Essays on Firm and Household Dynamics

by

Junke (Jake) Zhao

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Economics)

at the

University of Wisconsin-Madison

2014

Date of final oral examination: 5/27/2014

The dissertation is approved by the following members of the final oral committee:

Philip Dean Corbae, Professor, Economics and Finance

Erwan Quintin, Associate Professor, Real Estate

Mark Ready, Professor, Finance

Kenichi Fukushima, Assistant Professor, Economics

Nicolas Roys, Assistant Professor, Economics

Acknowledgments

My economics PhD journey began at the University of Texas at Austin. The first year was so tough and tedious that I had some serious misgivings, but fortunately, the second year field courses were able to reignite my interests. In particular, I developed a deep passion for computational economics and did well in the class. The professor, Dean Corbae, asked me at the end of the semester to be his research assistant in the spring. Then when he decided to move to the University of Wisconsin-Madison at the end of my second year, he asked me to accompany him. I was flattered by the opportunity and my move to Madison was probably the single most positive shock in my academic career. The professors at the University of Wisconsin were extremely active and they were able to provide invaluable guidance for my research. My advisor Dean Corbae and my committee members Erwan Quintin, Mark Ready, Kenichi Fukushima, and Nicolas Roys were immensely helpful and I am forever indebted to them. Dean and Erwan above all pushed me to become the best researcher that I could be in spite of my resistance at times.

The University of Wisconsin and the city of Madison truly hold a special place in my heart. My fiancée, Jiao Shi was a fellow Wisconsin economics PhD student and we spent many days and nights enjoying our time together in Madison. She has been supportive in all aspects of my life and I am eternally grateful to her. Her unwavering love has and will continue to make me a better person.

Finally, my parents, Zaiming Zhao and Lixia Li, and my younger brother, Yuke Zhao, provided a nurturing environment as I grew up where knowledge and learning is valued and where I constantly felt loved and appreciated. My parents were always willing to make sacrifices even if to marginally improve my life. I could not envision a better family and my PhD is just a small reflection of how great my family is.

Dissertation Abstract

Why do U.S. firms hold much more cash now than they did 30 years ago? Prior empirical studies have discovered a statistically significant positive relationship between firm cash holdings and cash flow volatility. Such findings, however, are subject to endogeneity problems. I construct a structural model of firm dynamics where cash provides a buffer against cash-flow shortfalls in the presence of costly external finance. My model finds that 63% of the increase in corporate cash holdings can be accounted for by the increase in cash flow volatility. The increase in cash flow volatility observed in the data arises from a decrease in the correlation between revenue and operating expenses. The model has a corresponding correlation parameter between the shocks on revenue and operating expenses and only this parameter is changed in the primary experiment. The decomposition of revenue and operating expenses is important and I show that other ways of modeling the cash flow volatility increase are both counterfactual and dampening. A regression using the model data then generates a coefficient on cash flow volatility similar to what was found in previous studies which suggests that the regression underestimates the true impact of volatility.

My second chapter shows that small Chinese manufacturing firms are less constrained than large Chinese manufacturing firms. Evidence of this finding comes from analyzing the cash flow sensitivity of investment and from estimating an investment Euler equation using GMM. My result is in direct contrast to most of the literature on financial constraints where small firms are found to be substantially more constrained. Small Chinese manufacturing firms are less constrained because they are far more labor intensive relative to large Chinese manufacturing firms and are not as reliant on capital investment for growth.

Why was the velocity of money so low after quantitative easing? My last chapter determines that 67% of the drop in the velocity of money can be attributed to the low wealth elasticity of demand for consumption goods and the high wealth elasticity of demand for financial assets by the people who hold the majority of their wealth in debt securities.

Contents

CHAPTER 1

Accounting for the Corporate Cash Increase

Abstract. Why do U.S. firms hold much more cash now than they did 30 years ago? Prior empirical studies have discovered a statistically significant positive relationship between firm cash holdings and cash flow volatility. Such findings, however, are subject to endogeneity problems. In this paper, I construct a structural model of firm dynamics where cash provides a buffer against cash-flow shortfalls in the presence of costly external finance. My model finds that 63% of the increase in corporate cash holdings can be accounted for by the increase in cash flow volatility. The increase in cash flow volatility observed in the data arises from a decrease in the correlation between revenue and operating expenses. The model has a corresponding correlation parameter between the shocks on revenue and operating expenses and only this parameter is changed in the primary experiment. The decomposition of revenue and operating expenses is important and I show that other ways of modeling the cash flow volatility increase are both counterfactual and dampening. A regression using the model data then generates a coefficient on cash flow volatility similar to what was found in previous studies which suggests that the regression underestimates the true impact of volatility. Finally, I investigate the response of cash holdings to policy changes and the consequences of cash restrictions on firm value.

1.1. Introduction

In the last 30 years, the cash-to-assets ratio of U.S. industrial firms has increased significantly. Bates et al. (2009) report that the cash-to-assets ratio more than doubles from 10.5% in 1980 to 23.2% in 2006 and has risen in every major industry. When firms are categorized by size as in Figure 1.1.1, it can be seen that the cash[1] buildup of small firms is even greater. For firms with less than 1 billion 2010 dollars in total assets, the cash ratio almost triples from 1980 to 2010. This upward trend in the cash ratio is clearly an important and compelling feature of the data. The trend also appears to be remarkably linear and has little correlation with the aggregate fluctuations in the business cycle.[2] For instance, cash increases during the recessions of the early 1980s and early 2000s, while cash decreases during the recession of the late 2000s. The cash ratio change is then indeterminate across firm size categories for the early 1990s recession.

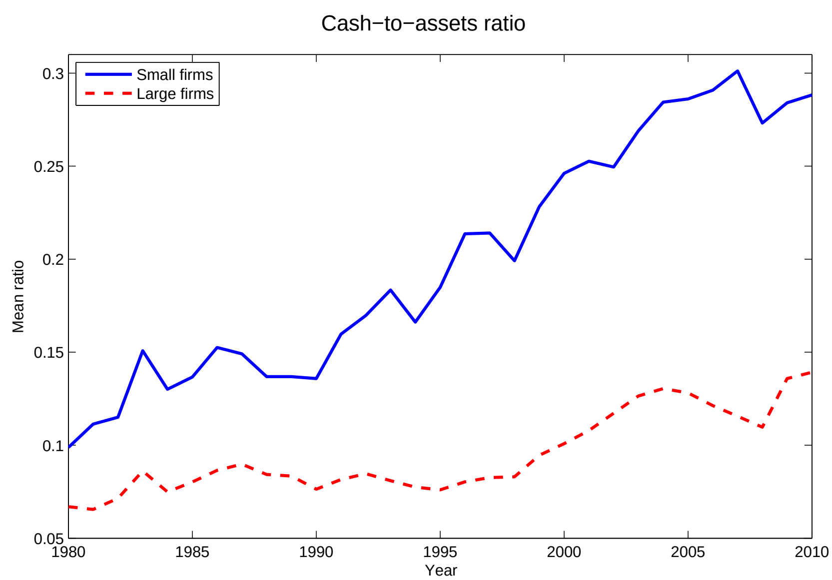


FIGURE 1.1.1. This figure plots the cash-to-assets ratio of firms categorized by size. Small (large) firms have less (more) than 1 billion 2010 dollars in total assets.

[1]The data definition of cash used in this paper is the cash and short-term investments variable (CHE) in Compustat.

[2]More precisely, the correlation between cash ratio growth and real GDP growth is -0.147 in the last 30 years.

In this same time period, cash flow volatility has also increased substantially as demonstrated in Figure 1.1.2. However, it should be emphasized that no direct causal link has been established in the empirical literature between cash and the volatility of cash flows. The correlation between cash and the volatility of cash flows suggests that a causal relationship is possible, but it is difficult to rule out confounding variables or even determine a causal direction, i.e. high cash holdings may instead cause high volatility in cash flows. In particular, it is hard to find good instruments which exogenously shift the volatility of cash flows but not the cash ratio and vice versa. A regression of cash on cash flow volatility almost always produces a significant coefficient, but at the same time, fails the Durbin-Wu-Hausman test for endogeneity with large F-statistics. The form of endogeneity of particular concern is simultaneity bias. It is straightforward to argue that cash flow volatility affects, in a causal way, most of the other regressors in standard cash regressions. These other regressors can also deliver significant coefficients if they are included. Therefore, even though the cash flow volatility coefficient has a large t-statistic, Bates et al. (2009) do not predict that the increase in cash is primarily due to the increase in cash flow volatility. They state, "holding all other variables constant, we infer that the average cash ratio increased by 2.1 percentage points from the 1980s to 2006 because of the increase in cash flow volatility." One of the advantages of a structural approach is that it is possible to clearly determine the direction of causation and disentangle the mechanisms behind the cash increase under a dynamic framework with rational expectations.

While the cash flow volatility has increased substantially over the last 30 years, it is interesting that revenue volatility and operating expenses volatility have not risen as seen in Figure 1.1.3. Rather, the correlation between revenue and operating expenses has declined as I find in Figure 1.1.4.[3] The decrease in the correlation between revenue and operating expenses is a possibly salient and important fact that has not been well investigated. In my

[3] Cash flow is mostly determined by revenue minus operating expenses. While there are other components such as interest, taxes, and depreciation, the variances and covariances contributed by these sources to cash flow variance is negligible.

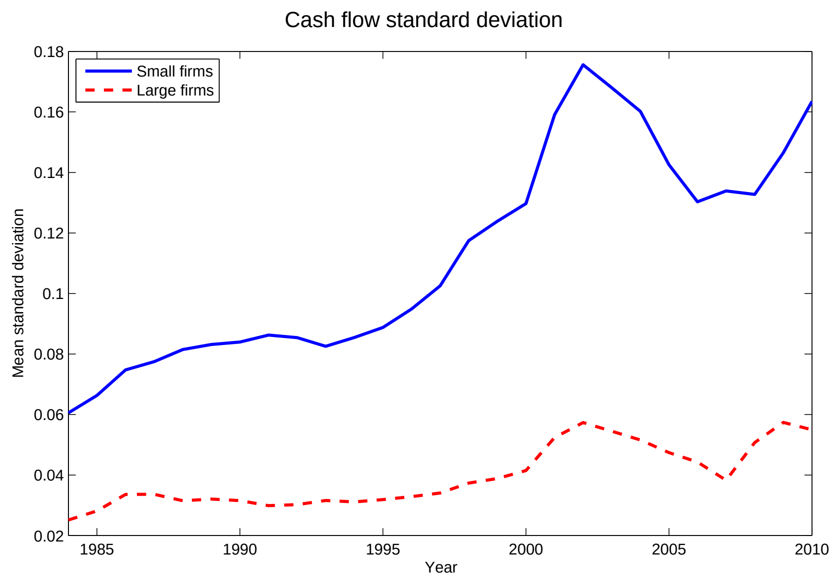


FIGURE 1.1.2. This figure plots the mean standard deviation of cash flow ((IB + DP)/AT) in 5 year panels where IB, DP, and AT refer to the income before extraordinary items, depreciation, and total assets variables in Compustat. The last year of the 5 year rolling panel is reported in the graph, and small (large) firms have less (more) than 1 billion 2010 dollars in total assets.

paper, the decrease in correlation is exogenous but plays an important role on how the model is constructed and estimated.[4] The correlation decrease occurs in every major industry and is also an independently interesting phenomenon that is explored in the appendix.

I introduce a buffer stock model of cash holdings with financing frictions where firms make dynamic capital, cash, equity flow, and exit decisions. The model is then taken to the data to determine that 63% of the increase in corporate cash holdings can be accounted for by the increase in cash flow volatility which arises from the decrease in correlation between revenue and operating expenses. The model has a corresponding correlation parameter between the shocks on revenue and operating expenses and only this parameter is changed to generate the model data. A regression using the model data then produces a coefficient on cash flow volatility similar to what was found in previous studies which indicates that standard cash regressions underestimate the true impact of volatility on cash holdings.

[4]Otherwise, the cash flow volatility increase would either have to come through a reduced-form and somewhat *ad hoc* cash flow shock, or through a counterfactual revenue or operating expenses volatility increase.

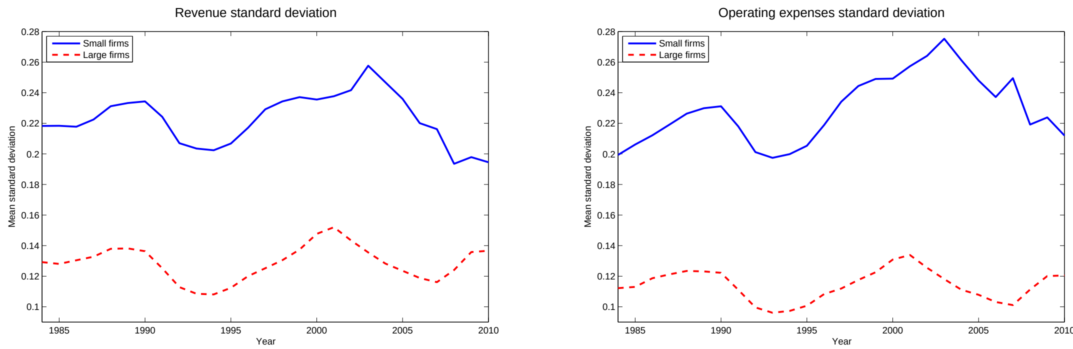


FIGURE 1.1.3. This figure juxtaposes the mean standard deviation of revenue (REVT/AT) and operating expenses (XOPR/AT) in 5 year panels where REVT, XOPR, and AT refer to the revenue, operating expenses, and total assets variables in Compustat. The last year of the 5 year rolling panel is reported in the graph, and small (large) firms have less (more) than 1 billion 2010 dollars in total assets.

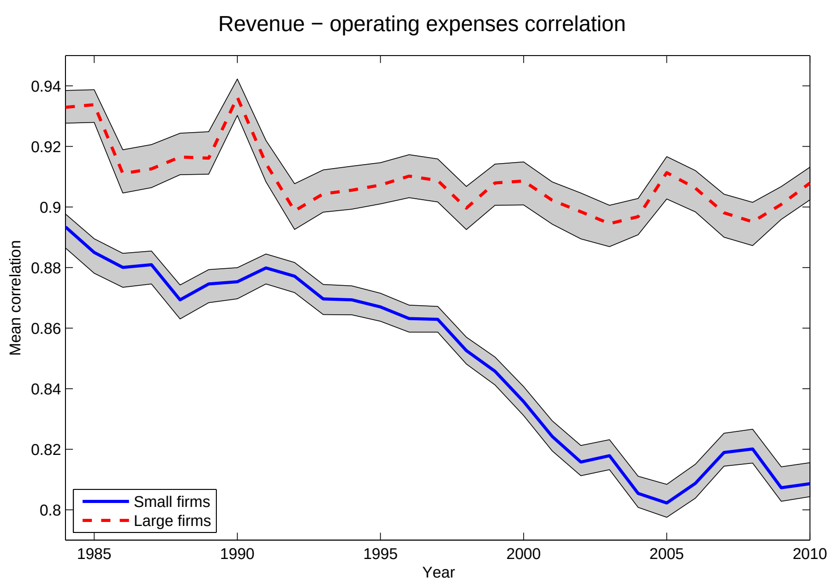


FIGURE 1.1.4. This figure displays the decrease in the mean correlation between revenue (REVT/AT) and operating expenses (XOPR/AT) in 5 year panels where REVT, XOPR, and AT refer to the revenue, operating expenses, and total assets variables in Compustat. The last year of the 5 year rolling panel is reported in the graph, and small (large) firms have less (more) than 1 billion 2010 dollars in total assets. The shaded area is two times the standard error above and below the mean.

The key mechanism is that, with a correlation decrease, revenue is no longer a natural hedge for operating expenses. In the past, when revenue fell, costs also fell, but now, when revenue falls, costs are less likely to fall. Therefore, this natural hedge occurs at a lesser degree which then translates to both more frequent and more severe negative cash flow events. Negative cash flow is especially harmful if cash is exhausted since the only options left to the manager are to sell off capital and/or raise costly external finance. Cash consequently acts as a buffer against cash flow shocks. Most other structural models in the corporate finance literature do not have the possibility of negative cash flows. However, management of negative cash flows and the implications of negative cash flows for default and exit are cited as central financial concerns by real world managers.[5] Other papers do not decompose revenue and operating expenses and they have difficulty producing the observed rise in corporate cash holdings with volatility alone. In particular, I show that increasing the cash flow volatility through an increase in revenue volatility produces counterfactual moments and a dampened effect on cash holdings.

In essence, my paper shows that the corporate cash increase can be mostly attributed to rational behavior in response to the idiosyncratic cash flow volatility increase. In fact, cash holdings are much less puzzling once the cash flow structure and shocks are modeled in a more comprehensive fashion. Using my model, I also demonstrate that policy attempts to motivate firms to invest or distribute their cash might have unintended consequences. Lowering the corporate tax rate or the real interest rate for instance increases investment and firm value but cash holdings increase as well. Finally, I show that cash restrictions can reduce firm value considerably.

The rest of the paper is organized as follows: Section 1.2 discusses my contribution relative to the literature, Section 1.3 presents the model, Section 1.4 provides intuition on the optimal cash policy, Section 1.5 details the results, Section 1.6 analyzes several policy experiments, Section 1.7 concludes, and the appendix contains the computational and normalization procedures as well as an analysis of the correlation decrease phenomenon.

[5] Lindsey and Carfang (April 4, 2013) conducted a quarterly survey of chief financial officers. In the survey, CFOs reported that paying down negative cash flows and financing capital expenditures are the two major uses of cash.

1.2. Literature

There are four widely-accepted motives for firms to hold cash - namely the transaction motive, the precautionary motive, the agency motive, and the tax motive. With the large increase in financial innovation in the last 30 years, it is quite surprising that the cash holdings of corporations have not decreased. Empirical papers such as Bates et al. (2009) indicate that the precautionary motive has an important but relatively small effect on the increase in the cash-to-assets ratio. On the other hand, they find insufficient evidence that agency conflicts make a meaningful contribution to the rise in the cash ratios of U.S. firms. Opler et al. (1999) also do not observe significant agency costs, but they do "find evidence that firms that do well tend to accumulate more cash than predicted by the static tradeoff model where managers maximize shareholder wealth." This indicates that the static tradeoff model is not rich enough to understand firm cash behavior and/or that there are behavioral explanations for the cash increase. They also noticed that derivative usage is quite rare (less than 10% of the observations) among S&P 500 firms. Derivative usage is rarer still among the small firms that I study in this paper. On the other hand, a recent paper by Nikolov and Whited (2011) argue that agency costs are important under certain assumptions on managerial incentives and contracts.

Han and Qiu (2007) construct a two-period model to study the role of the precautionary motive. They find that an increase in cash flow volatility increases the cash holdings of constrained firms but has no systematic effect on the cash holdings of unconstrained firms. In an infinite horizon structural model, every state in the ergodic distribution can be reached with nonzero probability in the future, so all firms are constrained to some degree. Therefore, it is more pertinent to think about the impact of a continuum of "constrainedness" in reference to the precautionary motive.

The effect of repatriation taxes on the cash holdings of multinational corporations was studied by Foley et al. (2007). Their paper concludes that repatriation taxes have a meaningful effect on multinational companies with big foreign tax spreads, but they cannot explain the cash buildup of other large firms or especially of small domestic firms. However, recall that small firms with under 1 billion 2010 dollars in total assets experience the largest increase in the cash ratio and they comprise 76.3% of Compustat firms. Also, over 90% of the firms under 1 billion 2010 dollars in total assets are classified as domestic corporations. Besides repatriation taxes, it is possible that the dramatic lowering of the corporate tax schedule over the postwar period is also a significant factor for the cash increase, but few papers on corporate cash holdings have directly investigated the quantitative effect of the tax decrease on the cash ratio. The tax effect is particularly ambiguous because of the different forces involved. A fall in the corporate tax rate diminishes the precautionary motive, which would reduce cash holdings. However, firms also tend to save out of increased cash flows.

My model in addition is able to generate a reduction in investment due to an increase in cash flow volatility as in Minton and Schrand (1999), and it is able to produce a wide cross-sectional distribution for the marginal value of cash as in Faulkender and Wang (2006) and Dittmar and Mahrt-Smith (2007) due to the rich shock structure and external finance costs.

The main structural focus in the corporate finance literature so far has been on the motivation to hold cash. For instance, Gamba and Triantis (2008) create a model where firms make dynamic debt and liquidity decisions. They find that financing frictions can cause firms to simultaneously borrow and lend which implies that cash is not just negative debt. On the other hand, Riddick and Whited (2009) focus on the cash flow sensitivity of cash, i.e. whether a firm tends to save or dissave out of cash flows. They find a negative propensity to save out of cash flow since firms in their model have large positive cash flows when they receive favorable profit shocks. The marginal value of capital increases with high profit shocks so that firms dissave to purchase more capital. The saving propensity is therefore not necessarily a good proxy to measure financial constraints and the costs of external finance. Bolton et al. (2011) then highlight the importance of the ratio of marginal

q to the marginal value of liquidity for the analysis of the investment and cash management problems.

Armenter and Hnatkovska (2012), Boileau and Moyen (2010), and Falato et al. (2013) also investigate the increase in cash holdings but they use different factors and mechanisms. Armenter and Hnatkovska (2012) state that firms hold more cash now because equity has become cheaper relative to debt. Boileau and Moyen (2010) look at the precautionary and transaction motives with a cash-in-advance structure which drives firm liquidity needs. Finally, Falato et al. (2013) assume that only tangible capital is pledgeable and they cite the rise in intangible capital usage as the explanation of the cash increase.

However, the dynamic firm decision in response to shocks is arguably the most fundamental problem. My model is based on the structural framework by Hennessy and Whited (2005) and Gomes (2001) which in turn are related to Cooley and Quadrini (2001) and Hopenhayn (1992). Overall, the specifics of my model and the questions I consider are quite different.

1.3. Model

1.3.1. Firm's problem. Assume that time is discrete and infinite, and that firms in the economy are risk-neutral. Firms are assumed to be owned by a representative risk-neutral agent that is not explicitly modeled. Firms in the economy are also heterogeneous but they face the same decision problems - therefore, I can refer to a single firm from now on without loss of generality. Let $k \in \mathbb{R}_+$ denote the capital stock and $m \in \mathbb{R}_+$ denote the cash holdings of the firm.[6] The firm comes into each period with these control variables as well as with revenue shock state variable z. The revenue shock $z \in [\underline{z}, \bar{z}] \equiv Z \subset \mathbb{R}_{++}$ is strictly positive, bounded, and has Markov transition function Γ.

[6]Previous versions of this paper contained debt as an additional continuous state variable and priced it competitively. This feature of the model engendered a lot of complexity and the results were not that different numerically. Corporate debt is unmodeled now and it is assumed that debt is rolled over each period without any frictions. On the other hand, I do not use net debt since the structure and assumptions on the collateral constraint can play a large role on the results.

The firm's production technology is assumed to exhibit decreasing returns to scale $\alpha < 1$ which implies that there exists a well-defined upper bound $\bar{k}$ on the optimal level of capital stock, where $\bar{k}$ will be defined later in the section. The firm's capital is therefore selected from the compact set $k \in [0, \bar{k}] \equiv K$.

Production is performed by the firm each period by using its capital to generate revenue. Operating expenses are proportional to the amount of capital used. This parsimonious specification encapsulates various costs the firm faces such as production costs, research and development costs, and selling and administrative expenses without modeling them separately.[7] The (operating) profit function is then,[8]

$$\pi(k, z, \eta_1, \eta_2; P) = P\eta_1 z k^\alpha - C(k, \eta_2) \tag{1.3.1}$$

where the cost function $C(k, \eta_2)$ has the form,

$$C(k, \eta_2) = \eta_2 c_v k + c_f. \tag{1.3.2}$$

The price $P \in \mathbb{R}_+$ can be thought of as the relative price of the homogenous consumption good to the price of capital. Note that the cost function has both variable and fixed components $c_v \in \mathbb{R}_+$ and $c_f \in \mathbb{R}_+$ respectively. In addition, the pair (η_1, η_2) is an i.i.d. random vector drawn from the truncated bivariate normal distribution,

(1.3.3)
$$\xi(1, 1, \sigma_1, \sigma_2, \rho, \underline{\eta}_1, \bar{\eta}_1, \underline{\eta}_2, \bar{\eta}_2) \sim \mathcal{N}\left[\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right] \text{ in } [\underline{\eta}_1, \bar{\eta}_1] \times [\underline{\eta}_2, \bar{\eta}_2]$$

[7] The estimation section will show that this way of modeling costs can approximate real-world cost dynamics reasonably well.

