with announcement period CAR being negative, below 5% and below 10%.

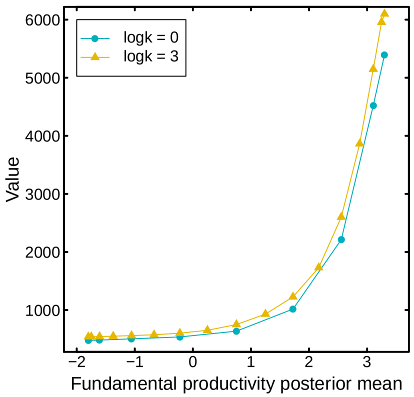
Figure 2: Stock returns before voluntary asset sales – evidence of learning over time



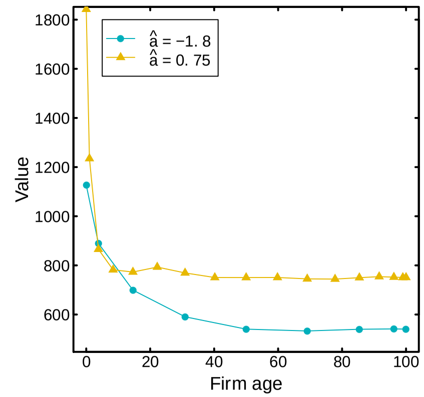
Notes. Figure plots the cumulative abnormal returns of firms before they sell corporate assets. Monthly abnormal return is calculated as firm stock return less return on CRSP value-weighted index. The sample includes voluntary asset sales between 1990 and 2003 in the SDC database, including mergers, acquisitions, and partial-firm asset sales. To exclude the firms who sell assets because they are in financial distress, asset sales whose documented purposes are to pay outstanding debt, raise cash through disposal, or sell a loss-making/bankrupt operation, are excluded from the sample.

Figure 3: Value function and firm investment decision

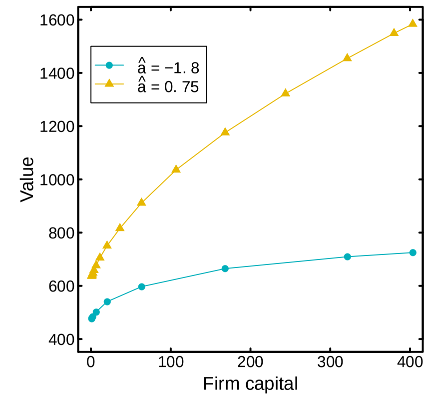
(a) Fundamental productivity posterior belief