[8] The data analogue of the profit function is corporate earnings before interest, taxes, depreciation, and amortization (EBITDA).

with mean $(1,1)$ and where $\underline{\eta}_1 = \underline{\eta}_2 = 0 = \underline{\eta}$ are the left and bottom truncation lines and $\bar{\eta}_1 = \bar{\eta}_2 = 2 = \bar{\eta}$ are the right and top truncation lines.[9] The truncations have virtually no numerical effect and are only needed to ensure that the revenue and costs are not negative.[10] To save space on notation, I will write $\xi(\sigma_1, \sigma_2, \rho)$ for the truncated bivariate normal distribution from now on. This correlated i.i.d. shock is introduced so that the model can mimic the decrease in the correlation between revenue and operating expenses observed in the data. If the firm has low operating margins, i.e. when mean revenues and expenses are much larger than mean profit, small changes in ρ can have powerful effects on the profit volatility and hence the cash flow volatility. The fundamental assumption here is that both persistent and transitory shocks may have important implications for real world firm dynamics.[11] The magnitude of the persistent or transitory component is then determined numerically. For example, the estimation may very well discover that $\sigma_1 = \sigma_2 = 0$ which would indicate that the transitory shock is an extraneous model feature. Of course, since the transitory shock is highlighted as an important part of the model, $\sigma_1 = \sigma_2 = 0$ is not what I find.

To recapitulate, the state vector of the firm at the beginning of the period is $\{k, m, z\}$ and profit $\pi(k, z, \eta_1, \eta_2; P)$ is generated after the realization of the transitory shocks $\{\eta_1, \eta_2\}$.

The firm also faces corporate taxes where the taxable income,

$$y(k, m, z, \eta_1, \eta_2; P) = \pi(k, z, \eta_1, \eta_2; P) - \delta k + r_f m \tag{1.3.4}$$

includes depreciation and interest, and δ is the capital depreciation per unit of time and r_f is the risk-free real interest rate.[12] Therefore, the net income is,

---

[9]This bivariate normal shock is a transitory shock with a specific structure. The right and top truncations are imposed simply for the sake of symmetry.

[10]In fact, σ_1 and σ_2 are always estimated to be small enough where the probabilities of $\eta_1 = 0$, $\eta_2 = 0$, $\eta_1 = 2$, or $\eta_2 = 2$ are numerically zero.

[11]Section 1.5.2 details the identification of the persistent and transitory shocks.

[12]Compustat firms hold most of their cash in interest bearing accounts or treasuries so that there is a small positive interest rate on cash.

(1.3.5) $$n(k, m, z, \eta_1, \eta_2; P) = (1 - \phi_c \tau_c) y(k, m, z, \eta_1, \eta_2; P)$$

where τ_c is the corporate tax rate and ϕ_c is the shorthand notation for the indicator $\mathbf{1}_{\{y(\cdot)\geq 0\}}$.[13] The cash flow,

(1.3.6) $$f(k, m, z, \eta_1, \eta_2; P) = n(k, m, z, \eta_1, \eta_2; P) + \delta k$$

simply adds back depreciation. The net income along with the current and next period capital and cash choices determine the equity flow of the firm. Therefore, the period equity flow to or from shareholders if the firm chooses to continue to operate and adjusts its capital to k' and its cash holdings to m' is,

$$e_I(k, k', m, m', z, \eta_1, \eta_2; P) = (1 - \phi_d \tau_d + \phi_\lambda \lambda) \{n(\cdot) + [k - k'] + [m - m']\}$$

where τ_d is the tax rate on a positive distribution (dividends) and λ is the equity flotation cost incurred per unit of negative equity flow (equity issuance). The function ϕ_d is the shorthand notation for the indicator $\mathbf{1}_{\{n(\cdot)+[k-k']+[m-m']\geq 0\}}$ and ϕ_λ is the shorthand notation for the indicator $\mathbf{1}_{\{n(\cdot)+[k-k']+[m-m']<0\}}$. In this class of models, it is straightforward to prove that the firm would never simultaneously distribute dividends and issue equity. One of the most important features of the profit function defined above is that the profit can be negative.[14] When the firm encounters negative net income, it must tap into its cash reserve or issue equity to maintain the same level of capital in the next period. Equity issuance is costly, so therefore, cash acts as a buffer stock against both transitory and persistent shocks. Distributing dividends in the current period and then issuing equity in the next period is

[13]In previous versions of this paper, the tax function was a more complicated arctangent function to emulate the real-world tax brackets. However, the complication was numerically unimportant because Compustat firms are large enough where they essentially face the top tax bracket whenever they have positive taxable income.

[14]Most other profit functions in the structural literature are weakly positive such as $\pi(k, z) = zk^\alpha$.

particularly expensive. The shareholders have to pay the distribution tax τ_d in the current period and then pay the per unit equity issuance cost λ in the next period if this occurs. On the other hand, the firm could have just retained the earnings without incurring additional taxes and equity issuance costs. The balance between the benefit and cost of holding cash is analyzed in the section on optimal cash policy.

If the firm instead chooses to exit, the equity flow is,

$$e_X(k, m, z, \eta_1, \eta_2; P) = (1 - \tau_d) \max \{\pi(\cdot) - \phi_c \tau_c y(\cdot) + s(1-\delta)k + (1 + r_f)m, 0\}$$

where shareholders receive a positive distribution after exiting if the cash plus the proceeds from selling the capital at the fire-sale price is more than enough to offset any negative cash flow. The coefficient $s \in [0, 1)$ is the fire-sale value of capital.

Finally, the equity flow for a potential entrant that chooses initial capital of k' and initial cash of m' is,

$$e_E(k', m') = (1 + \lambda) [-k' - m'].$$

Notice that the firm is purely equity financed at entry and pays per unit equity issuance cost λ.

1.3.2. Recursive formulation. The precise bound on capital $\bar{k} = \left(\frac{\bar{\eta}\bar{z}\alpha}{1+r_f+\underline{\eta}c_v+\delta}\right)^{\frac{1}{1-\alpha}}$ can be constructed directly from the first order condition on k' by assuming that the firm will receive the best possible shocks next period. The boundedness of the cash choice must also be proven. Let $\beta = \frac{1}{1+r_f(1-\tau_i)}$ be the discount factor where τ_i is the individual tax rate. So all that is needed for cash holdings to be bounded is $\tau_c > \tau_i \implies (1+(1-\tau_c)r_f) < (1+(1-\tau_i)r_f)$ which is a maintained assumption throughout the paper. That is, cash needs to be more valuable outside the firm than inside the firm at some level of holdings. Therefore m' is selected from the compact set $[0, \bar{m}] \equiv M$.

The value function of an incumbent that continues to operate is,

$$V_I(k,m,z,\eta_1,\eta_2;P) =$$
$$\max_{k',m'}\left\{e_I(k,k',m,m',z,\eta_1,\eta_2;P)+\beta\int\int V(k',m',z',\eta_1',\eta_2';P)d\xi(\sigma_1,\sigma_2,\rho)d\Gamma(z'|z)\right\}$$

subject to,

$$e_I(k,k',m,m',z,\eta_1,\eta_2;P) = (1-\phi_d\tau_d+\phi_\lambda\lambda)\left\{n(\cdot)+[k-k']+[m-m']\right\}.$$

The value function of an incumbent that exits is,

$$V_X(k,m,z,\eta_1,\eta_2;P) = e_X(k,m,z,\eta_1,\eta_2;P)$$

subject to,

$$e_X(k,m,z,\eta_1,\eta_2;P) = (1-\tau_d)\max\left\{\pi(\cdot)-\phi_c\tau_c y(\cdot)+s(1-\delta)k+(1+r_f)m,0\right\}.$$

Therefore the value function of an incumbent is,

$$V(k,m,z,\eta_1,\eta_2;P) = \max_{x'}\{V_I(k,m,z,\eta_1,\eta_2;P),V_X(k,m,z,\eta_1,\eta_2;P)\}.$$

The next period capital, cash, and exit decision rules for the incumbent are denoted $k' = \mathcal{K}(k,m,z,\eta_1,\eta_2;P)$, $m' = \mathcal{M}(k,m,z,\eta_1,\eta_2;P)$, and $x' = \mathcal{X}(k,m,z,\eta_1,\eta_2;P) \in \{0,1\}$ respectively. The exit decision rule for the incumbent $\mathcal{X}(k,m,z,\eta_1,\eta_2;P)$ is a discrete choice in $\{0,1\}$ where $x'=0$ implies that the firm continues to operate and $x'=1$ implies that the firm exits.

The value function of a potential entrant is,

$$V_E(z;P) =$$
$$\max_{k',m',x'}\left\{e_E(k',m')+\beta\int\int V(k',m',z',\eta_1',\eta_2';P)d\xi(\sigma_1,\sigma_2,\rho)d\Gamma(z'|z),0\right\}$$

subject to,

$$e_E(k', m') = (1 + \lambda)\,[-k' - m'].$$

The next period capital, cash, and entry decision rules for the potential entrant are denoted $k' = \mathcal{K}_E(z;P)$, $m' = \mathcal{M}_E(z;P)$, and $x' = \chi_E(z;P) \in \{0,1\}$ respectively. Similarly, the entry decision rule for the potential entrant $\chi_E(z;P)$ is a discrete choice in $\{0,1\}$ where $x' = 0$ implies that the potential entrant chooses to invest in capital and cash, and $x' = 1$ implies that the potential entrant does not choose to invest in capital and cash.[15] The important assumption here is that the potential entrant determines next-period capital and cash after z is realized. The potential entrant can also discover that the expected firm value is negative after the realization of z. This causes the potential entrant to not invest in capital and cash and not enter the economy.

1.3.3. Free entry. Assume that the potential entrant receives a z draw from the invariant distribution. It does not know the value of z before becoming a potential entrant. Therefore, the free entry condition is,

$$\begin{aligned}\int (1 - \chi_E(z;P))\Big\{&(1+\lambda)[-\mathcal{K}_E(z;P) - \mathcal{M}_E(z;P)] \\ &+ \beta \int\!\!\int V(\mathcal{K}_E(z;P), \mathcal{M}_E(z;P), z', \eta_1', \eta_2'; P)\, d\xi(\sigma_1, \sigma_2, \rho)\, d\Gamma(z'|z)\Big\} d\Gamma_E(z) \le c_E \end{aligned} \tag{1.3.7}$$

where $\Gamma_E(z)$ is the invariant distribution of z and c_E is the entry cost. The left side of the inequality is the expected value of the potential entrant prior to the knowledge of z. Recall that the potential entrant has the option of choosing $\chi_E(z;P) = 1$ which implies that it does not become an operational firm since there is no investment in capital and cash. This means that a potential entrant who does not invest in capital and cash never enters the economy and disappears immediately. However, every potential entrant pays the entry cost. More precisely, the "entry cost" is the cost paid to receive the z draw since a potential entrant can

[15] To economize on notation, I use χ and χ_E to refer to the exit and entry decision rules and $\mathcal{X}(k, m, z, \eta_1, \eta_2; P) = \chi_E(z;P) = 1$ always denotes that the firm or potential entrant leaves the economy.

pay this cost and not enter. The fixed costs of production are not incurred until the period after entry but the fixed costs nonetheless discourage potential entrants to enter with low z draws.

Again, the exact value of z is learned only after entry. This assumption induces larger firms to enter with a wide range of firm sizes and is a realistic model of entry into Compustat.[16] In contrast, the standard entry condition assumed in Hopenhayn (1992) and Gomes (2001) where the shock is learned after entry would cause all firms to enter with the same cash and capital. Entering firms tend to be smaller as well under this type of entry assumption and firms with low initial shock draws would immediately exit in the next period due to strong shock persistence.[17]

1.3.4. Distribution. The distribution μ can be computed by the following equation,

$$\begin{aligned} \mu'(k', m', z') = &\int\int\int I(k, m, z, \eta_1, \eta_2; P) d\xi(\sigma_1, \sigma_2, \rho) d\Gamma(z'|z) d\mu(k, m, z) \\ &+ M' \int\int (1 - \chi_E(z; P)) \mathbf{1}_{\mathcal{K}_E(z;P)=k'} \mathbf{1}_{\mathcal{M}_E(z;P)=m'} d\Gamma(z'|z) d\Gamma_E(z) \end{aligned} \tag{1.3.8}$$

where $I(k, m, z, \eta_1, \eta_2; P) \equiv (1 - \mathcal{X}(k, m, z, \eta_1, \eta_2; P)) \mathbf{1}_{\mathcal{K}(k,m,z,\eta_1,\eta_2;P)=k'} \mathbf{1}_{\mathcal{M}(k,m,z,\eta_1,\eta_2;P)=m'}$ is a combined indicator function and M' is the mass of potential entrants every period. Another way of writing the law of motion of μ is to define an operator T^* such that

$$\mu' = T^*(\mu, M'; P). \tag{1.3.9}$$

The T^* operator maps distributions to distributions and in equilibrium, $\mu = \mu' = \mu^*$.

1.3.5. Industry demand. Assume the industry price function has the form,

[16] Newly listed Compustat firms have a similar average size and size dispersion in comparison to existing firms.
[17] Only 1% of Compustat firms under 1 billion 2010 dollars in total assets exit after 1 year.

$$P = \frac{1}{Q^d}$$

where Q_d is the quantity demanded. Therefore the demand function is,

$$Q^d = \frac{1}{P}.$$

The specific form the demand function takes is unimportant as long as $\lim_{P \to 0^+} (Q^d) = \infty$ and $\lim_{P \to \infty} (Q^d) = 0$. The total quantity supplied by the firms in the economy is,

$$Q^s = \int \int \eta_1 z k^\alpha d\xi(\sigma_1, \sigma_2, \rho) d\mu(k, m, z).$$

And so the product market clears when $Q^d = Q^s$.

1.3.6. Incumbent timing.

(1) The firm comes into the period with state vector $\{k, m, z\}$.

(2) The transitory shocks $\{\eta_1, \eta_2\}$ are realized and profit $\pi(k, z, \eta_1, \eta_2; P)$ is generated.

(3) The firm chooses whether or not to exit. If the firm exits, there is possibly one last dividend distribution. If the firm continues to operate, then $k' > 0$ and m' are chosen.

(4) Dividend is distributed or equity is issued to shareholders depending on the sign of the equity flow.

(5) The next period revenue shock z' is realized.

1.3.7. Potential entrant timing.

(1) The potential entrant draws z from the invariant distribution and pays entry cost c_E.

(2) If $x' = 1$, then the firm never invests in capital and cash and does not enter into the economy. Otherwise, the firm chooses $k' > 0$ and m' and it is purely equity financed.

(3) The next period revenue shock z' is realized.

1.3.8. Equilibrium.

Definition 1. A stationary recursive competitive industry equilibrium is: a set containing (i) value functions $V_I(k, m, z, \eta_1, \eta_2; P)$, $V_X(k, m, z, \eta_1, \eta_2; P)$, and $V_E(z; P)$, (ii) decision rules for incumbents $k' = \mathcal{K}(k, m, z, \eta_1, \eta_2; P)$, $m' = \mathcal{M}(k, m, z, \eta_1, \eta_2; P)$, and $x' = \mathcal{X}(k, m, z, \eta_1, \eta_2; P)$, (iii) decision rules for potential entrants $k' = \mathcal{K}_E(z; P)$, $m' = \mathcal{M}_E(z; P)$, and $x' = \chi_E(z; P)$, (iv) a price P, and (v) an invariant distribution μ^* such that,

(1) The decision rules solve the value functions,
(2) The free entry condition (1.3.7) is satisfied,
(3) The invariant distribution $\mu^* = \mu = \mu'$ solves (1.3.8),
(4) And the product market clears $Q^d = Q^s$.

1.4. Optimal cash policy

The intuition behind the cash decision rule is explored in this section. The value function is not everywhere differentiable due to the discrete choice of exit. However, assuming differentiability of the value function and deriving the optimal cash policy can still offer many insights. First of all, the optimal cash policy is dependent on the state of the firm and the marginal value of a unit of cash to an incumbent that will continue to operate in the next period is,

$$\frac{\partial V(k, m, z, \eta_1, \eta_2; P)}{\partial m} = (1 - \phi_d \tau_d + \phi_\lambda \lambda)(1 + r_f - \phi_c \tau_c r_f). \tag{1.4.1}$$

Therefore, the marginal value of cash can vary greatly and cash is more valuable if the firm issues equity than if the firm pays out dividends. In fact, there are six different values of current-period cash. Cash is the most valuable with marginal value $(1 + \lambda)(1 + r_f)$ when the firm has negative taxable income and issues equity. As long as $(1 + \lambda)(1 + (1 - \tau_c)r_f) > (1 + r_f)$,[18] the next most valuable state for cash occurs when the firm has positive taxable

[18]This inequality holds for all of the parameterizations in the paper.

income and issues equity, and the marginal value is $(1+\lambda)(1+(1-\tau_c)r_f)$. Then, the third (fourth) most valuable state for cash occurs when the firm has negative (positive) taxable income and retains all earnings, and the marginal value is $(1+r_f)$ or $(1+(1-\tau_c)r_f)$ respectively. Finally, cash is the least valuable with marginal value $(1-\tau_d)(1+r_f)$ or $(1-\tau_d)(1+(1-\tau_c)r_f)$ when the firm has negative (positive) taxable income and distributes dividends. This ordering is consistent with intuition and generates a wide range of values for current-period cash.

The marginal value of a unit of next-period cash depends on the probability of ending up in the states just described. The first order condition with respect to m' is,

$$(1 - \phi_d \tau_d + \phi_\lambda \lambda) = \beta \int \int \frac{\partial V(k', m', z', \eta_1', \eta_2'; P)}{\partial m'} d\xi(\sigma_1, \sigma_2, \rho) \partial \Gamma(z'|z). \tag{1.4.2}$$

Plugging in envelope condition 1.4.1 gives,

$$(1 - \phi_d \tau_d + \phi_\lambda \lambda) = \beta \int \int (1 - \phi_d' \tau_d + \phi_\lambda' \lambda)(1 + r_f - \phi_c' \tau_c r_f) d\xi(\sigma_1, \sigma_2, \rho) \partial \Gamma(z'|z). \tag{1.4.3}$$

The left side of Equation 1.4.3 is the marginal value of shareholder distributions, retained earnings, or external finance while the right side of the equation is the shadow value of next-period cash. If the cash would otherwise be distributed, then the expected value of next-period cash only requires marginal value $(1-\tau_d)$. On the other hand, if the firm retains all earnings or needs equity, the expected values of next-period cash require higher marginal values of 1 or $(1+\lambda)$ respectively. If the firm needs equity, cash is very valuable today and it must be just as valuable tomorrow in expectation for the firm to hold on to the amount of cash given by m'.

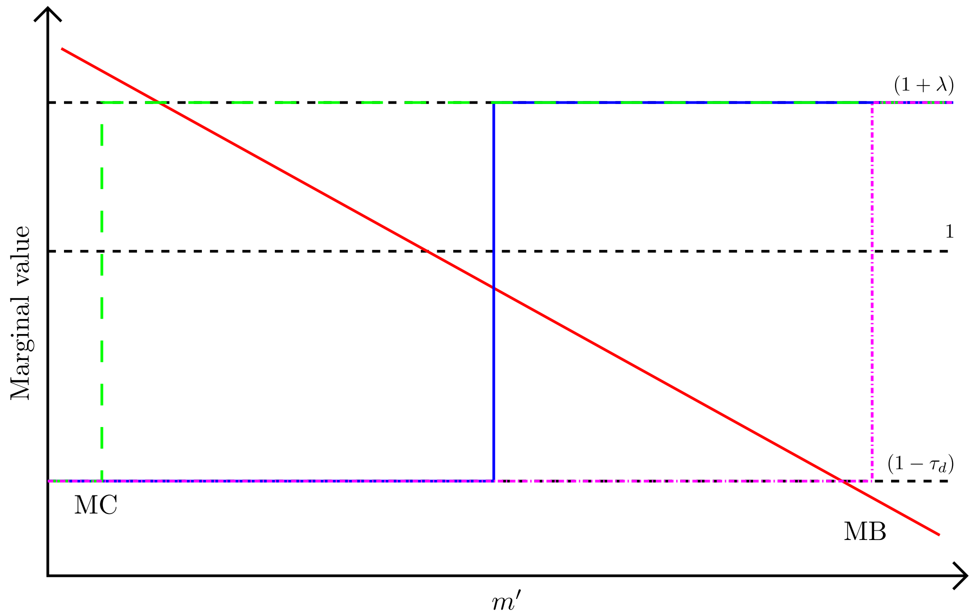


FIGURE 1.4.1. This figure graphs the optimal cash policy for different marginal cost functions.

Figure 1.4.1 illustrates the optimal cash policy given three different marginal cost functions of next-period cash. It can be seen that next-period cash holdings increase as the marginal cost decreases. The vertical lines indicate the value of m' where the firm switches from paying out dividends to issuing equity. The marginal benefit function can intersect the marginal cost functions at the horizontal lines or at the vertical line. When the intersection is at the horizontal lines, marginal benefit equals marginal cost and must have value $(1-\tau_d)$ or $(1+\lambda)$. In contrast, when the intersection is at the vertical line, the firm is in an inaction region where all earnings are retained and the marginal benefit of next-period cash can be anything between $(1-\tau_d)$ and $(1+\lambda)$. That is, the firm neither finds it worthwhile to distribute dividends nor issue equity.

1.5. Results

1.5.1. Parameterization. The z shock used for the estimated model is assumed to follow an AR(1) process in logs, i.e.

$$\log(z') = \phi \log(z) + \theta + \epsilon' \tag{1.5.1}$$

where $\phi \in (0,1)$, $\theta \in \mathbb{R}$, and $\epsilon' \sim N(0, \sigma_\epsilon^2)$. Some parameters are first set to values taken directly from the data or from the related literature as shown in Table 1. The risk-free real interest rate is found by using the 3 month treasury rate minus the rate of inflation and then averaged for the 1980-1984 period. The depreciation rate is set at the value found in Cooper and Haltiwanger (2006). The fire-sale value of capital and distribution tax rate are set to the values used in Hennessy and Whited (2005) and both are within the range commonly used in the literature. Finally, the top marginal U.S. corporate tax rate was 46% for the entire 5 year period and the price is normalized to 1.[19]

Outside parameters (1980-1984)		Value
r_f	Risk-free real interest rate	0.05
δ	Depreciation rate	0.069
s	Fire-sale value of capital	0.75
τ_i	Individual tax rate	0.296
τ_d	Distribution tax rate	0.12
τ_c	Corporate tax rate	0.46
P	Price	1

TABLE 1. This table lists the parameters taken from outside the model corresponding to the 1980-1984 time period.

1.5.2. Identification. There are 10 parameters that need to be estimated in the model, namely, the revenue returns to scale α, AR(1) in logs scale parameter θ, AR(1) in logs persistence parameter ϕ, AR(1) in logs standard deviation parameter σ_ϵ, standard deviations and correlation of the bivariate normal shock $\{\sigma_1, \sigma_2, \rho\}$, production costs $\{c_v, c_f\}$, and per unit equity issuance cost λ.

[19]This is a convenient trick since entry cost is assumed to be unobservable. Later on, entry cost will be set to the value found when $P = 1$. The important element is the change in price to satisfy the free entry condition and not the original normalization.

The identification is relatively straightforward for several initial parameters. The revenue returns to scale α is identified by the standard deviation of capital. The AR(1) in logs scale parameter θ is identified by mean revenue while the variable cost parameter c_v is identified by mean operating expenses. Finally, the fixed cost c_f is identified by the exit rate and the equity floatation cost λ is identified by mean equity issuance.

Revenue can be decomposed into $\log(\text{rev}) = \log(P\eta_1 z k^{\alpha}) = \log(P) + \log(\eta_1) + \log(z) + \alpha\log(k)$ and variable operating expenses can be decomposed into $\log(\text{vxp}) = \log(\eta_2 c_v k) = \log(\eta_2) + \log(c_v) + \log(k)$. Taking the appropriate variances and covariances of the revenue and variable operating expenses produces Table 2. With some algebra, it can be shown that the moments listed in the table can identify both persistent and transitory shock parameters.