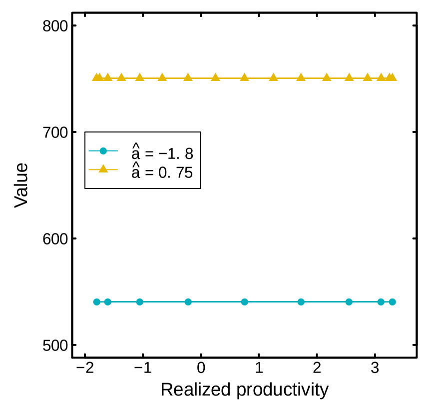
(b) Firm age



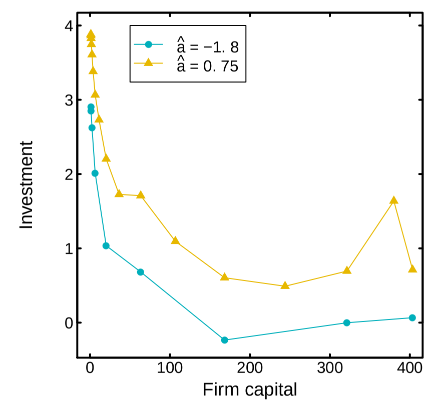
(c) Firm capital



(d) Firm realized productivity

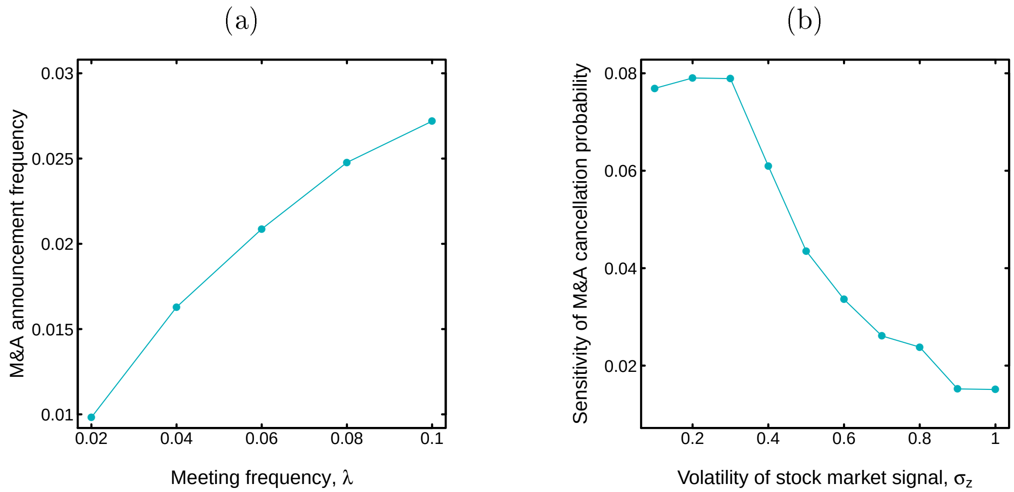


(e) Optimal investment decision



Notes.Figure plots the value function when $\lambda = 0.2, \sigma_z = 0.5$, and other parameters take values as in Table 1. Panel (a) plots the value function against manager's posterior mean of fundamental productivity $\hat{\theta}$. Panel (b) plots the value function against firm's age τ. Panel (c) plots the value function against firm capital. Panel (d) plots the value function against realized productivity. Panel (e) plots optimal investment against firm capital.

Figure 4: Identification of model parameters



Notes. Left panel plots the relationship between parameter λ and the frequency of MA deal announcements in simulated data, when $\sigma_z = 0.53$. Right panel plots the relationship between parameter σ_z and the sensitivity of deal cancellation probability with respect to announcement period return, that is, cancellation probability when announcement period return is less than -8% less that when announcement period return is less than 0%.

Appendix C

Numerical Solution to Bellman Equation

This Appendix describes how I numerically solve the Bellman equation to find the manager's optimal investment decision.

I approximate the value function by time iterations. I start by discretizing the state space. State variable τ_t takes values in the set $\{0, 1, ..., \bar{\tau} - 1\}$, where $\bar{\tau}$ is the maximum number of years of survival. I let $\hat{\theta}$ take values in finite set M, which contains grid points in the interval $[m - 2\sigma_0, m + 2\sigma_0]$. a take values in finite interval $[m - 2\sqrt{\rho^2\sigma_0^2 + \sigma_a^2}, m + 2\sqrt{\rho^2\sigma_0^2 + \sigma_a^2}]$. And let the capital stock grid be a multiplicative sequence so that $k^{i+1} = k^i/(1 - \delta)$ where k^{i+1} and k^i are adjacent values in the grid. The length of the intervals does not need to be extremely large, as extrapolation beyond the intervals turns out being accurate.

I set 16 grid points for each dimension. However, given the high dimension of the state space, I do not use the Cartesian grids, that is, the cross products of all one dimension grids. Rather, I use sparse grids (See Appendix D) to reduce the number of grid points. The value function is only defined on the grid points. Whenever needed, I use piece-wise linear interpolation and extrapolation to obtain the value function evaluated at arbitrary points. I start with a guess of V^0 over the grid:

$$V^0(\hat{\theta}, \tau, k, a) = \frac{1}{1 - \beta}\left[\exp(\rho\hat{\theta} + (1 - \rho)a + \phi k) + \tau\right] \tag{C.1}$$

Then I update the value function according to

$$V^{t+1}(\hat{\theta}_t, \tau_t, k_t, a_t) = \max_i \Big\{ exp(a_t)k_t^{\phi} - i_t^{\alpha} + \beta\lambda\mathbb{E}_t^* \max\{\Phi^t(\hat{\theta}'_M, \hat{\theta}_{t+1}, \hat{\theta}_T), 0\} + \beta\mathbb{E}_t^* V^t(\hat{\theta}_{t+1}, \tau_{t+1}, k_{t+1}, a_{t+1}) \Big\} \tag{C.2}$$