Moment	Components	Row
$\text{cov}(\tilde{\text{rev}}_t, \tilde{\text{rev}}_{t-1})$	$\text{cov}(\tilde{z}_t, \tilde{z}_{t-1}) + \alpha\text{cov}(\tilde{k}_{t-1}, \tilde{z}_t) + \alpha\text{cov}(\tilde{k}_t, \tilde{z}_{t-1}) + \alpha^2\text{cov}(\tilde{k}_t, \tilde{k}_{t-1})$	1
$\text{cov}(\tilde{\text{vxp}}_t, \tilde{\text{vxp}}_{t-1})$	$\text{cov}(\tilde{k}_t, \tilde{k}_{t-1})$	2
$\text{cov}(\tilde{\text{rev}}_{t-1}, \tilde{\text{vxp}}_t)$	$\text{cov}(\tilde{k}_t, \tilde{z}_{t-1}) + \alpha\text{cov}(\tilde{k}_t, \tilde{k}_{t-1})$	3
$\text{cov}(\tilde{\text{rev}}_t, \tilde{\text{vxp}}_{t-1})$	$\text{cov}(\tilde{k}_{t-1}, \tilde{z}_t) + \alpha\text{cov}(\tilde{k}_t, \tilde{k}_{t-1})$	4
$\text{var}(\tilde{\text{rev}}_t)$	$\text{var}(\tilde{\eta}_{1,t}) + \text{var}(\tilde{z}_t) + \alpha^2\text{var}(\tilde{k}_t)$	5
$\text{var}(\tilde{\text{vxp}}_t)$	$\text{var}(\tilde{\eta}_{2,t}) + \text{var}(\tilde{k}_t)$	6
$\text{cov}(\tilde{\text{rev}}_t, \tilde{\text{vxp}}_t)$	$\text{cov}(\tilde{\eta}_{1,t}, \tilde{\eta}_{2,t}) + \text{cov}(\tilde{k}_t, \tilde{z}_t) + \alpha\text{var}(\tilde{k}_t)$	7

Table 2. This table decomposes all the variances and covariances needed for the identification of persistent and transitory shock parameters. The tildes indicate log variables.

First, the variance and autocovariance of the AR(1) in logs shock process can also be written as $\text{var}(\tilde{z}_t) = \frac{\sigma_\epsilon^2}{1-\phi^2}$ and $\text{cov}(\tilde{z}_t, \tilde{z}_{t-1}) = \phi\frac{\sigma_\epsilon^2}{1-\phi^2}$ respectively. The covariances $\text{cov}(\tilde{k}_t, \tilde{z}_{t-1})$ and $\text{cov}(\tilde{k}_{t-1}, \tilde{z}_t)$ in turn can be found by subtracting α times Row 2 from Row 3 and Row 4. Note that $\text{cov}(\tilde{k}_{t-1}, \tilde{z}_t) = \text{cov}(\tilde{k}_t, \tilde{z}_{t+1})$ and the persistent shock $\tilde{z}_{t+1} = \phi^2\tilde{z}_{t-1} + (\phi + 1)\theta + \phi\epsilon_t + \epsilon_{t+1}$ can be rewritten by direct iteration. Therefore, $\text{cov}(\tilde{k}_{t-1}, \tilde{z}_t) = \phi\text{cov}(\tilde{k}_t, \tilde{z}_t) = \phi^2\text{cov}(\tilde{k}_t, \tilde{z}_{t-1})$ because the choice of k_t does not depend on ϵ_t or ϵ_{t+1}. Row 1 then pins down σ_ϵ, after the autocovariance of the shock process is expressed in terms of ϕ and σ_ϵ and the latter three terms are eliminated using Rows 2 through 4. The identification of σ_1, σ_2, and

ρ in the end comes from Rows 5 through 7 respectively since the terms not related to the bivariate normal shock are terms which have already been ascertained.

While the clean identification strategy outlined above may seem to suggest that the model does not need to be fully solved to find the parameters, in reality, each parameter has effects on multiple moments and everything is jointly determined. Identification simply comes from the fact that each parameter has stronger effects on certain moments.

1.5.3. Data. The data source is the Compustat North America Fundamentals Annual from 1980 to 2010. The focus of the paper is on industrial firms, and therefore regulated firms with Standard Industrial Classification (SIC) codes between 4,900 and 4,999 and financial firms with SIC codes between 6,000 and 6,999 are dropped. In addition, firms with under 10 million 2010 U.S. dollars in total assets and firms with missing or negative revenue, operating expenses, or assets are dropped from the sample. Firms with missing income, and firms with negative cash or capital are also dropped. At the end, missing cash and capital values are replaced with zero. There are a total of 129,507 firm-year observations remaining after the data is processed. The small firms which are the focus of this paper comprise 76.3% of the sample.

The moments used in the estimation are then normalized by the mean total assets A_T of small firms in the cross-section for each year. This form of normalization preserves the relative magnitudes of the variables while removing the real growth trend. A detailed analysis of the normalization is presented in the appendix.

I compute the cross-sectional statistics of the moments used in the identification by considering each firm-year as a data point. For example, a firm that was in Compustat for the first 4 years of the sample would contribute 4 data points. The cross-sectional covariances are also computed in this way where each firm-year revenue-expense pair is a data point. In the model, there is only one type of firm and the cross-sectional distribution is the same as if a single firm is simulated for a large number of periods.

Finally, the data definition of revenue, operating expenses, cash flow, capital, cash, and equity issuance are the REVT, XOPR, (IB + DP), PPENT, CHE, and (SSTK − PRSTKC) variables in Compustat respectively. Equity issuance is defined to be equity issuance net of repurchases (SSTK − PRSTKC) and cash flow is defined to be income before extraordinary items plus depreciation (IB + DP) as used in Riddick and Whited (2009).[20]

1.5.4. Estimation. The estimated parameters and moments matched are presented in Table 3 and Table 4. The parameters are quite reasonable overall. For example, the returns to scale parameter α is close to 1 since there is no labor in the model and low values of α generate counterfactually compressed distributions. There is also high persistence ϕ in the AR(1) process which then requires the scale and standard deviation parameters θ and σ_ϵ to be low. The bivariate normal i.i.d. shock estimate finds nonzero values for σ_1 and σ_2 which indicates that the transitory shock is important to matching the variance and covariance moments. In particular, σ_1 and σ_2 are needed to match the broad distribution of revenue and operating expenses respectively. The variable cost parameter c_v encapsulates many costs such as production costs, research and development costs, and selling and administrative expenses. Therefore c_v is estimated to be greater than 1. There is always some fear that the equity issuance cost must be unreasonably high to match the cash level in this class of models. Fortunately, λ is determined to have a sensible value of 4.22% which is within the range commonly found in the literature.

[20] Extraordinary items do not actually contribute much to the idiosyncratic cash flow volatility of firms (the difference of the means is less than 1% and the difference of the standard deviations is less than 2% when extraordinary items are included). I used the definition of cash flow from Riddick and Whited (2009) which was (IB + DP) but the moments would almost be the same if I used net income plus depreciation (NI + DP) instead.

Inside parameters (1980-1984)		Estimate	Std Error
α	Revenue returns to scale	0.939	0.0003
θ	AR(1) in logs scale parameter	0.0175	0.0024
ϕ	AR(1) in logs persistence parameter	0.984	0.0001
σ_ϵ	AR(1) in logs standard deviation parameter	0.0459	0.0075
σ_1	Standard deviation of bivariate shock on revenue	0.238	0.0102
σ_2	Standard deviation of bivariate shock on operating expenses	0.218	0.0077
ρ	Correlation of bivariate shock	0.967	0.0055
c_v	Variable cost	3.735	0.0014
c_f	Fixed cost	0.0146	0.0204
λ	Equity floatation cost	0.0422	0.0833
c_E	Entry cost	0.103	-

Table 3. This table lists the parameters estimated using the model corresponding to the 1980-1984 time period.

Overall, the moments are matched very closely and the only moment that is off by more than 5% is mean equity issuance. It is especially reassuring that the variances and covariances are matched well since the transitory shock and its identification are central to the results in this paper. Mean equity issuance is hard to match due to complex interactions. While the mean equity issuance is sensitive to a lowering of the per unit equity issuance cost λ, other moments such as the exit rate are also somewhat sensitive to changes in λ.

Moments (1980-1984)	Data	Model
Revenue mean	1.51	1.51
Revenue standard deviation	2.36	2.42
Operating expenses mean	1.37	1.37
Operating expenses standard deviation	2.21	2.16
Cash flow mean	0.090	0.086
Cash flow standard deviation	0.159	0.161
Capital mean	0.367	0.363
Capital standard deviation	0.567	0.562
Cash mean	0.0988	0.0983
Revenue - operating expenses covariance	5.20	5.22
Revenue autocovariance	5.19	5.42
Operating expenses autocovariance	4.58	4.36
Revenue - operating expenses$_{-1}$ covariance	4.86	4.83
Revenue$_{-1}$ - operating expenses covariance	4.87	4.89
Equity issuance mean	0.021	0.006
Exit rate	0.05	0.051

TABLE 4. This table lists the data moments from the 1980-1984 time period and the model moments which attempt to match them.

1.5.5. Decision rules. The equity flows implied by the capital and cash decision rules are graphed in Figure 1.5.1, Figure 1.5.2, and Figure 1.5.3. First, the equity flow is plotted along the capital and cash dimensions in Figure 1.5.1. For low values of capital and cash, the firm will issue equity, for medium values of capital and cash the firm will retain all earnings, and for high values of capital and cash the firm will distribute dividends. Next, along the persistent shock and capital dimensions, the firm will exit for low shock and capital values as seen in Figure 1.5.2. The empty locations on the surface plot are where the firm is better off exiting the economy. The most interesting thing about this graph is that dividend distributions peak around the middle persistent shock value. The reason is that, as the shock becomes higher, there is also the tendency for the firm to invest more. In this case, the investment propensity dominates the dividend distribution propensity for high values of z. Finally in Figure 1.5.3, the equity flow behavior along the η_1 and η_2 dimensions is very intuitive. A high (low) revenue transitory shock combined with a low (high) operating expenses transitory shock induce firms to distribute dividends (issue equity), while similar transitory shock values form the inaction region.

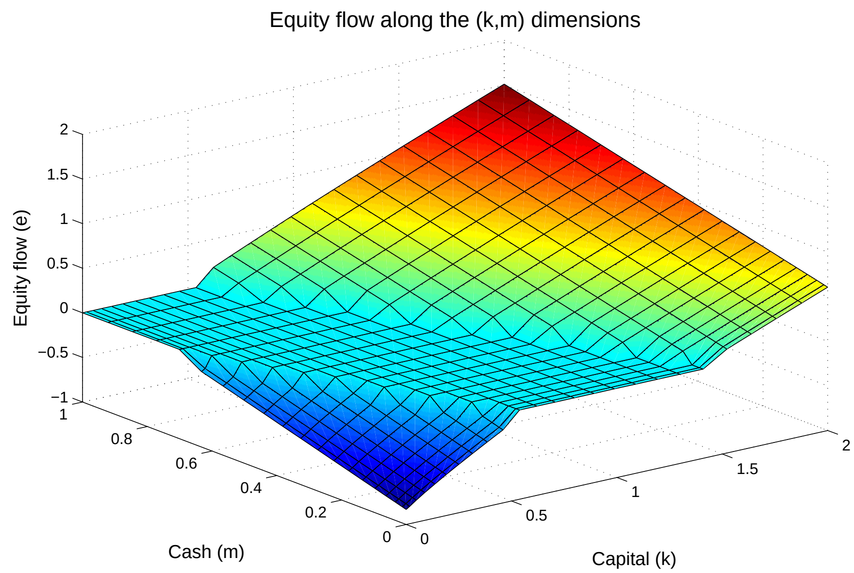


FIGURE 1.5.1. This figure graphs the equity flow along the (k, m) dimensions.

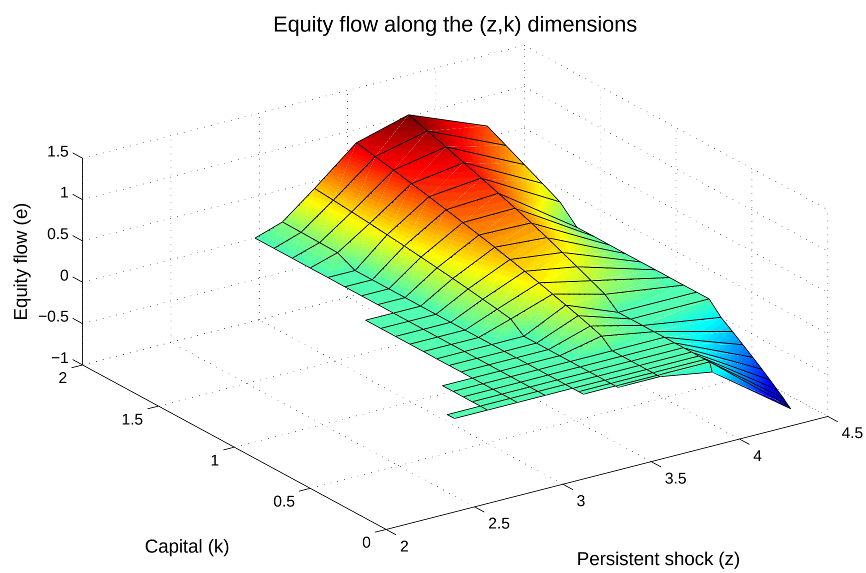


FIGURE 1.5.2. This figure graphs the equity flow along the (z, k) dimensions.

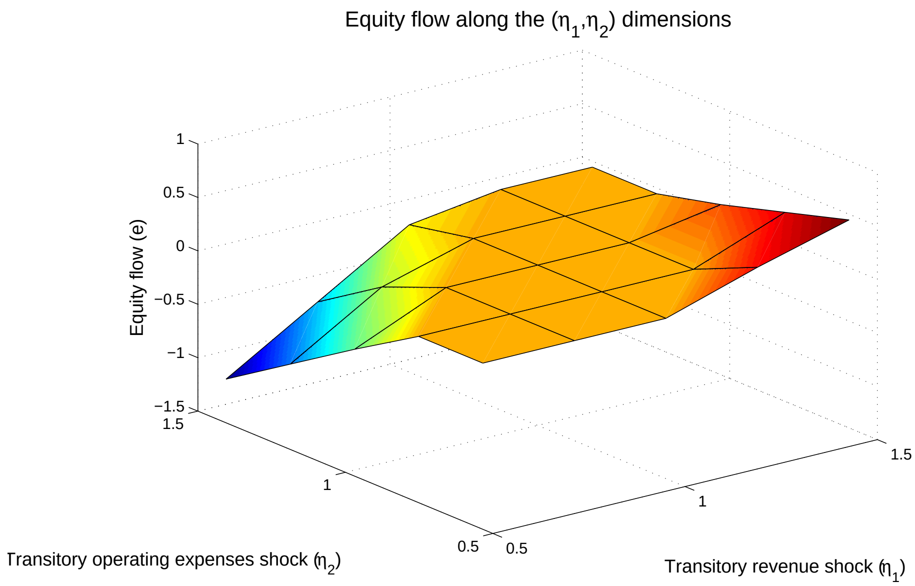


FIGURE 1.5.3. This figure graphs the equity flow along the (η_1, η_2) dimensions.

The investment policies $[k' - \delta(1-k)]$ are then graphed in Figure 1.5.4, Figure 1.5.5, and Figure 1.5.6. Note that investment does not depend much on cash for low or high values of current capital as seen in Figure 1.5.4. However, investment increases with cash for moderate amounts of current capital. On the other hand, along the persistent shock and capital dimensions in Figure 1.5.5, investment monotonically decreases with current capital. Since the conditional distribution of the AR(1) process in logs is lognormal and has a fat right tail, investment is considerably higher for the high values of z. Figure 1.5.6 demonstrates that the investment behavior along the transitory shock dimensions conforms well with intuition. Investment increases (decreases) with a high (low) η_1 shock and a low (high) η_2 shock.

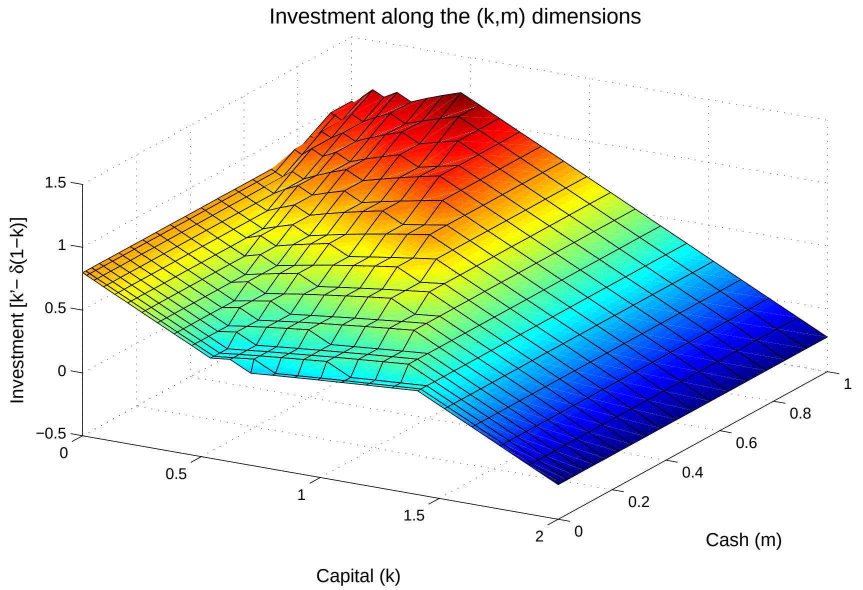


FIGURE 1.5.4. This figure graphs investment along the (k, m) dimensions.

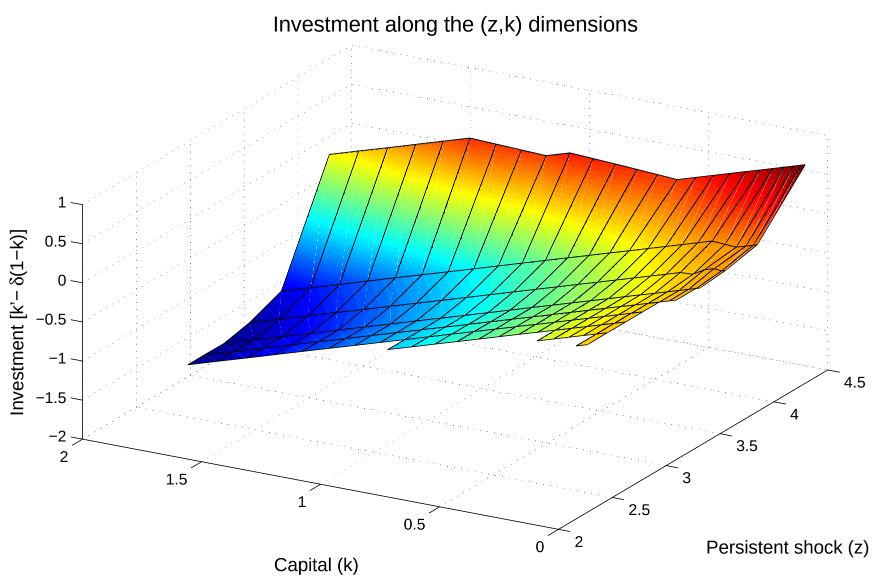


FIGURE 1.5.5. This figure graphs investment along the (z, k) dimensions.

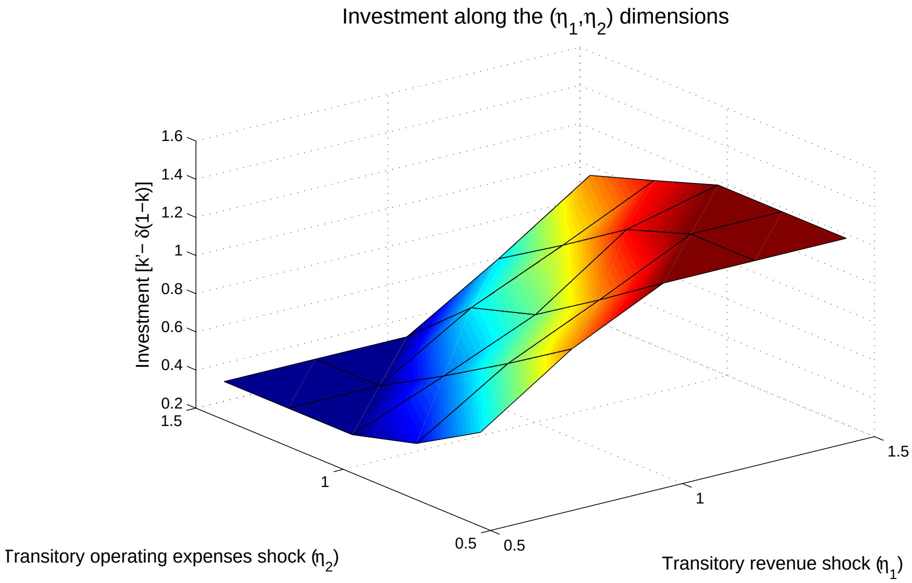


FIGURE 1.5.6. This figure graphs investment along the (η_1, η_2) dimensions.

Finally, the Compustat capital and cash distributions are juxtaposed with the model capital and cash distributions in Figure 1.5.7 and Figure 1.5.8. The general shapes of the Compustat distributions are captured nicely but of course the discreteness in the model does not allow for such a smooth decrease in proportion. In particular, the model capital distribution has a mass point just past 1.6 which the firms with the highest z value tend to choose. Also, a small fraction of the Compustat distributions actually extend out beyond the plotted histograms since the data contains an extremely diverse set of firms. The model capital and cash grids in contrast are set so that no firms are at the right endpoints.

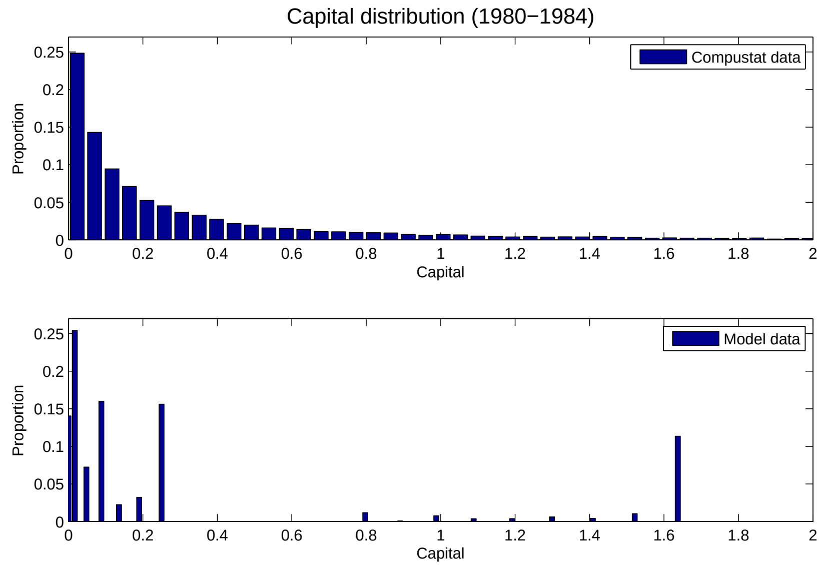


FIGURE 1.5.7. This figure compares the Compustat and model capital distributions.

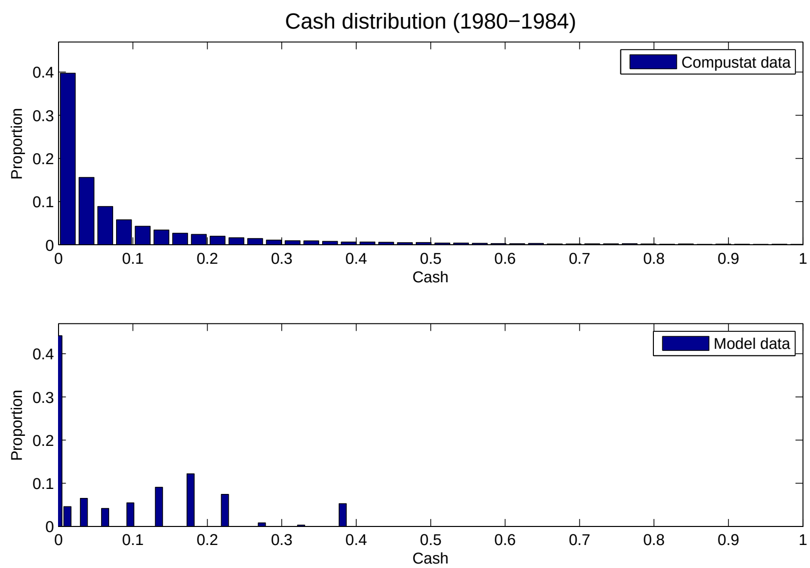


FIGURE 1.5.8. This figure compares the Compustat and model cash distributions.

1.5.6. Correlation decrease. The main experiment in this paper is performed in Table 5. It should be emphasized that the moments in the table are steady state moments and the model contains no aggregate shocks. Only the correlation ρ is decreased from the estimated value of 0.967 to a lower number. The parameter ρ is decreased to 0.862 so that the cash flow volatility in the model is matched exactly to the volatility observed in the data. The results are very good. In fact, almost every moment moves in the correct direction. However, while the standard deviation of revenue and operating expenses and the various covariances correctly move downward, these moments are significantly higher in the model experiment than in the data. The mean cash flow generated by the model is also somewhat higher than the mean cash flow observed in the data. These moments behave in this fashion because P adjusts upward to 1.014 in equilibrium to clear the goods market and the average quantity produced by each firm drops by 12%. On the other hand, if there were no equilibrium responses and the price was kept at 1, the standard deviation of revenue and operating expenses, the covariances, and the mean cash flow would be lower and closer to the data. The entry/exit rates in the model would be higher as well.

Moments (2006-2010)	Data	$\rho = 0.862$
Revenue mean	1.05	1.34
Revenue standard deviation	1.55	2.20
Operating expenses mean	0.95	1.20
Operating expenses standard deviation	1.44	1.94
Cash flow mean	0.043	0.081
Cash flow standard deviation	**0.241**	**0.241**
Capital mean	0.257	0.318
Capital standard deviation	0.447	0.504
Cash mean	0.2204	0.1750
Revenue - operating expenses covariance	2.21	4.23
Revenue autocovariance	2.04	4.50
Operating expenses autocovariance	1.76	3.45
Revenue - operating expenses$_{-1}$ covariance	1.87	3.90
Revenue$_{-1}$ - operating expenses covariance	1.89	3.98
Equity issuance mean	0.021	0.008
Exit rate	0.07	0.047

TABLE 5. This table presents the correlation decrease experiment.

Again, the fact that my model is an industry equilibrium means that there are equilibrium responses of relative prices and the firm size distribution to the change in the stochastic process. I find that the relative price of capital (the inverse of P) falls by 1.4%, entry/exit falls by 9%, and that firm size, as measured by the mean capital, decreases by 12%, while the coefficient of variation of capital, rises by 2.3%.

Table 6 then summarizes the cash increase resulting from the correlation decrease experiment. When ρ decreases from 0.967 to 0.862, 63% of the increase in cash in the last 30 years can be accounted for. The reasonable performance of the other moments acts as a test of the model and provides substantial confidence in the validity of the correlation mechanism.

Statistic	Data	$\rho = 0.862$
Cash in 1980-1984	0.0988	0.0983
Cash in 2006-2010	0.2204	0.1750
Percentage increase	123%	78%
Percentage accounted for	-	63%

Table 6. This table summarizes the behavior of cash from the correlation decrease experiment.

Table 7 shows that the decrease in correlation reduces investment and firm size where firm size is measured by the amount of capital holdings. Both investment and firm size drop because the increased volatility induces firms to substitute cash for capital for precautionary reasons. Cash flow and equity flow also decrease due to the reduction in firm size, and the coefficient of variation of size increases since volatility increases.

The first moment that behaves somewhat counterintuitively is firm value which remains roughly the same when the correlation decreases. The increased volatility increases firm value for struggling firms at the margin (low assets and/or shocks) since equity holders are residual claimants in good states of the world and have limited liability in bad states of the world. On the other hand, the increased volatility decreases firm value for decently performing firms (medium assets and/or shocks) since higher volatility just increases the chance that they will need costly external finance. However, the increased volatility increases firm value for very successful firms (high assets and/or shocks) since the best firms are even better now. For example, the firms in the bottom size tercile experience a 2.6% rise in value, the firms in the middle size tercile experience a 9.3% fall in value, and the firms in the top size tercile experience a 4.4% rise in value on average. The effect of volatility on mean firm value is mostly neutral once the price of the consumption good relative to the price of capital adjusts upward to clear the goods market. If there is no equilibrium response and the price does not adjust upward, the value drops for firms in any state, although, some firms are still more affected than others.

Finally, the entry/exit rate becomes lower which is also a bit counterintuitive. Again, the price rises in order to clear the goods market. This benefits the firms operating in the economy, and while the entry/exit rate is lower now, the firms that exit are worse than before. More precisely, the firms that choose to exit have lower expected value if they are forced (counterfactually) to stay in the economy when volatility is higher. Potential entrants in contrast have the same expected discounted value since the free entry condition must be satisfied.

	1980-1984 ($\rho = 0.967$)	2006-2010 ($\rho = 0.862$)	Percent change
Cash (m)	0.0983	0.1750	78.0%
Size (k)	0.3628	0.3177	-12.4%
CV of size	1.5499	1.5855	2.3%
Investment ($k' - (1-\delta)k$)	0.0206	0.0178	-13.6%
Cash flow (f)	0.0857	0.0812	-5.3%
Equity flow (e_I)	0.0565	0.0545	-3.5%
Price (P)	1.0000	1.0137	1.4%
Value (V_I)	1.0068	1.0077	0.1%
Entry/exit rate	0.0511	0.0465	-9.0%

Table 7. This table highlights the differences in various other important moments for high and low correlation economies.

Recall that positive equity flow is the same as dividend distribution and negative equity flow is the same as equity issuance. Table 8 tracks the change in average dividend distribution and equity issuance in an economy with low and high volatility. I break down the change for all firms, below median size firms, and above median size firms. Overall, firms distribute less dividends and issue more equity when the volatility rises.

The response across firm sizes is quite different however. For firms below (above) median size, the mean dividend distribution falls (rises) significantly. On the other hand, equity issuance increases for both firm size categories but increases more for large firms.

	1980-1984 ($\rho = 0.967$)	2006-2010 ($\rho = 0.862$)	Percent change
Dividend (all firms)	0.0638	0.0636	-0.3%
Dividend (below median)	0.0139	0.0129	-7.2%
Dividend (above median)	0.0943	0.1042	10.5%
Equity (all firms)	0.0055	0.0076	38.2%
Equity (below median)	0.0047	0.0056	19.1%
Equity (above median)	0.0060	0.0092	53.3%

TABLE 8. This table highlights the differences in the dividend distribution and equity issuance policies for high and low correlation economies.

1.5.7. Revenue volatility increase. Given the results in the previous subsections, one might ask, why shouldn't a revenue volatility increase be used to increase the cash flow volatility? In particular, what is the advantage of decomposing revenue and operating expenses? The intuition of revenue acting as a natural hedge for operating expenses was already outlined in earlier sections. Table 9 then addresses the numerical concerns. By just increasing the revenue volatility with a mean preserving spread on z to match the cash flow volatility increase, most of the moments are shown to be counterfactual. Some moments in fact are wildly counterfactual such as the standard deviations and covariances. For the mean preserving spread, each shock z is transformed to $\hat{z} = (1+\omega)z - \omega\bar{z}$ where $\omega \geq -1$ is the spread parameter and $\bar{z}$ is the mean of z. The ω needed to obtain the desired level of cash flow volatility is 0.85 and the implied equilibrium price is 0.820 in this experiment.

Cash does increase a small amount to 0.116, but clearly, achieving an increase in cash flow volatility with a revenue volatility increase is both counterfactual and dampening. The cash increase in the $\rho = 0.862$ correlation decrease experiment is more than quadruple the cash increase in the revenue volatility increase experiment.

Moments (2006-2010)	Data	$\omega = 0.85$
Revenue mean	1.05	1.93
Revenue standard deviation	1.55	3.33
Operating expenses mean	0.95	1.75
Operating expenses standard deviation	1.44	2.97
Cash flow mean	0.043	0.111
Cash flow standard deviation	**0.241**	**0.241**
Capital mean	0.257	0.464
Capital standard deviation	0.447	0.773
Cash mean	0.2204	0.1157
Revenue - operating expenses covariance	2.21	9.87
Revenue autocovariance	2.04	10.40
Operating expenses autocovariance	1.76	8.26
Revenue - operating expenses$_{-1}$ covariance	1.87	9.19
Revenue$_{-1}$ - operating expenses covariance	1.89	9.35
Equity issuance mean	0.021	0.011
Exit rate	0.07	0.068

TABLE 9. This table demonstrates the counterfactual and dampening nature of a cash flow volatility increase through an increase in revenue volatility.

1.5.8. Transition simulation. The cash transition simulation is plotted in Figure 1.5.9 and the price along the transition path is plotted in Figure 1.5.10. The simulation uses backwards induction from 2010 assuming a linear decrease of ρ from 0.967 to 0.862 over the last 30 years. The model economy experiences a steadier and more tempered increase in cash than the real world economy. Also, since the price P has risen, the relative price of capital has declined in the last 30 years.

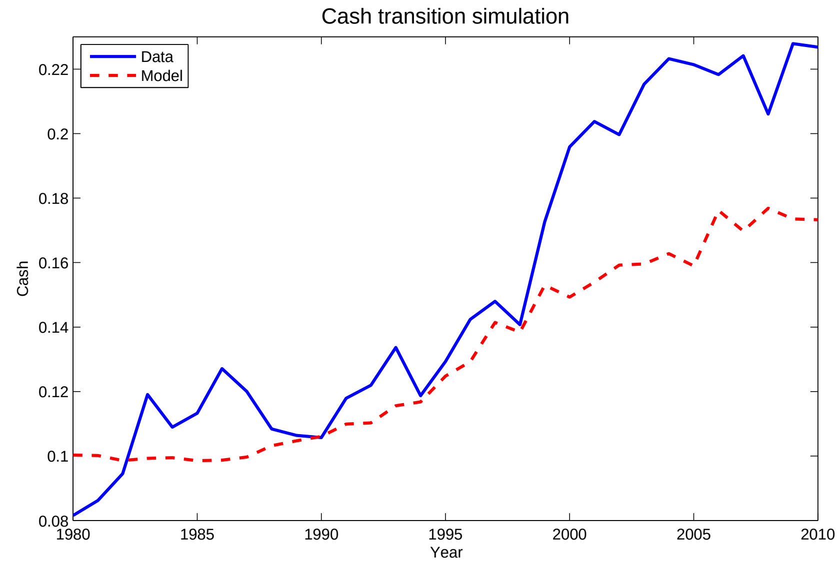


FIGURE 1.5.9. This figure plots the cash transition simulation.

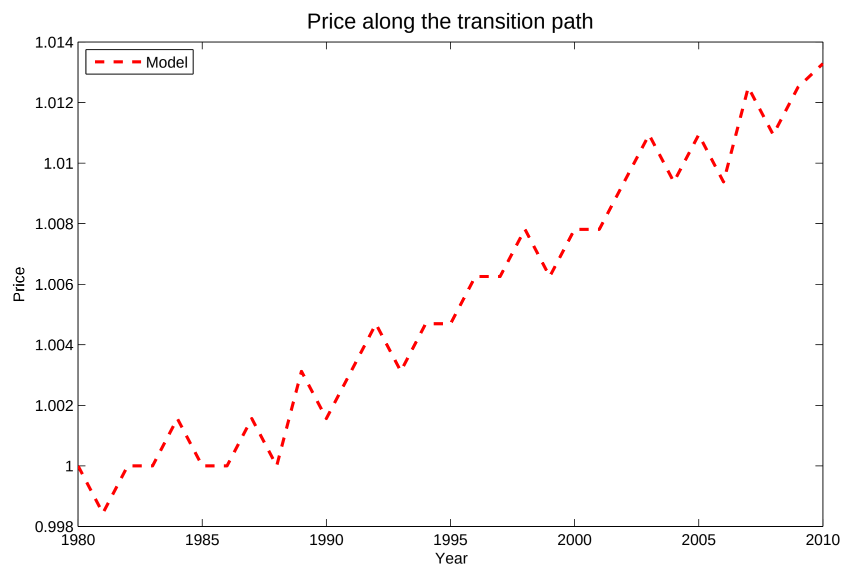


FIGURE 1.5.10. This figure plots the price along the transition path.

1.5.9. Regressions. Model data can then be produced to imitate Compustat data by using the transition computed in Section 1.5.8. A simplified regression using the Bates et al. (2009) regressors which have a model analogue can be performed on the Compustat data and on the model data. The results are detailed in Table 10. First note that all the signs of the coefficients are the same. The coefficients for cash flow volatility are also of similar magnitude for all three regressions. However, the coefficients of the other regressors have larger magnitudes for the model regressions. Since the model is a parsimonious description of the real world, these regressors contain more information about the dynamics in the model than the dynamics in the real world. The R^2 values of the model regressions are consequently larger than the R^2 value of the Compustat regression.

Cash	Data	Model	Model w/ ρ
Cash flow volatility	0.188***	0.191***	0.142***
Capital expenditure	-0.045***	-0.232***	-0.232***
Dividend dummy	0.017***	0.100***	0.102***
Market value	0.047***	0.009***	0.012***
ρ	-	-	-0.903***
R^2	0.276	0.258	0.296

TABLE 10. This table presents a simplified Bates et al. (2009) regression on the Compustat data and on the model generated data for the correlation decrease transition experiment. The last column includes unobservable ρ as a regressor for the regression on the model data. All coefficients are significant at the 1% level.

In the Compustat data, the cash flow volatility in 1980-1984 is 0.078 and in 2006-2010 is 0.103. So the regression using the Compustat data predicts a 0.5% increase in cash holdings over the last 30 years. Similarly, in the model data for the $\rho = 0.862$ transition experiment, the cash flow volatility is 0.083 in the first 5 year period and is 0.121 in the last 5 year period. So the regression using the model data predicts a 0.7% increase in cash holdings over the last 30 years. However, it known that *only* ρ is changed in the model from 0.967 to 0.862 and this change in ρ then increases cash flow volatility which ultimately generates the increase in cash. Therefore the regressions severely underpredict the contribution of cash flow volatility to the increase in cash. Simultaneity bias is the specific endogeneity issue at play here. An increase in cash flow volatility also increases capital expenditure and market value. Therefore the regression is picking up these effects as well even though the increase in cash flow volatility is the true source of causation. If unobservable ρ is added into the model regression, the regression predicts a 9.5% increase in cash holdings which then accounts for 78% of the total cash increase!

1.5.10. Real interest rate and corporate taxes. During the 30 year time period, the real interest rate has also decreased substantially as illustrated in Figure 1.5.11. A decrease in the real interest rate increases the discount rate from $\beta = \frac{1}{1+r_f(1-\tau_i)}$ to $\hat{\beta} = \frac{1}{1+\hat{r}_f(1-\tau_i)}$. Recall that the real return on cash is also pegged to r_f which means that $\frac{1}{1+r_f(1-\tau_i)}(1+r_f) > \frac{1}{1+\hat{r}_f(1-\tau_i)}(1+\hat{r}_f)$ if $r_f > \hat{r}_f$, i.e. the marginal benefit of cash decreases when it is isolated from the rest of the model dynamics. However firms actually tend to hold more cash when the real interest rate decreases because they also place a higher weight on the future cost of equity issuance.

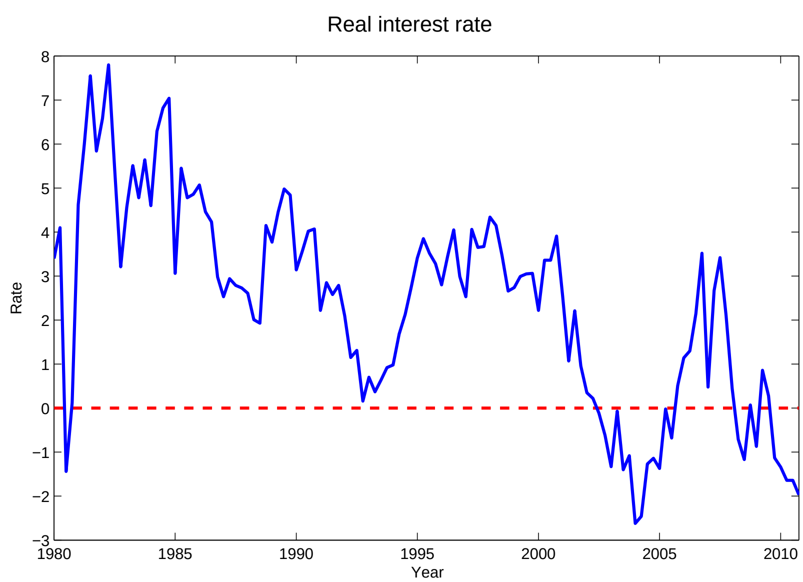


FIGURE 1.5.11. This figure plots the real interest rate over the last 30 years.

Table 13 was obtained by performing an estimation on the last 5 year period where r_f is decreased from 0.05 to 0.034, τ_i is decreased from 0.296 to 0.25, τ_c is decreased from 0.46 to 0.35, and c_E is kept at 0.103. Note that the interest rate was lowered to 3.4% instead of to the mean value observed in the last 5 years of the sample period. Compustat firms actually report the expectation of the risk-free rate and the mean expectation is 3.4% in the last 5 year period. This expectation is significantly higher than the observed rate for 2006-2010. However, the expectation is arguably a better approximation of the return and discount rate used in the firm decision. In the model, the realized return on cash for a few periods has very little numerical significance while the expectation on the future return and discount rate is very important. Unfortunately, only starting in 2002 was the expectation tracked in Compustat and so the realized real interest rate had to be used for the 1980-1984 period.

Outside parameters (2006-2010)		Value
r_f	Risk-free real interest rate	0.034
δ	Depreciation rate	0.069
s	Fire-sale value of capital	0.75
τ_i	Individual tax rate	0.25
τ_d	Distribution tax rate	0.12
τ_c	Corporate tax rate	0.35
c_E	Entry cost	0.103

Table 11. This table lists the parameters taken from outside the model corresponding to the 2006-2010 time period.

Inside parameters (2006-2010)		Estimate	Std Error
α	Revenue returns to scale	0.960	0.0005
θ	AR(1) in logs scale parameter	0.0230	0.0060
ϕ	AR(1) in logs persistence parameter	0.978	0.0002
σ_ϵ	AR(1) in logs standard deviation parameter	0.0408	0.0226
σ_1	Standard deviation of bivariate shock on revenue	0.244	0.0419
σ_2	Standard deviation of bivariate shock on operating expenses	0.243	0.0380
ρ	Correlation of bivariate shock	0.804	0.0140
c_v	Variable cost	3.253	0.0017
c_f	Fixed cost	0.0081	0.0625
λ	Equity floatation cost	0.0318	0.3198
P	Price	1.024	-

Table 12. This table lists the parameters estimated using the model corresponding to the 2006-2010 time period.

Moments (2006-2010)	Data	Model
Revenue mean	1.05	0.87
Revenue standard deviation	1.55	1.61
Operating expenses mean	0.95	0.79
Operating expenses standard deviation	1.44	1.46
Cash flow mean	0.043	0.051
Cash flow standard deviation	0.241	0.224
Capital mean	0.257	0.242
Capital standard deviation	0.447	0.433
Cash mean	0.2204	0.2293
Revenue - operating expenses covariance	2.21	2.32
Revenue autocovariance	2.04	2.37
Operating expenses autocovariance	1.76	1.88
Revenue - operating expenses$_{-1}$ covariance	1.87	2.08
Revenue$_{-1}$ - operating expenses covariance	1.89	2.15
Equity issuance mean	0.021	0.012
Exit rate	0.07	0.073

Table 13. This table lists the data moments from the 2006-2010 time period and the model moments which attempt to match them.

An estimation on the last 5 years finds that the correlation parameter decreases to 0.804. This is exciting because the estimation predicts a decline in correlation between revenue and operating expenses similar to the value used in the simple correlation decrease experiment.

1.6. Policy experiments

1.6.1. Corporate tax and real interest rate. The Obama administration has proposed that the top marginal corporate tax rate should be lowered to 28% as reported in Landler and Calmes (July 30, 2013). The prevailing idea is that a tax reduction along with a foreign tax holiday would propel firms to invest more and possibly hold less cash. Average investment does increase by 7.3% if this policy change is enacted. However, the model also predicts that average cash holdings rise by 11% as perhaps an unintended consequence.

Also, the real interest rate has hovered around 1% in the last few years. If the drop in the real interest rate suggests that there is a long-term shift in monetary policy, then the expectation would adjust as well. Thus, if the expected real interest rate drops from 3.4% to 1%, the model predicts that cash rises by 19%. Investment, on the other hand, increases by 6.8%.

1.6.2. Cash restrictions. Suppose that there are restrictions on cash. These restrictions may come from the government or from activist shareholders. First, starting from the parameter estimates for the 2006-2010 period, I look at the mean firm value when the option to hold cash is removed. As a baseline comparison, the mean firm value drops by 25% when no corporate cash holdings are allowed.

In the real world, a popular refrain is that firms should distribute excess cash. In my model, no cash is excess since all choices are fully rational. However, I can still run an experiment where firms are forced to distribute the cash that would not be necessary to cover any possible negative cash flows in the next period. When firms must distribute "excess" cash in this manner, mean firm value drops by 11% and mean cash drops by 35%. The period after the next period may require even more cash but accounting for the fact that the firm may need additional cash for many periods afterwards would entail not having a restriction at all at some point. The takeaway here is that cash restrictions can be quite harmful to firms.

1.7. Conclusion

The corporate cash increase is a phenomenon that has attracted a large amount of recent attention. This paper is an attempt to understand the phenomenon using a dynamic structural model with rational expectations. My model finds that 63% of the increase in corporate cash holdings can be accounted for by the increase in cash flow volatility through the mechanism of a correlation decrease between revenue and operating expenses. The correlation decrease observed in the data may be easily overlooked - however, careful attention to this issue might uncover other important insights.

In addition, I show that the standard regressions of cash on cash flow volatility may face endogeneity problems, and building a model to explain the data can provide a deeper understanding of firm behavior. Policies to induce firms to spend their cash such as lowering the corporate tax rate or the real interest rate increases firm value and investment but cash holdings increase as well. Finally, I establish that restrictions on cash can reduce firm value considerably.

References

Adda, Jerome and Russell Cooper, *Dynamic economics: quantitative methods and applications*, The MIT Press, 2003.

Almeida, Heitor, Murillo Campello, and Michael S. Weisbach, "The cash flow sensitivity of cash," *Journal of Finance*, 2004, *59*, 1777–1804.

———, ———, and ———, "Corporate financial and investment policies when future financing is not frictionless," *Journal of Corporate Finance*, 2011, *17*, 675–693.

Armenter, Roc and Viktoria Hnatkovska, "The macroeconomics of firms' savings," *mimeo*, 2012.

Bakke, Tor-Erik and Toni M. Whited, "Threshold events and identification: a study of cash shortfalls," *mimeo*, 2011.

Bates, Thomas W., Kathleen M. Kahle, and Rene M. Stulz, "Why do U.S. firms hold so much more cash than they used to?," *Journal of Finance*, 2009, *64*, 1985–2021.

Boileau, Martin and Nathalie Moyen, "Corporate cash savings: precaution versus liquidity," *mimeo*, 2010.

Bolton, Patrick, Hui Chen, and Neng Wang, "A unified theory of Tobin's q, corporate investment, financing, and risk management," *Journal of Finance*, 2011, *66*, 1545–1578.

———, ———, and ———, "Market timing, investment, and risk management," *Journal of Financial Economics*, 2013, *109*, 40–62.

Cooley, Thomas F. and Vincenzo Quadrini, "Financial markets and firm dynamics," *American Economic Review*, 2001, *91*, 1286–1310.

Cooper, Russell W. and John C. Haltiwanger, "On the nature of capital adjustment costs," *Review of Economic Studies*, 2006, *73*, 611–633.

Dittmar, Amy and Jan Mahrt-Smith, "Corporate governance and the value of cash holdings," *Journal of Financial Economics*, 2007, *83*, 599–634.

Duchin, Ran, "Cash holdings and corporate diversification," *Journal of Finance*, 2010, *65*, 955–992.

Dunne, Timothy, Mark J. Roberts, and Larry Samuelson, "Patterns of firm entry and exit in U.S. manufacturing industries," *RAND Journal of Economics*, 1988, *19*, 495–515.

Falato, Antonio, Dalida Kadyrzhanova, and Jae W. Sim, "Rising intangible capital, shrinking debt capacity, and the US corporate savings glut," *mimeo*, 2013.