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \theta_a(\tau_{t+1})\delta_a + \theta_e(\tau_{t+1})\delta_e + \theta_z(\tau_{t+1})\delta_z \tag{C.3}$$

$$\begin{pmatrix} \delta_a \\ \delta_e \\ \delta_z \end{pmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \sigma_a^2/\rho^2 + \hat{v}(\tau) & 0 & 0 \\ 0 & \sigma_e^2 + \hat{v}(\tau) & 0 \\ 0 & 0 & \sigma_z^2 + \hat{v}(\tau) \end{bmatrix}\right) \tag{C.4}$$

$$a_{t+1} \sim \mathcal{N}\Big((1-\rho)a_t + \rho\hat{\theta}_t, \sigma_a^2 + \rho^2\hat{v}(\tau)\Big) \tag{C.5}$$

$$Cov(\hat{\theta}_{t+1}, a_{t+1}) = \theta_a(\tau_{t+1})Var(a_{t+1})/\rho \tag{C.6}$$

In value function update, two places involve calculation of expectations. I approximate the second expectation, $\mathbb{E}_t^* V^t(\hat{\theta}_{t+1}, \tau_{t+1}, k_{t+1}, a_{t+1}) = \int V^t dF(\hat{\theta}_{t+1}, a_{t+1})$ using Gauss-Hermite quadrature with seven nodes (See Appendix E), since the conditional distribution of $(\hat{\theta}_{t+1}, a_{t+1})$, $F(\hat{\theta}_{t+1}, a_{t+1})$ is normal and given by the transitional probabilities described above. I approximate first expectation, the expected merger gain, utilizing law of iterated expectations,

$$\mathbb{E}_t^*\big[\max\{\Phi(\hat{\theta}'_M, \hat{\theta}_{t+1}, \hat{\theta}_T), 0\}\big] \tag{C.7}$$

$$= \mathbb{E}\Big[\mathbb{E}\big[max\{\Phi(\hat{\theta}'_M, \hat{\theta}_{t+1}, \hat{\theta}_T), 0\} \mid \hat{\theta}'_M, \hat{\theta}_{t+1}, a_{t+1}\big]\Big] \tag{C.8}$$

where the inner expectation is taken with respect to the cross-section distribution of firm characteristics, and the outer expectation taken with respect to $(\hat{\theta}'_M, \hat{\theta}_{t+1}, a_{t+1})$, which has a joint normal distribution. The inner expectation is approximated as an average across the $N = 300$ firms with characteristics distributed as $\Omega(\hat{\theta}, \tau, k, a)$ – a Monte Carlo method for calculating expectation. Specifically, I assume $\hat{\theta}$ and a are normally distributed with appropriate mean and variance that approximate the data, and τ and $\log k$ are uniformly distributed in the interval of the grid, and the four variables are independent. The outer expectation is approximated using Gauss-Hermite quadrature with seven nodes.

I stop value function iteration as soon as

$$\max_{(\hat{\theta}, \tau, k, a) \in G} |V^{t+1} - V^t| < 10^{-3} \tag{C.9}$$

Appendix D

Sparse Grids

This Appendix describes how I use sparse grids to reduce the computation burden caused by the high-dimensional state space.

The problem of interpolating multi-dimensional functions commonly arise in economics. In particular, when solving dynamic economic models, one needs to interpolate value function in terms of state variables. With few state variables, we can use tensor-product (i.e. Cartesian-product) rules. But tensor-product rules quickly become computationally infeasible when the number of state variables increases, a fact referred to as the curse of dimensionality. For example, if there are 10 grid points for each variable, a tensor product grid for 4 variables would contain 10^4 points.

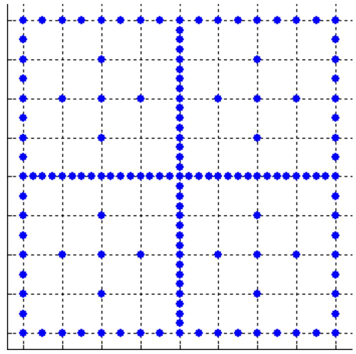
The sparse grids method, introduced by Smolyak (1963), select a subset of tensor-product grid that are more important for function values – the idea is we don't need interactions of high-order terms with high-order terms. A parameter, called the level of approximation, controls how many tensor-product elements are included into the sparse grid. The higher the level parameter, the better quality is the approximation.