Faulkender, Michael and Rong Wang, "Corporate financial policy and cash holdings," *Journal of Finance*, 2006, *61*, 1957–1990.

Foley, C. Fritz, Jay C. Hartzell, Sheridan Titman, and Garry Twite, "Why do firms hold so much cash? A tax-based explanation," *Journal of Financial Economics*, 2007, *86*, 579–607.

Gamba, Andrea and Alexander Triantis, "The value of financial flexibility," *Journal of Finance*, 2008, *63*, 2263–2296.

Gao, Xiaodan, "Corporate cash hoarding: the role of just-in-time adoption," *mimeo*, 2013.

Gomes, Joao F., "Financing investment," *American Economic Review*, 2001, *91*, 1263–1285.

Han, Seungjin and Jiaping Qiu, "Corporate precautionary cash holdings," *Journal of Corporate Finance*, 2007, *13*, 43–57.

Hennessy, Christopher A. and Toni Whited, "Debt dynamics," *Journal of Finance*, 2005, *60*, 1129–1165.

______ **and** ______, "How costly is external financing? Evidence from a structural estimation," *Journal of Finance*, 2007, *62*, 1705–1745.

Hopenhayn, Hugo A., "Entry, exit, and firm dynamics in long run equilibrium," *Econometrica*, 1992, *60*, 1127–1150.

Hopenhayn, Hugo and Richard Rogerson, "Job turnover and policy evaluation: a general equilibrium analysis," *Journal of Political Economy*, 1993, *101*, 915–938.

Jones, Charles I., "Economic growth and the relative price of capital," *Journal of Monetary Economics*, 1994, *34*, 359–382.

Karabarbounis, Loukas and Brent Neiman, "Declining labor shares and the global rise of corporate saving," *mimeo*, 2012.

Khurana, Inder K., Xiumin Martin, and Raynolde Pereira, "Financial development and the cash flow sensitivity of cash," *Journal of Financial and Quantitative Analysis*, 2006, *41*, 787–807.

Krusell, Per, "Investment-specific R&D and the decline in the relative price of capital," *Journal of Economic Growth*, 1998, *3*, 131–141.

Landler, Mark and Jackie Calmes, "Obama proposes deal over taxes and jobs," *The New York Times*, July 30, 2013.

Lindsey, Monie and Tony Carfang, "Treasury Strategies' Quarterly Corporate Cash Briefing," *Treasury Strategies Industry Research*, April 4, 2013.

McLean, R. David, "Share issuance and cash savings," *Journal of Financial Economics*, 2011, *99*, 693–715.

Minton, Bernadette A. and Catherine Schrand, "The impact of cash flow volatility on discretionary investment and the costs of debt and equity financing," *Journal of Financial Economics*, 1999, *54*, 423–460.

Nikolov, Boris and Toni M. Whited, "Agency conflicts and cash: estimates from a structural model," *mimeo*, 2011.

Opler, Tim, Lee Pinkowitz, Rene Stulz, and Rohan Williamson, "The determinants and implications of corporate cash holdings," *Journal of Financial Economics*, 1999, *52*, 3–46.

Pinkowitz, Lee, Rene Stulz, and Rohan Williamson, "Does the contribution of corporate cash holdings and dividends to firm value depend on governance? A cross-country analysis," *Journal of Finance*, 2006, *61*, 2725–2751.

_______, _______, and _______, "Multinationals and the high cash holdings puzzle," *mimeo*, 2012.

Riddick, Leigh A. and Toni M. Whited, "The corporate propensity to save," *Journal of Finance*, 2009, *64*, 1729–1766.

Subramaniam, Venkat, Tony T. Tang, Heng Yue, and Xin Zhou, "Firm structure and corporate cash holdings," *Journal of Corporate Finance*, 2011, *17*, 759–773.

Tauchen, George, "Finite state Markov-chain approximations to univariate and vector autoregressions," *Economics Letters*, 1986, *20*, 177–181.

1.8. Appendix

1.8.1. Computational algorithm.

(1) Set the grid to 25 points along the capital dimension where $k \in [0, 2]$, and 20 points along the cash dimension where $m \in [0, 1]$. Let the persistent shock z have 10 points and the transitory shock η_1 and η_2 have 5 points along each dimension. The persistent shock is discretized using the Adda-Cooper method and the transitory shock is discretized using the Tauchen method.

(2) Set an initial value for the price P.

(3) Solve for the decision rules and value functions.

(4) Find the entry cost for the economy. Then use bisection and repeat Step 3 to find the P which generates entry cost c_E.

(5) Set an initial value for the mass of entry M'.

(6) Solve for the invariant distribution.

(7) Find the quantity supplied for the economy. Then use bisection and repeat Step 6 to find the M' which generates quantity supplied $Q_s = Q_d$.

(8) Finally, the estimation is another outside loop which minimizes the mean squared distance between data moments and model moments.

1.8.2. Normalization. Recall that the profit function is,

$$\pi(k, z, \eta_1, \eta_2; P) = P\eta_1 z k^{\alpha} - \eta_2 c_v k - c_f.$$

Let A denote the mean total assets of firms in the economy and then normalize by dividing through by A to get,

$$\frac{\pi}{A} = \frac{P\eta_1 z k^{\alpha}}{A} - \frac{\eta_2 c_v k}{A} - \frac{c_f}{A}.$$

Let $\hat{z} = \frac{z}{A^{1-\alpha}}$ and rewrite the previous equation as,

$$\frac{\pi}{A} = P\eta_1 \hat{z}\left(\frac{k}{A}\right)^{\alpha} - \eta_2 c_v \left(\frac{k}{A}\right) - \frac{c_f}{A}.$$

Now assume that there is real growth in the economy up to time T which can be represented by,

$$G_T = \prod_{t=0}^{T}(1 + g_t)$$

where t is the time index, g_t is the per period growth rate, and $g_0 = 0$. Assume that $\hat{z}_T = G_T^{1-\alpha} z$ so that the profit function with real growth is,

$$G_T\pi = P\eta_1 \hat{z}_T (G_T k)^{\alpha} - \eta_2 c_v (G_T k) - G_T c_f.$$

Assume that the mean total assets of firms in the economy also grows at the same rate such that $A_T = G_T A$ is the mean total assets at time T. Therefore, the normalization now gives,

$$\frac{G_T\pi}{G_T A} = \frac{P\eta_1 \hat{z}_T (G_T k)^{\alpha}}{G_T A} - \frac{\eta_2 c_v (G_T k)}{G_T A} - \frac{G_T c_f}{G_T A}$$

which finally transforms to,

$$\frac{G_T\pi}{G_TA} = P\eta_1\hat{z}\left(\frac{G_Tk}{G_TA}\right)^{\alpha} - \eta_2 c_v\left(\frac{G_Tk}{G_TA}\right) - \frac{G_Tc_f}{G_TA}.$$

1.8.3. Decomposing the Data. The cost of goods sold (COGS) has become less correlated with revenue while the research and development expenses (RD) have become more correlated with revenue over the last 30 years (see Figure 1.8.1. However, the correlation between revenue and selling, general, and administrative expenses (SGA) have fluctuated with no general trend.

COGS compose around 70% of operating expenses (see Figure 1.8.2. Therefore, the decline in the revenue-COGS correlation is the primary source of the decrease in the correlation between revenue and operating expenses.

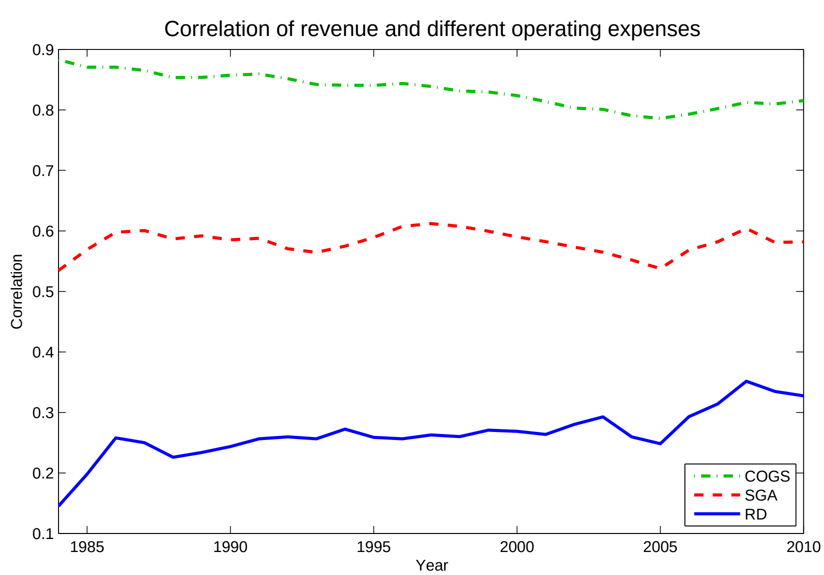


FIGURE 1.8.1. This figure breaks down the correlation between revenue and various types of operating expenses.

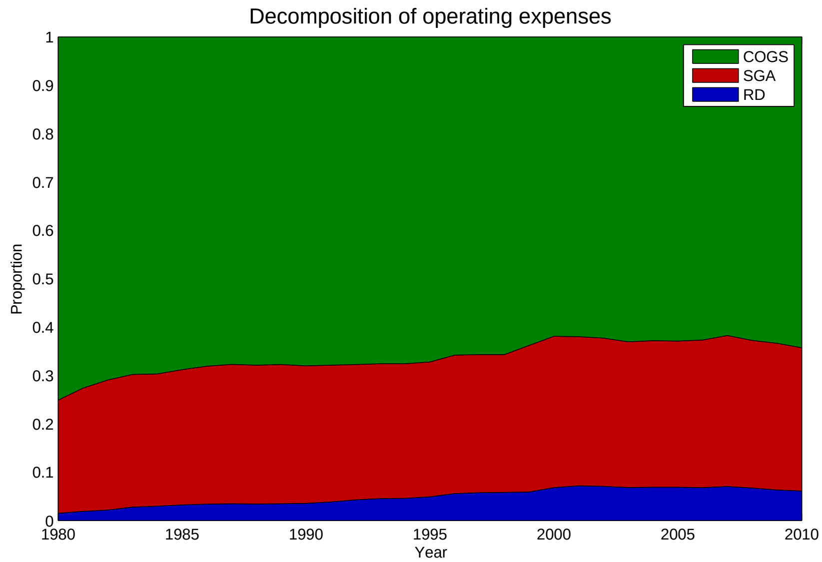


FIGURE 1.8.2. This figure breaks down the proportion of the various types of operating expenses.

Table 14 shows that the cash increase and correlation decrease occurred in every major Standard Industry Classification (SIC) industry. In fact, the industries which experienced the greatest cash increases also had the most significant correlation decreases between revenue and operating expenses.

	1980-1984			
Industry	# Firm-Year	Cash Ratio	Correlation	Domestic %
Agriculture	65	0.173	0.915	1.000
Mining	1194	0.109	0.738	0.995
Construction	287	0.088	0.960	0.999
Manufacturing	5084	0.115	0.933	0.991
Transportation	372	0.099	0.923	0.993
Communications	238	0.127	0.973	1.000
Technology	3027	0.147	0.880	0.990
Wholesale trade	826	0.087	0.964	0.994
Retail trade	1342	0.105	0.957	0.999
Services	1297	0.139	0.922	0.996
Healthcare	457	0.165	0.895	0.992

	2006-2010			
Industry	# Firm-Year	Cash Ratio	Correlation	Domestic %
Agriculture	50	0.184	0.712	0.980
Mining	696	0.160	0.497	0.945
Construction	119	0.153	0.949	0.933
Manufacturing	3688	0.348	0.743	0.927
Transportation	375	0.123	0.913	0.981
Communications	362	0.223	0.852	0.971
Technology	3610	0.331	0.852	0.868
Wholesale trade	389	0.121	0.961	0.936
Retail trade	781	0.159	0.953	0.985
Services	1207	0.255	0.820	0.934
Healthcare	870	0.316	0.868	0.943

TABLE 14. This table lists the number of firm-year observations, the mean cash ratio, the mean correlation between revenue and operating expenses, and the percentage of domestic income by industry for the first 5 years and the last 5 years of the sample period. The firms included in the statistics all have less than 1 billion 2010 dollars in total assets.

The correlation decrease is an interesting feature of the data and the decoupling of revenue and operating expenses can happen due to location and/or timing. In this section, I will explore the location explanation. Suppose that the bivariate normal shock had the following structure instead,

$$\eta_1 = \omega\eta_e + (1-\omega)\eta_w$$

$$\eta_2 = \nu\eta_e + (1-\nu)\eta_w$$

where $\omega \in [0,1]$ and $\nu \in [0,1]$. This would imply that there are regional components, namely east and west, to the shocks on revenue and operating expenses. The data suggests that revenue has become more global while operating expenses have remained quite local. Table 14 and Figure 1.8.3 also indicate that the industries which have become more global experienced the more substantial correlation declines. To be clear, this explanation is different from a cash increase due to repatriation taxes - rather, it is about the regional nature of the shocks. Pinkowitz et al. (2012) find that foreign tax holidays do little to reduce cash holdings which would imply that repatriation taxes do not have as large of an effect as found in Foley et al. (2007).

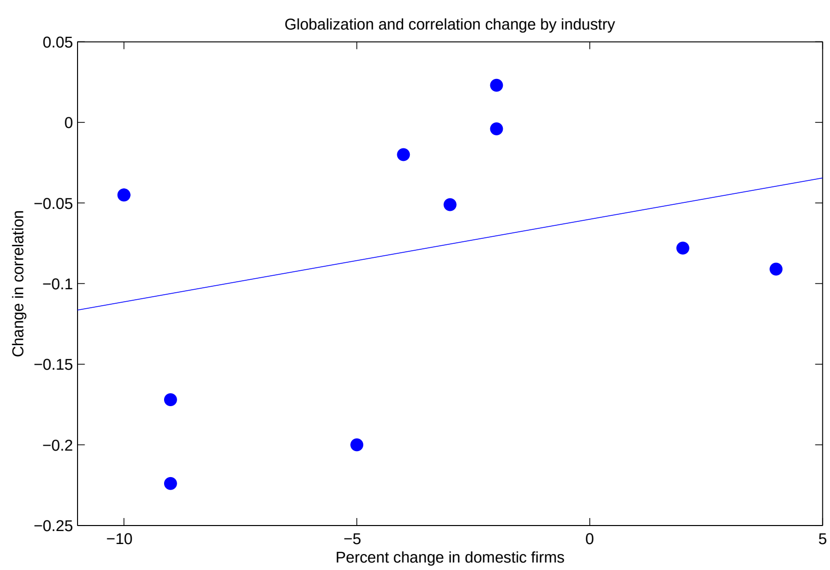


FIGURE 1.8.3. This figure plots the change in the percentage of domestic firms and the change in correlation between revenue and operating expenses by industry.

CHAPTER 2

Why are small Chinese manufacturing firms *less* constrained than large Chinese manufacturing firms?

Abstract. I show that small Chinese manufacturing firms are less constrained than large Chinese manufacturing firms. Evidence of this finding comes from analyzing the cash flow sensitivity of investment and from estimating an investment Euler equation using GMM. My result is in direct contrast to most of the literature on financial constraints where small firms are found to be substantially more constrained. Small Chinese manufacturing firms are less constrained because they are far more labor intensive relative to large Chinese manufacturing firms and are not as reliant on capital investment for growth.

2.1. Introduction

A negative relationship between the degree of financial constraints and firm size is almost canon in the financial constraints literature. Theory suggests that small firms, which also tend to be younger firms, face more asymmetric information. This relationship is so ubiquitous that it is often used to test the quality as well as to construct indexes which measure financial constraints.[1]

However, I find that a negative relationship does not hold for Chinese firms, and instead, small Chinese manufacturing firms actually appear to be less constrained than large Chinese manufacturing firms.[2] The data set that I use comes from China's National Bureau of Statistics (NBS) and it is a comprehensive sample of Chinese manufacturing firms from 1999 to 2007. With the data, I investigate the cash flow sensitivity of investment along the lines of Fazzari et al. (1988). Although there has been a large body of literature calling into question the usefulness of investment-cash flow regressions,[3] it is still a common starting point to analyze firm financing constraints. Specifically, measurement error in Tobin's q is a major contributing factor to the poor performance of investment-cash flow regressions. Annual sales growth therefore proxies for investment opportunities in the regressions that I run. My regressions on the Chinese firm data show that the cash flow sensitivity of investment increases dramatically with firm size holding investment opportunities constant while the opposite is true for US firms.

Taking the odd and counterintuitive behavior of the investment-cash flow regressions into account, I then employ the investment Euler equation methodology developed in Whited (1992), Bond and Meghir (1994), Love (2003), and Whited and Wu (2006) and tailor it for the Chinese data. In particular, I show that the positive relationship between the degree of financial constraints and firm size for Chinese firms comes from the fact that small Chinese

[1]For example, Hadlock and Pierce (2010) test the quality of several common financial constraint indexes and then they construct a new index based on firm size and age.
[2]But it should be emphasized that this finding does not imply that small Chinese firms are not constrained.
[3]Papers such as Kaplan and Zingales (1997), Erickson and Whited (2000), Gomes (2001), and Chen and Chen (2012) present evidence against investment-cash flow regressions.

firms are much more labor intensive than large Chinese firms. The labor share of production is 90% for the smallest firms and 46% for the largest firms. Given this huge difference in the cross-sectional labor share of production, it turns out that small Chinese firms do not have to undertake much capital investment at all to grow in a certain range. Small firms tend to operate in very labor intensive industries and can simply hire more workers to increase production in response to positive shocks. Another indication of financial constrainedness comes from the fact that the marginal value of additional capital is actually greater for large firms even though large firms have far more capital.

In development economics, a generally positive correlation has been observed between financial development and growth.[4] China however appears to be a counterexample. Guariglia et al. (2011) claim that the growth in Chinese firms in spite of a poorly developed financial system is due to the fact that Chinese firms retain a large amount of their earnings. They look at the cash flow sensitivity of total assets in particular. Private firms retain most of their earnings which suggests that these firms are quite financially constrained. State owned and collective enterprises in contrast do not have the same degree of sensitivities because they have better access to external capital markets. While retained earnings are important, I discover that the very structure of the production technology in China allow small firms, which face a high degree of asymmetric information, to need less capital investment.

The rest of the paper is organized as follows - Section 2.2 describes the literature in detail, Section 2.3 discusses the data, Section 2.4 runs the Fazzari et al. (1988) regressions on Chinese firms, Section 2.5 presents the investment Euler equation model, Section 2.6 performs the estimation, Section 2.7 provides GMM results, and Section 2.8 concludes.

2.2. Literature

Fazzari et al. (1988) first demonstrates that the cash flow sensitivity of investment can determine how constrained firms are. Firms face very different costs of external finance so firm heterogeneity is an important issue to consider. Financing constraints can help

[4] Levine (2005) provides a survey of financial development and growth.

explain the large variability of investment and should be accounted for in additional to "real" factors. Investment responds much stronger to sales than to the cost of capital in contrast to neoclassical theory. Low dividend firms retain most of their earnings. They also have the greatest cash flow sensitivity of investment and hence are the most constrained.

Kaplan and Zingales (1997) argue that there does not exist comprehensive theoretical or empirical evidence that the cash flow sensitivity of investment should increase monotonically with the degree of financing constraints. They look at the annual statements of 49 low dividend firms identified as constrained by FHP. From the financial statements, they find that less constrained firms in the "constrained" group actually have greater sensitivities. Splitting criteria based on preconceived notions (theoretical priors) of constrainedness can cause exactly opposite classifications. Finally, they argue that the FHP results are highly influenced by outliers.

Whited (1992) estimates an investment Euler equation. Financial variables should enter into the Euler equation through their effect on the debt constraint Lagrange multiplier. Firms are assumed to only issue debt in her model since it is usually the marginal source of funds. The existence of a presample bond rating provides the primary classification between constrained and unconstrained firms. This is in contrast to the FHP classification scheme which uses corporate dividend policies.

Bond and Meghir (1994) develop a formal test of liquidity constraints on firms using an Euler equation framework. Basically, the standard neoclassical Euler equation should only hold for unconstrained firms. The problem is that low dividend (constrained) firms tend to have a greater cash flow sensitivity of investment than standard neoclassical theory would predict.

Whited and Wu (2006) state that corporate returns moving together suggests that the firms might face similar financial constraints. They use GMM estimation of a investment Euler equation to compute the shadow value of relaxing financial constraints. GMM estimation avoids sample selection, simultaneity, and measurement error problems. They find that financial constraints are a source of priced risk, and that more constrained firms earn higher returns. Once financial constraints are controlled for, small firms no longer earn significantly higher returns. They also construct an index which classifies firms that are small, under invest, have low analyst coverage, and do not have bond ratings as financially constrained.

Gertler and Gilchrist (1994) look at the differential response between small and large firms to monetary policy. In particular, average inventories fall much more for small firms during periods of tight monetary policy. This behavior suggests that financial factors are at play. The data set they use is comprehensive for manufacturing firms but does not contain firm level detail. They find that even though small firms are defined to have 30% share of sales, they account for more than half of all manufacturing decline following tight monetary policy.

2.3. Data

The data comes from the National Bureau of Statistics (NBS) of China and contains a panel of firms from 1999 to 2007 inclusive. All firms with more than 5 million CNY in sales, or approximately 600 thousand USD during this time period, are required to provide detailed information regarding their financial situation. The information provided includes statistics such as employment numbers, income statement items, balance sheet items, and after 2004, cash flow items.

The available data only contains firms from the mining, manufacturing, and utility sectors. My paper will focus on the manufacturing sector because the mining sector is relatively small and operationally different from manufacturing while the utility sector is highly regulated in China. In addition, state-owned and collective corporations, which are also known as Township and Village Enterprises (TVEs), are removed from the sample. Each firm-year observation is classified as non-private corporations if the total state and collective paid in capital is greater than or equal to 50%.[5] I dropped firms with negative and missing sales, cost

[5] This type of classification is often used since official corporate ownership registrations can lag several years behind actual ownership changes. For instance, see Guariglia et al. (2011) for a similar approach.

of goods sold, depreciation, total assets, current assets, inventory, accounts receivable, fixed assets, total liabilities, total wages, and paid in capital. I also dropped firms with less than 5 million 1999 CNY in sales since there is likely significant undersampling and selection bias among the firms that are not required to report to the NBS. There are a total of 1,318,327 remaining firm-year observations after the data is filtered and cleaned.

I classify firms by size according to total assets and use percentile cutoffs on total assets for each year. The density breakdown is 5%, 10%, 15%, 20%, 20%, 15%, 10%, 5% and the cumulative density breakdown is 0-5%, 5-15%, 15-30%, 30-50%, 50-70%, 70-85%, 85-95%, 95-100% as seen in Table 1. This partition is quite attractive since mean total assets roughly doubles for each size group except for the largest size group due to the well-known right skewness of the firm size distribution. Table 2, on the other hand, presents summary statistics by year.

Percentile	# obs	Total assets (mil of 2005 CNY)	Capital (mil of 2005 CNY)	Sales (mil of 2005 CNY)	Number of employees
0-5	65950	1.9	0.7	13.1	68.5
5-15	131841	3.8	1.3	14.9	81.2
15-30	197769	6.5	2.2	18.1	97.8
30-50	263622	11.3	3.8	23.5	124.5
50-70	263660	21.7	7.5	35.3	171.8
70-85	197741	46.7	16.5	60.5	264.5
85-95	131833	121.3	42.5	129.8	465.1
95-100	65911	851.2	310.3	880.3	1464.6

Table 1. China summary statistics by firm size

Year	# obs	Total assets (mil of 2005 CNY)	Capital (mil of 2005 CNY)	Sales (mil of 2005 CNY)	Number of employees
1999	56252	76.3	30.6	63.0	304.4
2000	67537	77.8	29.8	70.5	301.2
2001	84526	73.0	28.1	68.8	281.9
2002	101582	70.4	26.6	70.5	268.5
2003	125626	72.1	26.1	78.8	263.3
2004	193307	59.1	20.9	69.7	221.0
2005	199747	68.9	24.5	85.9	238.7
2006	227377	70.7	24.7	93.5	230.8
2007	262373	71.5	24.1	98.8	222.3

Table 2. China summary statistics by year

For US data, I use Compustat and correspondingly drop all non-manufacturing sectors and missing and negative data. I also keep only the years from 1999 to 2007 inclusive and partition the firms according to the same densities. Of course, Compustat is a data set of public firms so that the average firm is much larger. However, the observed patterns by firm size can still be informative. Table 3 and Table 4 provides summary statistics by firm size and by year. There are 8 firm size categories and the last 3 Chinese firm sizes are approximately equivalent to the first 3 US firm sizes by the value of average total assets. It is unclear whether this comparison is completely valid though since similarly sized Chinese and US firms may face very different labor costs and have very different purchasing power. Therefore, a large Chinese firm may act more like a large public US firm rather than a small public US firm even if the value of total assets is more similar to the latter.