I start with $m(\mu) = 2^{\mu-1}+1$ points (including boundaries) for each one-dimensional set, where μ is the level of approximation. Then I construct the tensor product of the one-dimensional sets of points. Suppose i_1, i_2, i_3, i_4 are indices that correspond to dimension 1 to 4 (my model has $d = 4$ state variables), the following Smolyak rule tells us which tensor products are selected,

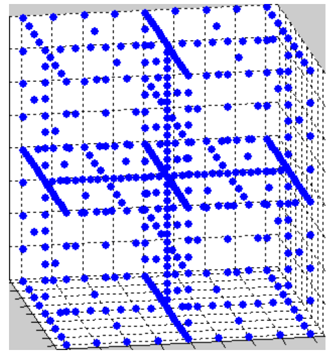
$$d \leq i_1 + i_2 + i_3 + i_4 \leq d + \mu \tag{D.1}$$

Figure D.1 plots the tensor product points selected in the sparse grid for 2-dimensional and 3-dimensional state space, with the level of approximation = 5.

Figure D.1: Sparse grid points for 2-dimensional and 3-dimensional state space



(a) dimension = 2, level = 5



(b) dimension = 3, level = 5

To interpolate multi-dimensional functions off the sparse grid points, following Bungartz and Griebel (2004) and Brumm and Scheidegger (2017), I use the piecewise d-linear interpolation. This is an analogue of piecewise linear interpolation in 1-dimensional case. We first define the standard hat function

$$\phi(x) = \begin{cases} 1 - |x|, \text{if } x \in [-1, 1] \\ 0, \text{otherwise.} \end{cases} \tag{D.2}$$

Then we define the 1-D basis function at each 1-D grid point i, with support $[(i-1)h, (i+1)h]$. For any arbitrary $x = (x_1, x_2, x_3, x_4)$, the 1-D basis function is

$$\phi_i(x) = \phi(\frac{x - ih}{h}) \tag{D.3}$$

The 4-D basis function at grid point $\mathbf{i} = (i_1, i_2, i_3, i_4)$ is the product of the 1-D basis functions,

$$\Phi_{\mathbf{i}}(x) = \prod_{j=1}^{4} \phi_{j,i}(x_j) \tag{D.4}$$

Then function value at arbitrary point can be approximated by a linear combination of the $\Phi_{\mathbf{i}}(x)$ associated with every grid point.

Appendix E

Gauss-Hermite Quadrature

This Appendix briefly describes the Gauss-Hermite quadrature method I used when calculating the expectations in value function iteration.

Gauss-Hermite quadrature approximates the expectation of a function of a single random variable $\tilde{X}$ that has normal distribution, [1] which requires only the computation of a weighted sum.

$$\mathbb{E}[f(\tilde{X})] = \int f(x)w(x)dx \approx \sum_{i=1}^{n} f(x_i)w_i \tag{E.1}$$

where the nodes $\{x_i\}$ and weights $\{w_i\}$ are chosen to minimize the approximation errors.

In value function iteration, I need to calculate expectation of continuation value with respect to $(\hat{\theta}_{t+1}, a_{t+1})$, and expectation of merger gain with respect to $(\hat{\theta}'_M, \hat{\theta}_{t+1}, a_{t+1})$. Tensor product principles can be applied to univariate Gaussian quadrature rules to develop quadrature rules for multivariate integration. Following Miranda and Fackler (2004), suppose that $\tilde{X}$ is a d-dimensional normal random variable (row vector) with mean vector μ and variance-covariance matrix Σ. Then $\tilde{X}$ is distributed as $\mu + \tilde{Z}R$, where R (upper triangular) is the Cholesky square root of Σ (e.g. $\Sigma = R^T R$) and $\tilde{Z}$ is a row d-vector of independent standard normal variates. If z_i, w_i are the degree n Gaussian nodes and weights for a standard normal variate, then an n^d degree approximation for $\tilde{X}$ may be constructed using tensor products. For example, in two dimensions the nodes and weights would take the form

$$x_{ij} = (\mu_1 + R_{11}z_i + R_{21}z_j, \mu_2 + R_{12}z_i + R_{22}z_j) \tag{E.2}$$