Percentile	# obs	Total assets (mil of 2005 USD)	Capital (mil of 2005 USD)	Sales (mil of 2005 USD)	Number of employees
0-5	1014	7.8	1.3	8.8	58.9
5-15	2018	21.0	3.7	22.6	142.8
15-30	3028	52.1	8.9	50.7	297.2
30-50	4038	142.2	26.7	136.1	751.0
50-70	4041	474.7	105.5	477.6	2292.9
70-85	3028	1628.5	419.8	1522.3	6966.5
85-95	2018	6436.9	1914.2	5920.8	23848.1
95-100	1005	48729.7	14102.5	41935.3	93921.7

Table 3. US summary statistics by firm size

Year	# obs	Total assets (mil of 2005 USD)	Capital (mil of 2005 USD)	Sales (mil of 2005 USD)	Number of employees
1999	2612	2456.1	714.5	2140.8	7281.3
2000	2576	2725.9	775.6	2436.4	8052.4
2001	2413	2883.6	801.6	2462.1	8332.2
2002	2267	3219.2	877.0	2665.0	8632.9
2003	2142	3685.0	1012.0	3082.6	8952.9
2004	2125	3950.6	1050.1	3494.1	9347.2
2005	2073	4036.1	1066.7	3650.1	9457.8
2006	2033	4224.6	1164.8	3917.1	9758.5
2007	1949	4442.1	1208.0	4095.1	9651.7

Table 4. US summary statistics by year

The main thing to note in the comparison of summary statistics is that the employment numbers and the sales to capital ratio look very different between Chinese and US firms. Small Chinese firms are much more labor intensive and have much less capital relative to the amount of sales or value added they are able to generate. The capital and labor share of output can then be estimated by imputing a rental price of capital since the NBS data includes annual wages paid by each firm. Table 5 lists the average capital share of output computed in this manner for different firm sizes.

Percentile	Capital share of production (100α)
0-5	9.7
5-15	14.7
15-30	19.2
30-50	23.8
50-70	29.5
70-85	35.4
85-95	41.1
95-100	54.1
Small	20.4
Large	46.5

TABLE 5. Capital share of production for Chinese firms

2.4. Regressions

I run the Fazzari et al. (1988) regressions on the Chinese and US data to investigate the behavior of investment-cash flow sensitivity across firm size. Table 6 and Table 7 report the regression variable summary statistics by firm size for China and US respectively. Again, some interesting features arise. For example, equity issuance increases with firm size in China while the opposite is true in the US. Cash is mostly the same across firm size for Chinese firms while cash is significantly higher for small US firms. Interest to sales (I to S) is very low for small Chinese firms even though interest to liabilities (I to L) is as high as in the US. In the US, interest to sales is comparatively much greater for small firms. This last fact seems to suggest that financial markets are substantially less developed and/or there is more asymmetric information in China. That is, Chinese firms which are apparently able to easily pay off their interest still face a high interest rate. Financial markets then are probably not close to being perfectly competitive.

Percentile	Cash flow	Growth	Cash	Capex	Issuance	I to S	I to L	L to A
0-5	24.1	15.1	21.4	-6.5	-0.7	0.3	5.5	49.7
5-15	17.3	16.3	20.5	2.6	1.0	0.4	4.0	54.7
15-30	14.2	17.0	20.0	7.2	2.1	0.6	3.3	56.3
30-50	11.6	17.3	20.0	10.5	2.9	0.8	2.8	57.2
50-70	10.4	18.3	20.6	12.6	3.5	1.1	2.5	57.1
70-85	9.4	19.2	21.0	13.7	3.7	1.4	2.3	57.0
85-95	9.1	20.1	21.9	13.5	3.8	1.9	2.2	57.0
95-100	9.4	22.2	23.5	13.6	3.3	2.3	2.1	58.5

Table 6. China regression variable summary by firm size (all values are multiplied by 100)

Percentile	Cash flow	Growth	Cash	Capex	Issuance	I to S	I to L	L to A
0-5	-29.6	0.2	34.8	-9.7	13.0	4.7	5.1	43.5
5-15	-13.4	5.4	32.8	-4.2	8.8	3.3	3.6	40.7
15-30	-6.7	7.2	34.2	-0.5	6.3	2.8	3.1	38.7
30-50	-0.3	9.6	33.3	5.8	4.2	2.5	2.8	36.8
50-70	5.6	9.8	25.4	7.4	1.9	2.7	3.0	44.4
70-85	8.0	9.5	17.6	6.0	-0.4	2.7	3.2	55.4
85-95	9.1	9.0	14.4	5.8	-1.1	2.4	2.9	59.4
95-100	10.8	9.4	14.4	6.5	-1.7	2.0	2.2	58.9

Table 7. US regression variable summary by firm size (all values are multiplied by 100)

Table 8 and Table 9 report the regression variable summary statistics by year for China and US respectively. The most striking phenomenon in this specific breakdown is that China clearly experiences a great deal of persistent growth from 1999 to 2007.

Year	Cash flow	Growth	Cash	Capex	Issuance	I to S	I to L	L to A
1999	-	-	17.8	-	-	2.1	3.7	58.9
2000	8.8	12.6	18.3	5.6	2.3	1.8	3.4	58.6
2001	8.5	6.6	19.3	6.0	2.4	1.4	3.1	57.4
2002	9.2	11.7	19.9	8.0	2.8	1.2	2.9	56.7
2003	9.8	17.9	20.6	10.7	2.9	1.1	2.8	56.5
2004	9.4	16.1	20.5	7.8	3.2	0.9	2.4	58.0
2005	11.9	21.5	21.2	14.8	3.0	0.9	2.8	56.1
2006	12.9	19.6	21.5	10.6	3.0	0.8	2.8	55.4
2007	14.8	21.8	22.0	10.6	2.8	0.9	3.1	54.8

Table 8. China regression variable summary by year (all values are multiplied by 100)

Year	Cash flow	Growth	Cash	Capex	Issuance	I to S	I to L	L to A
1999	-	-	22.3	-	-	2.8	3.5	47.6
2000	1.1	13.2	24.8	7.9	5.6	3.1	3.9	45.8
2001	-3.6	-2.2	26.0	2.6	3.4	2.9	3.4	46.0
2002	-2.5	0.1	26.1	-2.6	1.9	2.9	3.0	46.6
2003	0.3	7.5	27.5	-1.2	3.6	2.7	2.8	44.7
2004	2.2	14.7	28.4	2.9	3.7	2.6	2.7	43.6
2005	1.8	11.8	29.1	2.8	2.7	2.5	2.6	44.3
2006	1.5	12.3	29.1	7.8	2.7	2.6	2.7	44.9
2007	0.7	11.4	29.5	8.8	2.3	2.9	2.9	44.8

Table 9. US regression variable summary by year (all values are multiplied by 100)

Finally, Table 10 and Table 11 report the overall regression variable summary statistics for China and US respectively. The number of outliers are reduced by winsorizing at the 1% tails when appropriate and by normalizing growth and capital expenditure with the annual midpoint value of sales and capital respectively.

Statistic	Cash flow	Growth	Cash	Capex	Issuance	I to S	I to L	L to A
# obs	902950	902950	1318317	902422	902950	1223352	1218334	1318317
Mean	11.8	18.3	20.8	10.4	2.9	1.1	2.9	56.4
Stdev	16.5	42.0	18.6	51.6	16.4	2.1	7.2	26.9
Min	-13.6	-199.7	0.0	-200.0	-55.0	0.0	0.0	1.2
Max	95.1	199.5	100.0	200.0	76.3	89.8	100.0	122.8

TABLE 10. China regression variable summary (all values are multiplied by 100)

Statistic	Cash flow	Growth	Cash	Capex	Issuance	I to S	I to L	L to A
# obs	16639	16639	20016	16626	14569	18582	18552	19971
Mean	0.1	8.3	26.7	3.5	3.3	2.8	3.1	45.4
Stdev	23.1	32.8	24.2	34.5	15.2	6.5	3.8	26.3
Min	-105.1	-194.2	0.0	-200.0	-20.1	0.0	0.0	5.1
Max	35.0	193.0	100.0	200.0	93.4	97.6	93.6	140.9

TABLE 11. US regression variable summary (all values are multiplied by 100)

The investment-cash flow regression is modeled as,

$$I_{it} = a_0 + a_1 CF_{it} + a_2 SG_{it} + a_3 ISG_{it} + a_4 CASH_{it} + a_5 ITL_{it} + a_6 LTA_{it} + \nu_{it}$$

where ν_{it} is the error term. In addition, CF is cash flow, SG is sales growth, ISG is industry sales growth, $CASH$ is cash and short-term investments, and LTA is the liabilities to assets ratio. Small firms are below median size firms and large firms are above median size firms. Table 12 presents the FHP cash flow sensitivity of investment results. Small Chinese manufacturing firms have about one third as much sensitivity relative to large Chinese manufacturing firms after controlling for investment opportunities. In contrast, small US firms have a high sensitivity and large US firms almost have no cash flow sensitivity of investment.

	China				US			
Regressor	Small	t-stat	Large	t-stat	Small	t-stat	Large	t-stat
CF	0.106	20.2	0.276	39.6	0.196	7.3	0.028	0.7
SG	0.161	58.8	0.220	103.4	0.308	16.7	0.497	23.9
ISG	0.166	11.6	0.116	10.4	0.053	1.2	-0.058	-2.0
$CASH$	0.288	52.7	0.287	59.3	0.157	5.9	0.190	8.2
ITL	-0.143	-12.4	-0.167	-8.5	-0.287	-1.8	-0.704	-3.8
LTA	0.010	2.7	-0.033	-10.4	-0.089	-3.8	-0.026	-1.7
$cons$	-0.061	-16.3	0.008	2.7	-0.011	-0.7	0.016	1.3

Table 12. Investment-cash flow regression

The fixed effects regression has the form,

$$I_{it} = a_0 + a_1 CF_{it} + a_2 SG_{it} + a_3 ISG_{it} + a_4 CASH_{it} + a_5 ITL_{it} + a_6 LTA_{it} + f_i + \nu_{it}$$

where f_i is the firm specific fixed effect. The cash flow coefficient changes significantly for small Chinese firms in the fixed effects regression which indicates that small Chinese firms often experience a persistent level of investment. Once this persistence is controlled for, they experience a higher sensitivity to cash flow similar to the level for large firms. The US sensitivities in contrast are still ordered in the same way as before.

	China				US			
Regressor	Small	t-stat	Large	t-stat	Small	t-stat	Large	t-stat
CF	0.226	19.1	0.215	16.6	0.181	3.7	0.028	0.5
SG	0.108	27.8	0.170	60.5	0.249	11.7	0.407	15.2
ISG	0.013	0.5	-0.373	-23.7	0.036	0.7	-0.136	-3.9
$CASH$	0.490	45.0	0.498	55.0	0.562	10.6	0.500	9.7
ITL	-0.252	-10.3	-0.576	-15.0	-0.855	-4.0	-2.340	-5.0
LTA	-0.150	-13.6	-0.144	-16.7	-0.184	-4.5	-0.014	-0.3
$cons$	0.012	1.4	0.142	22.4	-0.075	-2.9	0.010	0.3

Table 13. Investment-cash flow regression with fixed effects

2.5. Model

Assume that time is discrete and infinite and assume that firms in the economy are risk-neutral.[6] Let each firm be indexed by i and each time period be indexed by t. Also let K_{it} ($K_{i,t+1}$) denote beginning (end) of period capital stock and L_{it} denote beginning of period labor. The profit function is given by $\pi(K_{it}, L_{it}, z_{it})$ where $z_{it} \in [\underline{z}, \bar{z}] \subset \mathbb{R}_+$ is the idiosyncratic productivity shock. Firms invest or disinvest in capital each period as well and the law of motion of capital is defined as,

$$I_{it} = K_{i,t+1} - (1-\delta)K_{it}$$

where δ is the rate of depreciation. When firms invest or disinvest, they face a capital adjustment cost which is given by the function $\psi(I_{it}, K_{it})$.

Firms in addition can borrow and the beginning (end) of period debt is denoted by $B_{it} \subset \mathbb{R}_+$ ($B_{i,t+1} \subset \mathbb{R}_+$) and the interest rate is given by $r + \phi$ where r is the real interest rate and ϕ is an additional loan maintenance cost.

The dividend payment in each period is defined as,

$$d_{it} = \pi(K_{it}, L_{it}, z_{it}) - \psi(I_{it}, K_{it}) - I_{it} - (1 + r + \phi)B_{it} + B_{i,t+1} \tag{2.5.1}$$

subject to,

$$d_{it} \geq \underline{d}_{it} \tag{2.5.2}$$

[6]The model is based on Whited and Wu (2006) with modifications tailored to better describe the Chinese manufacturing data.

$$B_{i,t+1} \leq \bar{B}_{i,t+1} \tag{2.5.3}$$

where λ_{it} is the Lagrange multiplier for dividend constraint (2.5.2) and γ_{it} is the Lagrange multiplier for debt constraint (2.5.3). The Lagrange multiplier λ_{it} can be interpreted as the shadow cost of relaxing the dividend constraint which is also the shadow cost of equity issuance if $\underline{d}_{it}$ is negative. Firms maximize the expected discounted dividend to shareholders and the value of the firm is therefore,

$$V_{i0} = \mathrm{E}_{i0} \sum_{t=0}^{\infty} \left(\frac{1}{1+r} \right)^t d_{it} \tag{2.5.4}$$

The value of the firm can also be written recursively as a Bellman equation,

$$V = d + \frac{1}{1+r} \mathrm{E}[V'] \tag{2.5.5}$$

subject to (2.5.2) and (2.5.3). The first order conditions with respect to K, L, K', and B' are,

$$K : V_K = (1+\lambda)(\pi_K(K, L, z) + (1-\delta)(\psi_I(I, K) + 1) - \psi_K(I, K)) \tag{2.5.6}$$

$$L : \pi_L(K, L, z) = 0 \tag{2.5.7}$$

$$K' : 0 = (1+\lambda)(-\psi_I(I, K) - 1) + \frac{1}{1+r} \mathrm{E}[V'_K] \tag{2.5.8}$$

$$B' : 0 = (1+\lambda) - \frac{1}{1+r}\mathrm{E}[(1+\lambda')(1+r+\phi)] - \gamma \tag{2.5.9}$$

Equation (2.5.6) and (2.5.8) combines to produce the Euler equation,

$$\begin{aligned} &(1+\lambda)(\psi_I(I,K)+1) \\ &= \frac{1}{1+r}\mathrm{E}\left[(1+\lambda')(\pi_K(K',L',z') + (1-\delta)(\psi_I(I',K')+1) - \psi_K(I',K'))\right] \end{aligned} \tag{2.5.10}$$

which is a second order difference equation. The left hand side of the Euler equation is the cost of investing today which depends on the adjustment cost and the shadow cost of external funds. While the right hand side is the value of not investing and waiting until tomorrow which depends on the future adjustment cost, the shadow cost of future external funds, marginal product of future capital, and the discount rate. The optimal investment policy simply balances out the costs and benefits.

2.6. Estimation

The Euler equation (2.5.10) transforms to,

$$\begin{aligned} &\epsilon' + \psi_I(I,K)+1 \\ &= \frac{1}{1+r}\left[\left(\frac{1+\lambda'}{1+\lambda}\right)(\pi_K(K',L',z') + (1-\delta)(\psi_I(I',K')+1) - \psi_K(I',K'))\right] \end{aligned} \tag{2.6.1}$$

when the expectations operator is replaced by an expectational error ϵ' and $\mathrm{E}[\epsilon'] = 0$ and $\mathrm{E}[(\epsilon')^2] = \sigma^2$.

Assume that the production function exhibits constant returns to scale and has the form,

$$\pi(K,L,z) = zK^{\alpha}L^{1-\alpha} - wL \tag{2.6.2}$$

where w is the wage rate. The marginal production function with respect to capital is,

$$\pi_K(K, L, z) = \frac{\alpha z K^{\alpha} L^{1-\alpha}}{K} \tag{2.6.3}$$

The marginal production function with respect to labor is,

$$\pi_L(K, L, z) = \frac{(1-\alpha) z K^{\alpha} L^{1-\alpha}}{L} - w \tag{2.6.4}$$

Let the adjustment cost function have the form,

$$\psi(I, K) = \theta \frac{I^2}{K} \tag{2.6.5}$$

which is the functional form adopted in Cooper and Haltiwanger (2006). Then equation (2.6.1) transforms to,

$$\begin{aligned} &\epsilon' + 2\theta \frac{I}{K} + 1 \\ &= \frac{\Lambda}{1+r} \left(\frac{\alpha Y'}{K'} + (1-\delta) \left(2\theta \frac{I'}{K'} + 1 \right) + \theta \left(\frac{I'}{K'} \right)^2 \right) \end{aligned} \tag{2.6.6}$$

where $\Lambda = \frac{1+\lambda'}{1+\lambda}$ and Y is value added.

Let the shadow cost of raising new equity be parameterized as,

$$\lambda_{i,t+1} = b_1 CF_{i,t+1} + b_2 SG_{i,t+1} + b_3 ISG_{i,t+1} + b_4 CASH_{i,t+1} + b_5 LTA_{i,t+1} \tag{2.6.7}$$

where CF is cash flow, SG is sales growth, ISG is industry sales growth, $CASH$ is cash and short-term investments, and LTA is the liabilities to assets ratio. The only regressor in the FHP regression that is not used as a determinant of the financial constraints Lagrange multiplier is the interest to liabilities ratio ITL. This parametric assumption on the Lagrange multiplier goes beyond the formal structure of the model but it allows for the estimation of a single investment Euler equation based on observables. The parameterization is based on Whited (1992), Love (2003), and Whited and Wu (2006) and uses similar observables. A more complicated model that endogenizes all these variables is beyond the scope of this paper.

2.7. GMM Results

GMM estimation of the investment Euler equation 2.6.1 gives,

$$\lambda_{i,t+1} = b_1 CF_{i,t+1} + b_2 SG_{i,t+1} + b_3 ISG_{i,t+1} + b_4 CASH_{i,t+1} + b_5 LTA_{i,t+1}$$

where $\lambda_{i,t+1}$ is the Lagrange multiplier on the external finance constraint. Table 14 summarizes the GMM estimation results. Both small and large Chinese firms are less financially constrained when cash flow, industry growth, and cash rises. On the other hand, both small and large Chinese firms are more financially constrained when individual firm growth increases. Finally, the coefficient for the liabilities to assets ratio has different signs for small versus large firms. This suggests that small firms are actually less constrained when debt increases. One explanation for this phenomenon is that being able to borrow provides a positive signal for small firms.

Regressor	Small	Stderr	Large	Stderr
θ	0.3834	0.1275	0.0004	0.0003
b_1	-0.2110	0.0004	-7.1098	0.0000
b_2	0.1939	0.0001	1.0633	0.0000
b_3	-1.4299	0.0000	-2.8160	0.0001
b_4	-0.0138	0.0005	-0.8411	0.0000
b_5	-0.4218	0.0001	2.5302	0.0001

Table 14. GMM estimation of the investment Euler equation

Using the estimated values for b_1, b_2, b_3, b_4, and b_5, I can construct Table 15 which provides a summary of the financial constraints index for small and large Chinese firms. Smaller firms are in fact less constrained than larger firms given the estimates.

Firm size	# obs	Mean	Stdev	Min	Max
Small	376872	-0.493	0.154	-1.522	0.345
Large	480948	0.261	1.331	-9.069	5.405

Table 15. GMM estimation of the investment Euler equation

2.8. Conclusion

In contrast to most of the literature on financial constraints, I find that small Chinese manufacturing firms are less constrained than large Chinese manufacturing firms. While this result may seem very counterintuitive at first, I show that small Chinese firms are much more labor intensive and do not need a large amount of capital investment to grow. Small firms also have a lower marginal value of capital. Therefore, the labor to capital ratio and the marginal value of capital may be important issues to consider when looking at financial constraints across different firm sizes. Future research on Chinese firms, on the other hand, might investigate the implications of such wide cross-sectional differences in labor intensity. For example, how much of this difference is due to selection? How and why do firms transition into greater capital intensity as they grow larger?

References

Bond, Stephen and Costas Meghir, "Dynamic Investment Models and the Firm's Financial Policy," *Review of Economic Studies*, 1994, *61*, 197–222.

Carpenter, Robert E. and Bruce C. Petersen, "Is the Growth of Small Firms Constrained by Internal Finance?," *Review of Economics and Statistics*, 2002, *84*, 298–309.

Chen, Huafeng and Shaojun Chen, "Investment-cash flow sensitivity cannot be a good measure of financial constraints: Evidence from the time series," *Journal of Financial Economics*, 2012, *103*, 393–410.

Chen, Jean J., "Determinants of capital structure of Chinese-listed companies," *Journal of Business Research*, 2003, *57*, 1341–1351.

Chow, Clement Kong-Wing and Michael Ka Yiu Fung, "Small Business and Liquidity Constraints in Financing Business Investment: Evidence from Shanghai's Manufacturing Sector," *Journal of Business Venturing*, 2000, *15*, 363–383.

Cooper, Russell W. and John C. Haltiwanger, "On the Nature of Capital Adjustment Costs," *Review of Economic Studies*, 2006, *73*, 611–633.

Erickson, Timothy and Toni M. Whited, "Measurement Error and the Relationship between Investment and q," *Journal of Political Economy*, 2000, *108*, 1027–1057.

Fazzari, Steven M., R. Glenn Hubbard, and Bruce C. Petersen, "Financing Constraints and Corporate Investment," *Brookings Papers on Economic Activity*, 1988, *19*, 141–206.

Gertler, Mark and Simon Gilchrist, "Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms," *Quarterly Journal of Economics*, 1994, *109*, 309–340.

Gollin, Douglas, "Getting Income Shares Right," *Journal of Political Economy*, 2002, *110*, 458–474.

Gomes, Joao F., "Financing Investment," *American Economic Review*, 2001, *91*, 1263–1285.

Guariglia, Alessandra, Xiaoxuan Liu, and Lina Song, "Internal finance and growth: Microeconometric evidence on Chinese firms," *Journal of Development Economics*, 2011, *96*, 79–94.

Hadlock, Charles J. and Joshua R. Pierce, "New Evidence on Measuring Financial Constraints: Moving Beyond the KZ Index," *Review of Financial Studies*, 2010, *23*, 1909–1940.

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

Kaplan, Steven N. and Luigi Zingales, "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?," *Quarterly Journal of Economics*, 1997, *112*, 169–215.

Kashyap, Anil K., Owen A. Lamont, and Jeremy C. Stein, "Credit Conditions and the Cyclical Behavior of Inventories," *Quarterly Journal of Economics*, 1994, *109*, 565–592.

Levine, Ross, "Finance and Growth: Theory and Evidence," in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, 1 ed., Vol. 1 2005, chapter 12, pp. 865–934.

Love, Inessa, "Financial Development and Financing Constraints: International Evidence from the Structural Investment Model," *Review of Financial Studies*, 2003, *16*, 765–791.

Poncet, Sandra, Walter Steingress, and Hylke Vandenbussche, "Financial constraints in China: Firm-level evidence," *China Economic Review*, 2010, *21*, 411–422.

Whited, Toni M., "Debt, Liquidity Constraints, and Corporate Investment: Evidence from Panel Data," *Journal of Finance*, 1992, *47*, 1425–1460.

———, "External finance constraints and the intertemporal pattern of intermittent investment," *Journal of Financial Economics*, 2006, *81*, 467–502.

——— **and Guojun Wu**, "Financial Constraints Risk," *Review of Financial Studies*, 2006, *19*, 531–559.

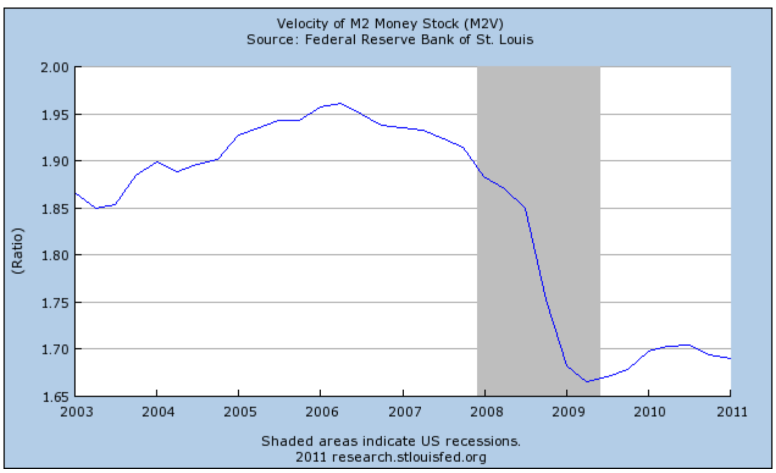
CHAPTER 3

Quantitative Easing and the Velocity of Money

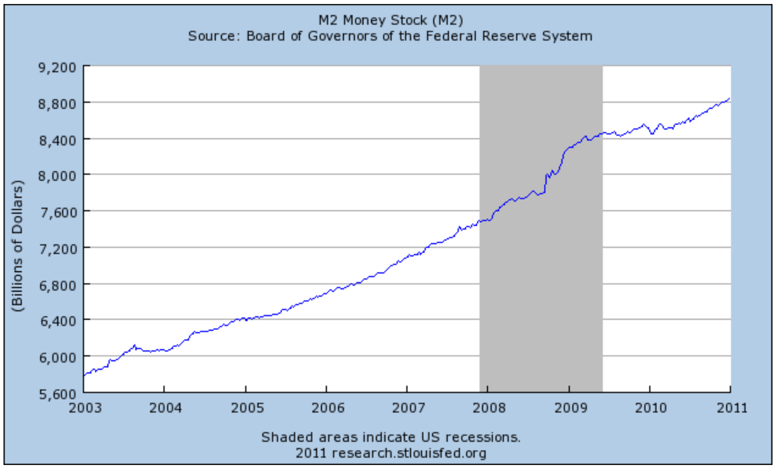
Abstract. Why was the velocity of money so low after quantitative easing? Standard models predict that the velocity of money should be much higher than what is observed in the data. I find that people who hold the majority of their wealth in debt securities are much richer on average. The liquidity injection by the Federal Reserve through the purchase of debt securities has a dampening effect on money velocity since holders of these assets tend to switch to holding stocks or cash rather than consume more. In my model, heterogeneous households hold different types of assets. After estimating the degree of heterogeneity in the households, my model determines that 67% of the drop in the velocity of money can be attributed to the low wealth elasticity of demand for consumption goods and the high wealth elasticity of demand for financial assets by the people who hold the majority of their wealth in debt securities.

3.1. Introduction

Between late 2008 and early 2010, and between late 2010 and mid 2011, the U.S. Federal Reserve purchased large amounts of treasury securities, mortgage-backed securities, and other forms of debt. These purchases were done through two rounds of "quantitative easing." Predictably, this loose monetary policy pumped some liquidity into the financial system and propped up asset prices. However, quite puzzling, the velocity of money dropped to historical lows and the rate of inflation also remained low. The graph below illustrates the precipitous drop in the velocity of money between 2008 and 2009.



The classical equation under the quantity theory of money is $MV = PY$ where M is the money stock, V is the velocity, P is the price level, and Y is the real output. Therefore under conventional wisdom, an increase in the money stock mostly increases the price level while having little effect on velocity and real output as demonstrated empirically in Lucas (1980). *So why was the velocity of M2 money stock so low even after two rounds of quantitative easing?* The difference in the case of quantitative easing is that the agents which are targeted by monetary policy are bond holders. From the estimation, I show that the bond holders have distinctly different preferences than holders of equity, money, or debt. The bond holders that are affected on the margin by quantitative easing choose to convert bonds to equity or money, but when the previous period bond holders become money holders, they tend to keep money holdings high rather than spend more on consumption. This induces a sort of an unconventional liquidity trap, i.e. a liquidity trap unshaped by negative expectations.



The model in this paper relies on a joint distribution over the discount factor and risk aversion coefficient to generate the extremely heterogeneous behavior of agents. Individual discounting and risk aversion has been well investigated in the economic literature but very few papers have looked at both concurrently. Perhaps as a result, most dynamic economic models introduce little preference variation in discounting and risk aversion. A representative agent is generally used and heterogeneity instead comes from other sources such as a transitional earnings process or some other idiosyncratic shock. The drawback is that a possibly important form of heterogeneity is ignored. For instance, using telephone insurance micro-data, Cicchetti and Dubin (1994) discover significant heterogeneity in the degree of risk aversion in the population. Other studies such as Cameron and Gerdes (2007) use a survey-based approach for the estimation of heterogeneity in discounting and risk aversion pertaining specifically to climate change, and they also find a large amount of variation in preferences.

An economy is said to have market segmentation when non-empty subsets of the population participate in different financial markets. Most previous work in this area assumes that the market segmentation is exogenous, but Chatterjee and Corbae (1992), Alvarez, Atkeson, and Kehoe (2001), and Khan and Thomas (2009) were able to generate endogenous market segmentation by imposing fixed costs to bond market transactions. Chatterjee and

Corbae show that changes in the steady-state growth rate of the money stock have a negative effect on real interest rates. Alvarez, Atkeson, and Kehoe produce a negative relation between expected inflation and real interest rates. Khan and Thomas extend the literature by manufacturing a nontrivial distribution of money and time-varying market segmentation. Khan and Thomas also demonstrate that the price level can gradually adjust following a monetary shock. All three papers find persistent liquidity effects due to the fixed costs. In contrast to the previous models, my paper induces endogenous market segmentation with a joint distribution over the discount factor and risk aversion coefficient and with "forcing" parameters. In my opinion, heterogeneity in preferences is as natural of a mechanism as fixed costs of trading to generate endogeneity. A model with heterogeneity in preferences also allows different questions to be answered.

With the Survey of Consumer Finances (SCF), I can observe household financial asset choices conditional on the person's earnings state and net financial worth. I then construct a structural model of household financial behavior with pseudo-life-cycle properties and use simulated method of moments to estimate the distribution of discounting and risk aversion and other endogenous parameters. Since different discount factors and coefficients of relative risk aversion generate different value functions and decision rules, modern computational methods and resources are used as a necessity to determine the parameter estimates and standard errors. These methods and resources are largely untapped by economists but are widely employed in other disciplines. In this case, I show the power of scientific computing for estimation in the presence of high agent heterogeneity and an extensive state space.

Finally, using the parameter estimates, I simulate the model with and without quantitative easing to determine the effect of quantitative easing on financial asset flows and the velocity of money.

3.2. Base Model Economy

The base model is a partial equilibrium model with a unit measure of agents who have the option of participating in different financial markets. The time-separable utility for each individual is $u_i(c)$ where the utility function is strictly increasing, strictly concave, and differentiable. Since I will use household level data from the SCF and PSID, the terms household, individual, and agent will be used interchangeably to refer to the atomic unit of the decision maker from now on.

Let the discount rate and the coefficient of relative risk aversion come from a truncated bivariate normal distribution,

$$\begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix} \sim \mathcal{T}\left(\mu_\beta, \mu_\gamma, \sigma_\beta^2, \sigma_\gamma^2, \rho, \underline{\nu}_\beta, \bar{\nu}_\beta, \underline{\nu}_\gamma, \bar{\nu}_\gamma\right)$$

where β_i is left truncated at $\underline{\nu}_\beta$ and right truncated at $\bar{\nu}_\beta$, and γ_i is left truncated at $\underline{\nu}_\gamma$ and right truncated at $\bar{\nu}_\gamma$. The parameters $\mu_\beta, \mu_\gamma, \sigma_\beta^2, \sigma_\gamma^2$ are the pre-truncation means and variances.

Let the economy-wide bond returns Θ_b be i.i.d. $\mathcal{N}(\mu_b, \Sigma_b^2)$ distributed, and let the economy-wide stock returns Θ_s be i.i.d. $\mathcal{N}(\mu_s, \Sigma_s^2)$ distributed. Like in the capital asset pricing model (CAPM), Σ_b^2 and Σ_s^2 is the systematic or market risk. The unsystematic or idiosyncratic risk is σ_b^2 and σ_s^2 for bonds and stocks respectively. The systematic and unsystematic risk are assumed to be uncorrelated, and the total distribution of bond and stock returns is therefore $\theta_b \sim \mathcal{N}(\mu_b, \Sigma_b^2 + \sigma_b^2)$ and $\theta_s \sim \mathcal{N}(\mu_s, \Sigma_s^2 + \sigma_s^2)$.

The money return rate is constant and set to μ_m, and the debt price is also constant and set to $\frac{1}{\mu_d}$. There are no financial intermediaries and default in this model so the debt price has to be set exogenously. The amount of inflation is ι and it is also the expectation on future inflation. This is admittedly a strong assumption and requires that the simulated method of moments be performed during a period with relatively stable inflation. The asset choice is $\zeta \in \{b, s, m, d\}$ and let b, s, m, and d denote the choice of bonds, stocks, money, and debt respectively while $a \in [0, \infty)$ is the numerical amount that is held of the asset. The earnings e evolve according to a Markov process with transition probability $\pi(e'|e)$ where $e, e' \in [0, \infty)$.

In each period, a person has a discrete choice of whether to hold bonds, stocks, money, or debt into the next period. Therefore, the person can only hold one asset at a time. If bonds or stocks are chosen, there is a probability $\Lambda_b \in [0, 1)$ and $\Lambda_s \in [0, 1)$ respectively that his holdings will not be convertible to consumption goods or to other asset types in the next period. The indicator variables $\lambda_b \in \{0, 1\}$ and $\lambda_s \in \{0, 1\}$ turn on when he is forced to keep his stock or bond holdings. The forcing parameters can be interpreted as the inconvenience value of bonds and stocks over money, and the primary benefit of the forcing parameters is two-fold. On the one hand, I avoid introducing another state variable, on the other, the forcing parameters are continuous and allow the model capture a small period of forced holding when the discretized time interval is large. The forcing parameters also engender one source of endogenous persistence and give the model a cash-in-advance flavor without introducing complicated timing issues. Agents have motivation to hold money, even if they are not very risk adverse, to insure against idiosyncratic earnings risk - in the context of the quantity theory of money, this is known as the transaction motive. It may be argued that imposing fixed costs is a better approach than forced holding but the counterargument is that the data cannot separately identify the time cost and the monetary cost. The fixed costs in Chatterjee and Corbae, Alvarez, Atkeson, and Kehoe, and Khan and Thomas are necessary to produce endogenous market segmentation but my model already generates this feature with heterogeneous preferences. The assumption of fixed costs of trading is also probably not accurate for those who have a large amount of assets. Therefore the additional computational complexity that the incorporation of fixed costs creates is in my opinion not worth the questionable benefit.

The forcing variables are essentially a Calvo-like assumption. However, the benefit of the Calvo-like assumption comes without the normal problems associated with Calvo pricing when it is applied to firms. The time cost is often explicit for financial transactions. For instance, automated clearing house (ACH) transactions take days to clear, brokerage automated customer account transfers (ACAT) take weeks to clear, and many mutual funds and bond funds require investors to hold the funds for a long period of time (months to years) or face penalties. Also, there is frequently a trade-off between a time cost or a monetary cost to a transaction. And the monetary cost of a transaction is analogous to the menu cost of a price change by a firm.

The pseudo-life-cycle property of the model comes from the condition that the net financial assets of any agent are reset to zero with probability $\Phi \in [0, 1)$. The indicator variable $\phi \in \{0, 1\}$ turns on when the agent is reset. The reset is performed in between periods so that the agent's earnings process and steady state characteristics are not affected. One interpretation is that the agent dies in between periods and is immediately replaced by another person with the same earnings process. The reduced-form estimates indicate that age of the decision maker is perhaps the most important driving force behind household saving. In the model, an agent who has the tendency to save will tend to accumulate assets as he "ages" even when he is faced with the possibility of a reset. Another interpretation is that $\frac{1}{\Phi}$ is a measure of household financial myopia, i.e. households behave as though their financial assets will become useless and their debt will be wiped after $\frac{1}{\Phi}$ expected years. Depending on the magnitude of Φ, this friction may create substantial *ex post* inefficiency but the decision problem is still time-consistent. It should be noted that the net asset or debt after a person dies is assumed to be absorbed by the government. This assumption is not grievous since the total amount of money gained or lost from the economy is numerically insignificant due to the low estimates of Φ, and transferring the assets gained or lost back to the population adds considerable computational complexity. Overall, the reset parameter has a much greater effect on the decision rules.

Giving agents the capability to take on as much debt as they want, even in an environment with full commitment and with natural borrowing constraints, causes the mean debt in the model to be much higher than the mean debt in the data. The initial versions of the model did not have a debt constraint but it was clear that agents tended to borrow around an order of magnitude higher than the mean debt in the data. Therefore, the model now has a constraint $\delta \in [0, \infty)$ where a person can only take on a maximum debt of δe, where e is the current earnings realization. If the current earnings shock causes his debt constraint to drop below the debt carried over from the previous period, he still can retain his debt load or pay some of the debt off but he cannot take on more debt. The motivation for this type of restriction comes from the behavior of real world loan markets in the most basic sense. The simplest way of imposing a borrowing constraint is with just a lower bound for all agents like in Huggett (1993), but if the lower bound is binding, there will be a mass of agents on the lower bound and the debt distribution will be quite limited.

The household solves the functional equation below to maximize lifetime expected utility

$$v_i(\zeta, a, e, \theta_b, \theta_s, \lambda_b, \lambda_s, \phi; \mu_b, \mu_s, \mu_m, \mu_d, \Sigma_b^2 + \sigma_b^2, \Sigma_s^2 + \sigma_s^2, \iota, \Lambda_b, \Lambda_s, \Phi, \delta)$$

$$(3.2.1) \quad = \max_{(\zeta', a') \in \Gamma(\cdot;\cdot)} \begin{cases} u_i(c) + \beta_i \mathrm{E}[v_i(\zeta', a', e', \theta_b', \theta_s', \lambda_b', \lambda_s', \phi'; \cdot)|e] & \text{if } \zeta = b, \lambda_b = 0 \\ u_i(e) + \beta_i \mathrm{E}\left[v_i\left(b, \frac{\theta_b a}{\iota}, e', \theta_b', \theta_s', \lambda_b', \lambda_s', \phi'; \cdot\right) \middle| e\right] & \text{if } \zeta = b, \lambda_b = 1 \\ u_i(c) + \beta_i \mathrm{E}[v_i(\zeta', a', e', \theta_b', \theta_s', \lambda_b', \lambda_s', \phi'; \cdot)|e] & \text{if } \zeta = s, \lambda_s = 0 \\ u_i(e) + \beta_i \mathrm{E}\left[v_i\left(s, \frac{\theta_s a}{\iota}, e', \theta_b', \theta_s', \lambda_b', \lambda_s', \phi'; \cdot\right) \middle| e\right] & \text{if } \zeta = s, \lambda_s = 1 \\ u_i(c) + \beta_i \mathrm{E}[v_i(\zeta', a', e', \theta_b', \theta_s', \lambda_b', \lambda_s', \phi'; \cdot)|e] & \text{if } \zeta = m \\ u_i(c) + \beta_i \mathrm{E}[v_i(\zeta', a', e', \theta_b', \theta_s', \lambda_b', \lambda_s', \phi'; \cdot)|e] & \text{if } \zeta = d \end{cases}$$

where $\Gamma(\cdot;\cdot) = \Gamma(\zeta, a, e, \theta_b, \theta_s; \mu_b, \mu_s, \mu_m, \mu_d, \lambda, \Sigma_b^2 + \sigma_b^2, \Sigma_s^2 + \sigma_s^2, \Lambda_b, \Lambda_s, \Phi, \delta)$, and the functional equation is subject to the following constraints

$$(3.2.2) \quad c + a' \mathbf{1}_{(\zeta' \in \{b,s,m\})} + \frac{a'}{\mu_d} \mathbf{1}_{(\zeta'=d)} \leq e + \frac{\theta_b a \mathbf{1}_{(\zeta=b)} + \theta_s a \mathbf{1}_{(\zeta=s)} + \mu_m a \mathbf{1}_{(\zeta=m)} + a \mathbf{1}_{(\zeta=d)}}{\iota}$$

$$(3.2.3) \quad \min\{-\delta e, a\} \leq a' \mathbf{1}_{(\zeta'=d)}$$

if $\{\zeta, \{\lambda_b, \lambda_s\}, \phi\} \in \{\{b, s, m, d\} \times 0 \times 0\}$,

$$a' = \frac{\theta_b a}{\iota} \text{ or } a' = \frac{\theta_s a}{\iota} \tag{3.2.4}$$

if $\{\zeta = b, \lambda_b = 1, \phi = 0\}$ or $\{\zeta = s, \lambda_s = 1, \phi = 0\}$ respectively, and

$$c + a'\mathbf{1}_{(\zeta' \in \{b,s,m\})} + \frac{a'}{\mu_d}\mathbf{1}_{(\zeta'=d)} \leq e \tag{3.2.5}$$

$$-\delta e \leq a'\mathbf{1}_{(\zeta'=d)} \tag{3.2.6}$$

if $\phi = 1$.

Expanded out, the recursive formulation of the expected present value of choosing to hold bonds is,

$$\begin{aligned}
\beta_i \mathrm{E}\left[v_i(b, a', e', \theta_b', \theta_s', \lambda_b', \lambda_s', \phi'; \cdot)|e\right] &= \Phi\beta_i \mathrm{E}[v_i(b, 0, e', \theta_b', \theta_s', \lambda_b', \lambda_s', \phi'; \cdot)|e] \\
&+ (1-\Phi)\beta_i \Big((1-\Lambda_b)\mathrm{E}[v_i(b, a', e', \theta_b', \theta_s', \lambda_b', \lambda_s', \phi'; \cdot)|e] \\
&+ \Lambda_b \Phi \Big(\mathrm{E}[u_i(e')|e] + \beta_i \mathrm{E}[v_i(\zeta'', 0, e'', \theta_b'', \theta_s'', \lambda_b'', \lambda_s'', \phi''; \cdot)|e'] \Big) \\
&+ \Lambda_b (1-\Phi) \Big(\mathrm{E}[u_i(e')|e] + \beta_i \mathrm{E}[v_i(b, a'', e'', \theta_b'', \theta_s'', \lambda_b'', \lambda_s'', \phi''; \cdot)|e'] \Big) \Big)
\end{aligned}$$

The recursive formulation of the expected present value of choosing to hold stocks is the same except for a change in the forcing parameter and conditional probability, i.e. the appropriate b's should be replaced by s's. In contrast, the expansion of the expected present value of choosing to hold money is,

$$\begin{aligned}
\beta_i \mathrm{E}\left[v_i(m, a', e', \theta_b', \theta_s', \lambda_b', \lambda_s', \phi'; \cdot)|e\right] &= \Phi\beta_i \mathrm{E}[v_i(m, 0, e', \theta_b', \theta_s', \lambda_b', \lambda_s', \phi'; \cdot)|e] \\
&+ (1-\Phi)\beta_i \mathrm{E}[v_i(m, a', e', \theta_b', \theta_s', \lambda_b', \lambda_s', \phi'; \cdot)|e]
\end{aligned}$$

The recursive formulation of the expected present value of choosing to go into debt is similarly derived by replacing the appropriate m's by d's.

The resulting decision rules are $g_i(\zeta, a, e, \theta_b, \theta_s; \mu_b, \mu_s, \mu_m, \mu_d, \lambda, \Sigma_b^2+\sigma_b^2, \Sigma_s^2+\sigma_s^2, \Lambda_b, \Lambda_s, \Phi, \delta)$ $\in \{b, s, m, d\}$ and $h_i(\zeta, a, e, \theta_b, \theta_s; \mu_b, \mu_s, \mu_m, \mu_d, \lambda, \Sigma_b^2 + \sigma_b^2, \Sigma_s^2 + \sigma_s^2, \Lambda_b, \Lambda_s, \Phi, \delta) \in [0, \infty)$.

Definition 2. A stationary recursive competitive equilibrium is a set of decision rules and a set of individual and aggregate laws of motion such that the policy functions g_i and h_i solve the individual's optimization problem (1) subject to (2)-(6) for all i.

The main disadvantage of the discrete choice of asset type is that the households in the model do not hold a portfolio. In Section 6, I discuss in detail the impact computational limitations have on a more general framework and offer a comparison of the discrete choice approach to the popular two-stage approach. However, the following proposition shows that an agent does not necessarily need a portfolio option to hold different asset types at different states.

Proposition 3. *Suppose the economy has two periods and the agent can only work in the first period has no earnings in the second period. Let the mean return rates of asset types be ordered by* $\mu_m < \mu_b < \mu_s$ *and let the variances be ordered by* $0 = \sigma_m^2 < \Sigma_b^2 + \sigma_b^2 < \Sigma_s^2 + \sigma_s^2$*. Let the minimum return rate for all three asset types be censored at a value greater than zero. The means and variances should also satisfy the property* $P(r|s) < P(r|b) < P(r|m)$ *for* $r < R$*, i.e. the riskier asset types have a higher probability of realizing any return rate less than some bound* R*. If the discount factor is greater than zero and the utility function has the additional assumption that* $\lim_{c \to 0} u_i'(c) = -\infty$*, then for each individual there exists a partition of the wealth states with at least two nonempty blocks where the individual holds a different asset type in each block of the partition.*

Proof. The agent will clearly choose to hold assets into the second period since he would receive $-\infty$ utility otherwise. Without loss of generality, I will show that bonds or stocks can be held by the same individual depending on the wealth state. From Jensen's inequality I have that $\mathrm{E}[u_i(a')|b] < u_i(\mathrm{E}[a'|b]) = u_i(\mu_b a') < u_i(\mathrm{E}[a'|s]) = u_i(\mu_s a')$ by the strict concavity of u_i. I want to first prove that $\mathrm{E}[u_i(a')|b] < \mathrm{E}[u_i(a')|s]$ for $a' > \bar{A}$, and then I want to prove that $\mathrm{E}[u_i(a')|b] > \mathrm{E}[u_i(a')|s]$ for $a' < \underline{A}$.

Since the utility function is strictly increasing and strictly concave, I have that for any $\epsilon > 0$ there exists an $a' > \bar{A}$ such that $u_i(\mathrm{E}[a'|s]) - \epsilon < \mathrm{E}[u_i(a')|s]$. Let $\epsilon = u_i(\mathrm{E}[a'|s]) - \mathrm{E}[u_i(a')|b] > 0$ to get $u_i(\mathrm{E}[a'|s]) - (u_i(\mathrm{E}[a'|s]) - \mathrm{E}[u_i(a')|b]) = \mathrm{E}[u_i(a')|b] < \mathrm{E}[u_i(a')|s]$.

Since $\lim_{c \to 0} u_i'(c) = -\infty$, I have that for any finite N there exists an $a' < \underline{A}$ such that $\mathrm{E}[u_i(a')|s, r < R] - \mathrm{E}[u_i(a')|b, r < R] < N$. Let $N = \mathrm{E}[u_i(a')|b, r \geq R] - \mathrm{E}[u_i(a')|s, r \geq R]$ to get $\mathrm{E}[u_i(a')|s, r < R] - \mathrm{E}[u_i(a')|b, r < R] < \mathrm{E}[u_i(a')|b, r \geq R] - \mathrm{E}[u_i(a')|s, r \geq R] \implies \mathrm{E}[u_i(a')|b] > \mathrm{E}[u_i(a')|s]$. □

3.3. Data

The primary source of data is the 2007 Survey of Consumer Finances (SCF). The SCF reports the estimated amount of equity and liquid asset holdings for each household. The equity assets are composed of directly-held stocks, stock mutual funds (full value for mostly stock mutual funds, half value for combination mutual funds), individual retirement accounts and Keoghs invested in stock (full value if mostly invested in stock, half value if split between stocks and bonds or stocks and money market funds, one third value if split between stocks, bonds, and money market funds), other managed assets with equity interest such as annuities, trusts, and managed investment accounts (full value if mostly invested in stock, half value if split between two asset types), and thrift-type retirement accounts invested in stock (full value if mostly invested in stock, half value if split between stocks and interest earning assets). The liquid assets are composed of money market funds, checking accounts, savings accounts, call accounts, and physical money holdings. Certificates of deposit are unclassifiable in the model since it has both bond and money properties so I chose to ignore this asset with a relatively low total value.