and

$$w_{ij} = w_i w_j \tag{E.3}$$

[1] Gauss-Hermite quadrature applies to the weighting function $w(x) = exp(-x^2)$, as opposed the weighting function for the standard normal density $w(x) = exp(-x^2/2)/\sqrt{2\pi}$

$$\mathbb{E}f(\tilde{X}) = \mathbb{E}f(\mu + zR) = (\frac{1}{\sqrt{2\pi}})^d \int_{-\infty}^{\infty} f(\mu + zR) exp(-z^2/2) dz \tag{E.4}$$

$$= \pi^{-d/2} \int_{-\infty}^{\infty} f(\mu + \sqrt{2}\eta R) exp(-\eta^2) d\eta \tag{E.5}$$

$$\approx \pi^{-d/2} \sum_i \sum_j w_{ij} f(x_{ij}) \tag{E.6}$$

Bibliography

Andrews, Isaiah, Matthew Gentzkow, and Jesse M. Shapiro (2017), "Measuring the Sensitivity of Parameter Estimates to Estimation Moments," Quarterly Journal of Economics, 132, 1553–1592.

Bakke, Tor-Erik and Toni M Whited (2010), "Which Firms Follow the Market? An Analysis of Corporate Investment Decisions," Review of Financial Studies, 23, 1941–1980.

Bennett, Benjamin, René Stulz, and Zexi Wang (2020), "Does the Stock Market Make Firms More Productive?," Journal of Financial Economics, 136, 281–306.

Bento, Pedro and Diego Restuccia (2017), "Misallocation, Establishment Size, and Productivity," American Economic Journal: Macroeconomics, 9, 267–303.

Bond, Philip, Alex Edmans, and Itay Goldstein (2012), "The Real Effects of Financial Markets," Annual Review of Financial Economics, 4, 339–360.

Brumm, Johannes and Simon Scheidegger (2017), "Using Adaptive Sparse Grids to Solve High-Dimensional Dynamic Models," Econometrica, 85, 1575–1612.

Bungartz, Hans-Joachim and Michael Griebel (2004), "Sparse Grids," Acta Numer, 13, 147–269.

Chen, Kaiji and Zheng Song (2013), "Financial Frictions on Capital Allocation: A Transmission Mechanism of TFP Fluctuations," Journal of Monetary Economics, 60, 683–703.

Chen, Xia, Jarrad Harford, and Kai Li (2007), "Monitoring: Which institutions matter?," Journal of Financial Economics, 86, 279–305.

Datta, Sudip, Mai P Iskandar-Datta, and Kartik Raman (2001), "Executive compensation and corporate acquisition decisions," Journal of Finance, 56.

David, Joel (2017), "The Aggregate Implications of Mergers and Acquisitions," Working Paper.

David, Joel M, Hugo A Hopenhayn, and Venky Venkateswaran (2016), "Information, Misallocation, and Aggregate Productivity," Quarterly Journal of Economics, 131, 943–1005.

Durnev, Art, Randall Morck, and Bernard Yeung (2004), "Value-enhancing Capital Budgeting and Firm-specific Stock Return Variation," Journal of Finance, 59, 65–105.

Edmans, Alex, Sudarshan Jayaraman, and Jan Schneemeier (2017), "The Source of Information in Prices and Investment-Price Sensitivity," Journal of Financial Economics, 126, 74–96.

Erickson, Timothy and Toni M Whited (2012), "Treating Measurement Error in Tobin's q," Review of Financial Studies, 25, 1286–1329.

Feng, Ying (2018), "Firm Life-Cycle Learning and Misallocation," Working Paper.

Hayek, Friedrich (1945), "The Use of Knowledge in Soceity," American Economic Review, 35, 519–530.

Hennessy, Christopher A and Toni M Whited (2007), "How Costly is External Financing? Evidence from a Structural Estimation," Journal of Finance, 62, 1705–1745.

Hite, Gailen, James E Owers, and Ronald C Rogers (1987), "The Market for Interfirm Asset Sales: Partial Sell-offs and Total Liquidations," Journal of Financial Economics, 18, 229–252.

Hopenhayn, Hugo A (1992), "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," Econometrica, 60, 1127–1150.

Hsieh, Chang-Tai and Peter J Klenow (2009), "Misallocation and Manufacturing TFP in China and India," Quarterly Journal of Economics, 124, 1403–1448.

Jain, Prem (1985), "The Effect of Voluntary Sell-off Announcements on Shareholder Wealth," Journal of Finance, 40, 209–224.

John, Kose and Eli Ofek (1995), "Asset Sales and the Increase in Focus," Journal of Financial Economics, 37, 105–126.

Jovanovic, Boyan (1979), "Job Matching and the Theory of Turnover," Journal of Political Economy, 87, 972–990.

Jovanovic, Boyan (1982), "Selection and the Evolution of Industry," Econometrica, 649–670.

Kau, James B, James S Linck, and Paul H Rubin (2008), "Do Managers Listen to the Market?," Journal of Corporate Finance, 14, 347–362.

Lambrecht, Bart and Stewart C Myers (2007), "A Theory of Takeovers and Disinvestment," Journal of Finance, 62, 809–845.

Lang, Larry, Annette Poulsen, and Rene Stulz (1995), "Asset Sales, Firm Performance, and the Agency Costs of Managerial Discretion," Journal of Financial Economics, 37, 3–38.

Lehn, Kenneth M and Mengxin Zhao (2006), "CEO Turnover after Acquisitions: Are Bad Bidders Fired?," Journal of Finance, 61, 1759–1811.

Li, Kai, Chi-Yang Tsou, and Chenjie Xu (2020), "Learning and the Capital Age Premium," Working Paper.

Lipster, Robert and Albert N Shiryaev (2001), Statistics of Random Processes. Springer.

Luo, Yuanzhi (2005), "Do Insiders Learn from Outsiders? Evidence from Mergers and Acquisitions," Journal of Finance, 60, 1951–1982.

Maksimovic, Vojislav and Gordon Phillips (2001), "The Market for Corporate Assets: Who Engages in Mergers and Asset Sales and Are There Efficiency Gains?," Journal of Finance, 56, 2020–2065.

Malmendier, Ulrike and Geoffrey Tate (2008), "Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction," Journal of Financial Economics, 89, 20–43.

Michaelides, Alexander and Serena Ng (2000), "Estimating the Rational Expectations Model of Speculative Storage: A Monte Carlo Comparison of Three Simulation Estimators," Journal of Econometrics, 96, 231–266.

Midrigan, Virgiliu and Daniel Yi Xu (2014), "Finance and Misallocation: Evidence from Plant-Level Data," American Economic Review, 104, 422–58.

Miranda, Mario J and Paul L Fackler (2004), Applied Computational Economics and Finance. MIT press.

Mitchell, Mark L and Kenneth Lehn (1990), "Do Bad Bidders Become Good Targets?," Journal of Political Economy, 98, 372–398.

Moeller, Sara B, Frederik P Schlingemann, and Rene M Stulz (2005), "Wealth Destruction on a Massive Scale? A Study of Acquiring-firm Returns in the Recent Merger Wave," Journal of Finance, 60.

Morck, Randall, Andrei Shleifer, and Robert W Vishny (1990), "Do Managerial Objectives Drive Bad Acquisitions?," Journal of Finance, 45, 31–48.

Oksendal, Bent (2003), Stochastic Differential Equations: An Introduction with Applications. Springer.

Roll, Richard (1986), "The Hubris Hypothesis of Corporate Takeovers," Journal of Business, 197–216.

Shimer, Robert and Lones Smith (2000), "Assortative Matching and Search," Econometrica, 68, 343–369.

Smolyak, Sergey A (1963), "Quadrature and Interpolation Formulas for Tensor Products of Certain Classes of Functions," Dokl. Akad. Nauk, 148, 1042–1045.

Sockin, Michael (2015), "Not So Great Expectations: A Model of Growth and Informational Frictions," Working Paper Univerisity of Texas at Austin.

Tan, Yue, Li Gao, and Hui Zhou (2015), "Stock Price Expectation, Learning and M&A Termination: Evidence from Major Asset Restructuring Event," Research on Economics and Management, 36.

Van Binsbergen, Jules H and Christian C Opp (2019), "Real Anomalies," Journal of Finance, 74, 1659–1706.