From each household's equity estimate, I can compute an estimate for the total bond holdings by adding together directly-held bonds, directly-held stocks, pooled investment funds, retirement assets, and other managed assets and then subtracting the equity estimate. This procedure overestimates the bond holdings somewhat. Money market funds are likely the largest component in the assets that are erroneously included in the bond estimate. Overall, the slight overestimation of bond holdings should not have a large impact on the model since the erroneously included assets are likely to be interest bearing but the true return variance might be a bit lower.

Installment loans and credit card balances are the types of debt that are included in the debt total. Although mortgage debt is the largest component of household debt, it is not a financial asset and it usually does not impact short-term consumption decisions. A house as an asset is much less liquid and does not directly substitute with the other types of assets considered in the model. If mortgage debt and home equity values are introduced into the model, household beliefs on future home prices and time-cost parameters must also be incorporated. On the other hand, home equity lines of credit were also not included in the debt total since the SCF does not report how much of this credit line was actually used.

The net financial worth is found by adding up bond, stock, and money holdings and then subtracting the debt of each household. The households with negative net financial worth are necessarily debt choosers in the model. Table 1 presents summary statistics from the 2007 SCF. Table 2 reports a logistic regression of various regressors on the probability that a household has non-negative net financial worth. Age is the most significant predictor of non-negative financial worth and so it gives a strong motivation for the pseudo-life-cycle model. Earnings is also a significant factor which the model accounts for. The other regressors however are beyond the scope of the model. The households with non-negative financial worth are placed in the bond, stock, and money holder category depending on the dominant asset choice. Table 3 reports an ordered logistic regression of various regressors on the asset choice (ordered by increasing return risk). Here, the net financial worth is the most significant predictor of the dominant asset choice. The next two most significant regressors are age and earnings which again the model accounts for.

Main Asset	Fraction	Median Worth	Mean Worth
Bonds	0.1841	67,130	258,263
Stocks	0.2430	134,800	550,997
Money	0.2829	2,000	44,804
Debt	0.2899	-10,908	-19,834

TABLE 1. Asset choices and net financial worth from the 2007 SCF.

$fin\_worth \geq 0$	Estimate	Robust Std. Err.	z-stat	p-value
log(*earnings*)	0.0444	0.0176	2.52	0.012
age	0.0430	0.0031	13.79	0.000
educ	0.0755	0.0160	4.72	0.000
married	0.0384	0.0932	0.41	0.680
kids	-0.0597	0.0377	-1.58	0.113
white	-0.5323	0.2151	-2.47	0.013
black	-0.8242	0.2434	-3.39	0.001
hispanic	-0.5062	0.2576	-1.97	0.049
constant	-2.0449	0.4250	-4.81	0.000

TABLE 2. A logistic regression of various regressors on the probability that a household has non-negative net financial worth. The omitted race dummy variable is *other* which includes Asians and Pacific Islanders as the largest component.

asset_choice	Estimate	Robust Std. Err.	z-stat	p-value
log(fin_worth)	0.5880	0.0262	22.42	0.000
log(*earnings*)	0.0720	0.0252	2.85	0.004
age	-0.0229	0.0034	-6.75	0.000
educ	0.0358	0.0207	1.73	0.083
married	-0.0463	0.1025	-0.45	0.652
kids	-0.0091	0.0468	-0.19	0.846
white	0.4186	0.2368	1.77	0.077
black	0.2397	0.2775	0.86	0.388
hispanic	0.1178	0.2897	0.41	0.684
Money/Bonds Cut	5.6607	0.5243		
Bonds/Stocks Cut	7.2858	0.5349		

TABLE 3. An ordered logistic regression of various regressors on the asset choice (ordered by increasing return risk). The omitted race dummy variable is *other* which includes Asians and Pacific Islanders as the largest component.

The 2005-2007 Panel Study of Income Dynamics (PSID) is then used to find the earnings transition. All sources of earnings for each household are included such as labor, foodstamp, welfare, pension, retirement, social security, and annuity income. Including all these sources of income substantially reduces the zero income problem of the model when the utility functions have a CRRA form. The earnings process is then discretized to transitions between the mean household earnings of each decile as shown in Table 4.

Decile Number	Mean Earnings	Decile Number	Mean Earnings
1st decile	3,445	6th decile	46,742
2nd decile	12,748	7th decile	59,094
3rd decile	20,294	8th decile	74,195
4th decile	28,054	9th decile	96,675
5th decile	36,549	10th decile	191,174

TABLE 4. The mean inflation adjusted earnings of households in each decile from the 2005 and 2007 PSID.

All the households weights of the SCF and PSID were taken into account with great care. The PSID weights were not constant from 2005-2007 so they were averaged for each household.

From 2002 to 2007, the inflation remained in the 2% to 3% range. The mean inflation was 2.69% and so this is the forward inflation estimate for the households in the model.

The mean bond return and the mean stock return is set to the mean annual return of the 2002-2007 Barclays Capital U.S. Aggregate Bond Index[1] and the mean annual return of the 2002-2007 Russell 3000 Index respectively as shown in Table 5. Similarly, the standard deviation of the bond return and the standard deviation of the stock return is set to the annual standard deviations of the corresponding indices. After the initial burn-in period to generate an approximate steady state, the final few periods of the bond and stock returns are set to the actual annual returns of the 2002-2007 Barclays Capital U.S. Aggregate Bond Index and the 2002-2007 Russell 3000 Index. The argument for using this specific period is that the start of the period is close to the end of the dot-com crash. The economy was

[1]Formerly known as the Lehman Brothers Aggregate Bond Index.

relatively stable and "normal" during this period since there was a mild recession as well as a moderate recovery. The mean bond and stock returns during this period are also not too different from postwar mean bond and stock returns but should offer a better estimate of household beliefs.

Period	Bond Return	Stock Return	Inflation
2002-2003	1.1025	0.7846	1.0162
2003-2004	1.0410	1.3106	1.0215
2004-2005	1.0414	1.1195	1.0284
2005-2006	1.0280	1.0779	1.0334
2006-2007	1.0446	1.1386	1.0326
2007-2008	1.0665	1.0519	1.0294
Mean	1.0540	1.0805	1.0269
Std. Dev.	0.0269	0.1710	0.0067

Table 5. Bond and stock returns from the 2002-2007 Barclays Capital U.S. Aggregate Bond Index and Russell 3000 Index as well as 2002-2007 inflation.

3.4. Identification

- The parameters μ_β, σ_β, and δ are jointly identified by the fraction of people in debt, the median amount of debt, and the mean amount of debt. The debt constraint first pulls the debt level of the households to the correct range. The distributional parameters for the discount factor μ_β and σ_β determine the fraction of the population that borrow and the distribution of debt (i.e. the median-mean ratio) respectively. Households with very low discount factors have an inclination to borrow as much as possible while households with moderately low discount factors have an inclination to borrow only when a low earnings shock is received.
- The distributional parameters for the coefficient of risk aversion μ_γ and σ_γ are jointly identified by the fraction of people with stock holdings, the median amount of stock holdings, and the mean amount of stock holdings. If the discount factor is fixed at a relatively high value, those who are less risk averse are more likely to choose stocks over bonds or money.

- The forcing parameters Λ_b and Λ_s are jointly identified by the fraction of people who hold money, the median amount of money holdings, and the mean amount of money holdings. Without the forcing parameters, the low return rate of money relative to bonds and stocks causes a very small fraction of households to choose money unless the risk aversion parameter is ridiculously high. In contrast, with the forcing parameters, this equity and bond premium puzzle almost disappears entirely. The characteristics of bond and stock holders are generally different so when there are two types of forcing parameters, the households that switch to become money holders can come from any combination of the original bond holder population and the original stock holder population. The diverse set of households which substitute money for bonds and stocks help generate the broad distribution of money holders. Finally, as mentioned in Section 2, the fixed costs of market transactions cannot be separately identified from the forcing parameters so that including both types of costs into a model is not necessarily beneficial.
- The reset parameter Φ is identified by the mean amount of bond holdings. The households that choose bond holdings tend to have high discount factors and high risk aversion coefficients, and as a result, they accumulate large amounts of bond holdings which are far beyond the mean bond holdings in the data. Imposing a hard ceiling on the amount of bond holdings will easily reduce the mean but the ceiling will also distort the distribution of bond holdings. Again, this is why the pseudo-life cycle property of the model is attractive.
- If there are a lot of wealthy people who choose bonds or money (high β and high γ), and poor people who choose stocks or borrow (low β and low γ), ρ will likely be positive. If there are a lot of wealthy people that choose stocks (high β and low γ), and poor people that choose bonds or money (low β and high γ), ρ will likely be negative. I am left with the median bond holdings as the only unused moment but the identification of ρ comes from a lot of other sources as well (i.e.

the overindentification conditions above). The correlation coefficient arguably has the weakest specific identification by a moment; however, the overall identification may be quite strong.

3.5. Estimation and Results

Let the utility function have the specific form,

$$u_i(c) = \begin{cases} \frac{c^{1-\gamma_i}-1}{1-\gamma_i} & \text{if } \gamma_i \neq 1 \\ \ln(c) & \text{if } \gamma_i = 1 \end{cases}$$

Then let the discount factor be left truncated at $\underline{\nu}_\beta = 0$ and right truncated at $\bar{\nu}_\beta = 0.99$, and let the coefficient of relative risk aversion be left truncated at $\underline{\nu}_\gamma = 0$ and right truncated at $\bar{\nu}_\gamma = 10$. The right truncation of the CRRA is necessary to ensure the utility functions are numerically well-defined.

The idiosyncratic bond and stock return variance is set to $\sigma_b^2 = \sigma_s^2 = 0$. The idiosyncratic risk would ideally be nonzero; however, the data does not identify these parameters well since an increased return risk does not alter the mean and median of a symmetric distribution of returns.

The debt price $\frac{1}{\mu_d}$ is set to the inverse of the mean bond return rate $\frac{1}{\mu_b}$ and the return rate of money is set to $\mu_m = 1$.

The simulated method of moments uses the eleven moments (bond holder fraction, stock holder fraction, money holder fraction, bond median, stock median, money median, debt median, bond mean, stock mean, money mean, and debt mean) to estimate the nine parameters (μ_β, σ_β, μ_γ, σ_γ, ρ, Λ_b, Λ_s, Φ, and δ). In Section 4, it was easier to talk about identification in terms of the debt holder fraction but the bond holder fraction was actually used in the estimation.

3.5.1. SMM. Table 6 reports the results of the SMM estimation and it can be seen that the model moments are very close to the data moments especially for an overidentified system.

Statistic	Target	Model	Parameter	Estimate
Bond holder fraction	0.1841	0.2014	μ_β	1.1611
Stock holder fraction	0.2430	0.2862	σ_β	2.8956
Money holder fraction	0.2829	0.1466	μ_γ	0.9707
Debt holder fraction*	0.2899	0.3658	σ_γ	0.2180
Bond median	0.6713	0.7155	ρ	0.9098
Stock median	1.3480	1.2065	Λ_b	0.3525
Money median	0.0200	0.0487	Λ_s	0.2909
Debt median	0.1091	0.2530	Φ	0.0175
Bond mean	2.5826	1.9603	δ	0.3984
Stock mean	5.5100	2.2339		
Money mean	0.4480	0.9677		
Debt mean	0.1983	0.2830		

TABLE 6. Results from overidentified simulated method of moments estimation. The asset values are measured in hundreds of thousands. *This statistic was left out of the estimation due to colinearity but reported here for the sake of completeness.

Asset Choice	β Mean	β Std. Dev.	γ Mean	γ Std. Dev.
Bonds	0.9064	0.0511	1.5797	0.6083
Stocks	0.9389	0.0369	1.6087	1.0152
Money	0.9374	0.0408	1.4698	0.7663
Debt	0.8611	0.0705	0.8555	0.3593

TABLE 7. The distribution of discount factors and risk aversion coefficients conditional on the asset choice.

Asset Choice	Earnings Median	Earnings Mean	Earnings Std. Dev.
Bonds	1.4731	1.0508	0.4819
Stocks	0.4193	0.6242	0.4326
Money	0.0819	0.2274	0.2421
Debt	0.2401	0.3823	0.3893

TABLE 8. The distribution of earnings conditional on the asset choice. The earnings are measured in hundreds of thousands.

3.5.2. Simulated Asset Flows. The impact of quantitative easing on household asset flows is an important question. Without considering all the general equilibrium effects, quantitative easing from the household's point of view is essentially a one time boost in the bond return. At the same time, it causes a lower expected future bond yield and a higher expected future inflation. Heterogeneity in preferences is arguably the most natural way to determine how bond holders will react by possibly substituting other asset types for bonds. The households that marginally preferred bonds over stocks prior to quantitative easing might switch over to stocks while the households that marginally preferred bonds to money might switch over to money. Without heterogeneity, a large mass of households will behave similarly but it is clear from the data that the responses are much more diverse.

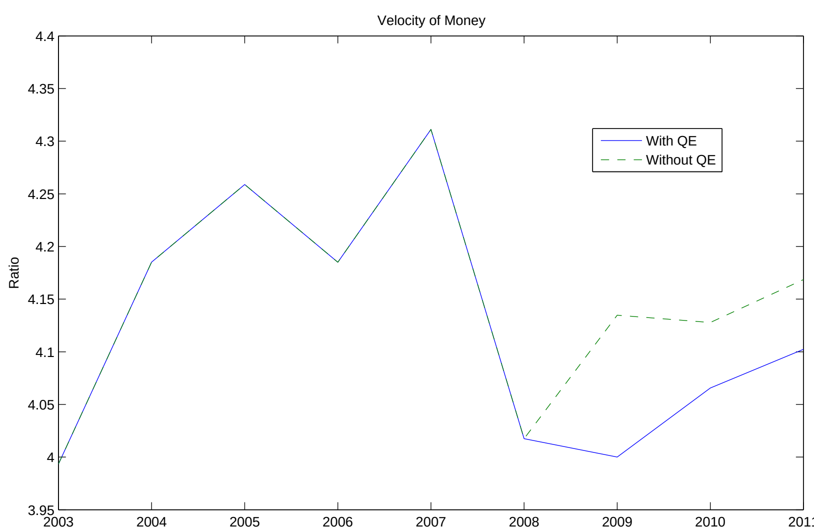
The bulk of the first round of quantitative easing in the United States occurred from the end of 2008 to the beginning of 2009. The bond return rate from the first quarter of 2008 to the first quarter of 2009 is 1.1236. I suppose that all the bond returns above the mean bond return of 2002-2007 is attributed to quantitative easing. Therefore 0.0696 is the one time boost in the bond return. I also suppose that the future bond return belief is decreased by 1% and the future inflation belief is increased by 1%. All this is hypothetical but it should provide some intuition on the behavior of the households. Using the model, I simulate the asset decisions with and without quantitative easing to compose Table 9.

Statistic	End of 2007	End of 2008 without QE	End of 2008 with QE
Bond holder fraction	0.2014	0.2026	0.1611
Stock holder fraction	0.2862	0.2885	0.3050
Money holder fraction	0.1466	0.1407	0.1602
Debt holder fraction	0.3658	0.3682	0.3737
Bond median	0.7155	0.6280	0.0557
Stock median	1.2065	1.2065	1.5396
Money median	0.0487	0.0894	0.1925
Debt median	0.2530	0.2530	0.2530
Bond mean	1.9603	1.9870	1.3211
Stock mean	2.2339	2.1941	2.5528
Money mean	0.9677	1.0217	1.2797
Debt mean	0.2830	0.2821	0.2796

TABLE 9. Simulated asset flows due to quantitative easing.

It can be seen that the households with a large amount of bond holdings switch over to equities when there is quantitative easing. Therefore quantitative easing increases the fraction of households that hold stocks as well as the mean stock amount since it is the rich households that convert bonds to stocks. The households in the previous period that held bonds received a large positive shock but holding bonds again is unattractive due to the lower relative return and due to the concavity of the utility function. There are also more households that take up debt because the debt price is set at $\frac{1}{\mu_b}$ and μ_b decreases by 1%. Inflation on the other hand increases by 1% which makes taking on debt even more attractive.

3.5.3. Velocity of Money. The average 2008-2011 bond yield dropped to 3.77% and the average bond return (yield + capital gains) during the same period was 6.04%. I use the new yield as the new expectation on the bond return following quantitative easing - leaving all else fixed. The graph below indicates that with quantitative easing, the velocity of money drops significantly.



The following table offers a more precise look at how the bond holders adjust their asset choices from 2008 to 2011.

Statistic	Without QE	With QE
Old mean bond holdings	1.9603	1.9603
New bond fraction	0.5987	0.3308
New stock fraction	0.1925	0.4552
New money fraction	0.0954	0.1006
New debt fraction	0.1135	0.1135
New mean holdings	2.0735	2.6547
New mean money holdings	2.1790	2.9067

Table 10. The bond holder behavior where the new statistics are computed for households which were bond holders in 2008.

Note that without quantitative easing, a higher fraction of households remain bond holders in 2011. In contrast, with quantitative easing, households hold greater net assets due to the fiscal stimulus and the money holdings are particularly high when compared with the normal average. This suggests that the original bond holders continue to hold a large amount of assets when they become money holders thereby reducing velocity.

3.6. Computation

Even though I observe conditional choice probabilities, I do not pursue the faster conditional choice probability (CCP) framework of Hotz and Miller (1993) or the nested pseudo-likelihood (NPL) framework of Aguirregabiria and Mira (2002) because the joint distribution of β_i and γ_i conditional on the agent's asset choice is unknown. The CCP and NPL estimators require the integration over the conditional joint distribution of β_i and γ_i. I would need to run simulations to approximate the conditional joint distribution but doing so requires the decision rules. The decision rules of course can only be computed by integration over the conditional joint distribution of β_i and γ_i.

For the popular two-stage approach as in Erosa and Ventura (2002), the individual only has to determine the amount of assets to carry into the next period. Then a second stage allows the individual to choose a portfolio. This dramatically simplifies the problem and reduces the intertemporal choice to the standard dynamic programming problem. However, it should be emphasized that the two-stage approach is still an approximation since the individual might like to choose a portfolio with a different total value in each state. Requiring the individual to perform a two-stage choice may also greatly reduce *ex post* efficiency. It should be noted that my model with forcing parameters is poorly suited to the two-stage framework since a portfolio would need to be carried intertemporally when an agent is forced to hold bonds or stocks.

Producing a portfolio with the traditional dynamic programing methods is nearly impossible without methodological advances. Even with two asset types and just 100 grid points, the value function for each individual will have 10 thousand elements and the transition matrix will therefore have 100 million elements. Solving this value function once is difficult enough for most computers. The model in this paper requires four asset types and a distribution of agents each with a different value function. Then the simulated method of moments outer loop is iterated thousands of times to find the minimum error value.

The computational burden for SMM is significant since each agent will have a different β_i and γ_i but it is highly parallelizable. I parallelize the program with a message passing interface (MPI) and use optimized basic linear algebra subprograms (BLAS) to speed up vectorized code. Around one thousand combined CPU core hours were needed to find the parameter estimates.

3.7. Conclusion

On the surface, quantitative easing seems like a promising monetary tool to simultaneously create liquidity in the financial markets and stimulate the economy when interest rates are close to zero. However, I show that the population that is targeted by quantitative easing is especially disinclined to generate money velocity. Therefore, monetary policies aimed at other segments of the population may be more effective.

With a simple framework and a careful treatment of data this paper is able to rigorously model complex financial market behavior. However, there is much left to be done in analyzing the general equilibrium spillover effects of quantitative easing on the government, firms, and households. Cross country effects of the financial markets may also be an area that will yield fruitful future research.

References

Adda, Jerome and Russell Cooper, *Dynamic economics: quantitative methods and applications*, The MIT Press, 2003.

Aguirregabiria, V. and P. Mira, "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, 2002, *70*, 1519–1543.

Alvarez, F., A. Atkeson, and P. J. Kehoe, "Money, Interest Rates, and Exchange Rates with Endogenously Segmented Markets," *Journal of Political Economy*, 2002, *110*, 73–112.

Cameron, T. A. and G. R. Gerdes, "Discounting versus risk aversion: their effects on individual demands for climate change mitigation," *mimeo*, 2007.

Chatterjee, S. and D. Corbae, "Endogenous Market Participation and the General Equilibrium Value of Money," *Journal of Political Economy*, 1992, *100*, 615–646.

Cicchetti, C. J. and J. A. Dubin, "A Microeconometric Analysis of Risk Aversion and the Decision to Self-Insure," *Journal of Political Economy*, 1994, *102*, 169–186.

Erosa, A. and G. Ventura, "On Inflation as a Regressive Consumption Tax," *Journal of Monetary Economics*, 2002, *49*, 761–795.

Friedman, M., "The Demand for Money: Some Theoretical and Empirical Results," *Journal of Political Economy*, 1959, *67*, 327–351.

Gould, J. P. and C. R. Nelson, "The Stochastic Structure of the Velocity of Money," *American Economic Review*, 1974, *64*, 405–418.

Guvenen, F. and B. Kuruscu, "Does Market Incompleteness Matter for Asset Prices?," *Journal of the European Economic Association*, 2006, *4*, 484–492.

Hetzel, R. L. and Y. P. Mehra, "The Behavior of Money Demand in the 1980s," *Journal of Money, Credit and Banking*, 1989, *21*, 455–463.

Hodrick, R. J., N. Kocherlakota, and D. Lucas, "The Variability of Velocity in Cash-in-Advance Models," *Journal of Political Economy*, 1991, *99*, 358–384.

Hotz, V. J. and R. A. Miller, "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, 1993, *60*, 497–529.

Huggett, M., "The risk-free rate in heterogeneous-agent incomplete-insurance economies," *Journal of Economic Dynamics and Control*, 1993, *17*, 953–969.

Khan, A. and J. K. Thomas, "Inflation and Interest Rates with Endogenous Market Segmentation," *mimeo*, 2009.

Kiyotaki, N. and R. Wright, "On Money as a Medium of Exchange," *Journal of Political Economy*, 1989, *97*, 927–954.

Lucas, R. E., "Two Illustrations of the Quantity Theory of Money," *American Economic Review*, 1980, *70*, 1005–1014.

Mankiw, N. G. and L. H. Summers, "Money Demand and the Effects of Fiscal Policies," *Journal of Money, Credit and Banking*, 1986, *18*, 415–429.

3.8. Appendix

3.8.1. Computational Algorithm.

(1) Initialize the asset grids and come up with initial guesses for the parameters.

(2) In the SMM loop, first separately discretize the truncated distribution for the discount factor $\beta_i \sim \mathcal{T}\left(\mu_\beta, \sigma_\beta^2, \nu_\beta\right)$ and risk aversion $\gamma_i \sim \mathcal{T}\left(\mu_\gamma, \sigma_\gamma^2, \nu_\gamma\right)$ using the equal areas approach. The cumulative distribution function for the truncated normal distribution is

$$F(x;\mu,\sigma^2,\nu_a,\nu_b) = \frac{\Phi\left(\frac{x-\mu}{\sigma}\right) - \Phi\left(\frac{\nu_a-\mu}{\sigma}\right)}{\Phi\left(\frac{\nu_b-\mu}{\sigma}\right) - \Phi\left(\frac{\nu_a-\mu}{\sigma}\right)}$$

where Φ is the normal CDF. This procedure constructs a two dimensional grid of the truncated bivariate normal distribution for β_i and γ_i when $\rho = 0$.

(3) When $\rho \neq 0$, the bivariate normal CDF will be needed to find the probability that an agent will land in each cell. There does not exist an analytical form for the

bivariate normal CDF so it has to be approximated using the bivariate normal PDF

$$f(x, y; \mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho) =$$
$$\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right]\right)$$

(1) Compute the interpolated return probabilities for each asset type using the parameters μ_b, μ_s, μ_m, μ_d, Σ_b^2, Σ_s^2, σ_b^2, σ_s^2, and ι.

(2) Solve for the value functions and decision rules using (1)-(6).

(3) Simulate the economy for a burn in period using the mean return rates for all asset types.

(4) Simulate the final few periods of the bond and stock returns using actual data from the 2002-2007 Barclays Capital U.S. Aggregate Bond Index and the 2002-2007 Russell 3000 Index.

(5) Compute the model moments and find the error value $(M_m - M_d)'\hat{W}(M_m - M_d)$ where M_m are the model moments, M_d are the data moments, and $\hat{W}$ is the weighting matrix.

(6) Use a different guess and go back to Step 1 until the minimum error value is found with a global optimization algorithm.