Credit vs. demand constraints: the determinants of US firm-level investment over the business cycles from 1977 to 2011

Christian Schoder*
The New School for Social Research

March 21, 2012

Abstract

The paper studies empirically how demand and credit constraints affect US firm-level investment expenses over the cycle. A dynamic econometric specification of capital accumulation including sales growth, Tobin's q, the cash flow-capital ratio and the cost of capital as covariates is fitted by a rolling window System GMM estimator using quarterly Compustat data on publicly traded US corporations from 1977 to 2011 in order to obtain time-varying coefficients. We find that the demand related measures, sales growth and Tobin's q, have strongly pro-cyclical effects on investment, whereas this does not hold for cash flow or the cost of capital. Our results suggest that investment was (a) unconstrained in the 1981 recession, (b) demand constrained in the 1982 recession, (c) demand constrained during the 1990 downturn, (d) credit constrained in the 1995 and 1998 investment stagnations, (e) demand-and-credit constrained in the 2001 recession, and (f) demand constrained in the recent 2009 recession.

Keywords: investment, credit constraints, business cycles, panel estimation, System GMM
JEL Classification: D22, E22, G32

*Address: Department of Economics, The New School for Social Research, 6 East 16th Street, New York, NY 10003. Email: schoc152@newschool.edu.

1 Introduction

To overcome the recession following the financial crisis of 2007-09, US authorities have pursued policy measures aimed at maintaining a steady flow of credit from financial markets to businesses mainly through quantitative easing, the implementation of new lending facilities and the purchase of toxic assets by the public. Despite an aggressive interest rate policy as well as a stimulus package, typically Keynesian economists have expressed concerns that demand management has not been pursued to a sufficient extent as non-financial business investment declined tremendously in 2009/10.

The debate on the effectiveness of fiscal stimulus packages is primarily based on estimations of the multiplier effects of government spending.[1] While this has the advantage of analyzing the overall impact of fiscal policy on GDP growth, it has a severe downside: The multiplier effect is typically not identified as government spending responds to output growth. This issue contributes to the inconclusiveness of the studies conducted which manifests in a wide range of estimated multipliers.[2] Therefore, the present paper seeks to follow a different route and analyzes how investment responded to changes in demand and financial market conditions in recent periods of economic distress. The elasticity of investment with respect to demand and credit market conditions allows for some qualitative inference on the marginal effectiveness of fiscal and monetary policy.

Naturally, the effectiveness of demand stabilizing and credit-flow sustaining policies depends on the reasons why business investment declines. Measures that increase the willingness of the banking sector to lend may proof ineffective if firms do not want to expand their capital stock due to large spare capacities and low demand expectations, a situation we refer to as *demand-constrained* investment. On the other hand, expansionary conventional

[1] See, among others, Romer and Bernstein (2009), Cogan et al. (2010) and Christiano et al. (2009) on the multiplier effects of the stimulus packages during the recent crisis.

[2] While Romer and Bernstein (2009) estimate a multiplier of 1.6 for the *American Recovery and Reinvestment Act*, Cogan et al. (2010) come up with an estimate below one.

monetary policy as well as fiscal policy may involve small multiplier effects on investment if firms' access to external finance is constrained, i.e. if they are *credit constrained*. Moreover, both demand and credit constraints on investment may be binding which requires a policy mix of both demand stabilizing and credit-flow sustaining policy measures in order to be effective.

The emphasis of policy on credit-market conditions follows from the prevailing paradigm in investment theory which derives credit constraints on investment from information asymmetries on the capital market which drive a wedge between the costs of internal and external funds (Greenwald et al. 1984; Myers and Majluf 1984; Jensen and Meckling 1976). Problems of *moral hazard* arise as debtors have an incentive to engage in riskier investments once they have received the external funds. A risk premium arises which is reinforced by *adverse selection*, i.e. the displacement of risk averse investors relative to risk taking investors. Since the agency cost of investment which reflects the cost difference between external and internal funds is inversely related to a firm's net worth, the well-known *financial accelerator* theory pioneered by Bernanke and Gertler (1989) predicts that credit constraints exhibit a strong counter-cyclical pattern.

Introducing adjustment costs to the model, investment demand is argued to be determined by the expected marginal profitability of investment which is completely characterized by Tobin's marginal q, i.e. the ratio of the market value of an additional unit of capital to its replacement cost (Tobin 1969). Due to the unsatisfying performance of Tobin's q in explaining investment, empirical studies often include accelerator terms such as current sales as a proxy for demand expectations and capacity utilization in the investment function (cf. Fazzari et al. 1988).

Even though credit constraints may plausibly move counter-cyclically, they are irrelevant for economic policy if they are not binding and investment is constrained by low demand expectations. In fact, the possibilities of demand and credit constraints give rise to four different

investment regimes: *demand constrained*, *credit constrained*, *demand-and-credit constrained* and *unconstrained*. This paper seeks to empirically identify these investment regimes for the US corporate business sector since 1977 as well as to analyze the policy measures taken during the various recessions to manage demand and maintain the flow of credit in the context of the investment regimes identified.

We motivate different investment regimes by the aid of a simple asymmetric information model of investment. We argue that sales growth and Tobin's q are appropriate variables reflecting investment opportunities. Further, cash flow is directly related to net worth and, therefore, inversely related to agency costs. We argue that, in practice, the cash-flow effect on investment is a good measure for the firms' credit constraints.

To study how the investment regimes prevailing in the US since the mid-1970s changed with the business cycle and over time, we estimate a dynamic linear investment function for a panel of US corporate businesses including sales growth, Tobin's q, cash flow and the cost of capital as covariates. A rolling window regression with a width of 5 quarters has been applied in order to track the changes in the coefficients over time. Since the time dimension of the window is small while the individual dimension is large, we fit the dynamic model using the System GMM estimator for panel data developed by Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998) as traditional fixed effects and pooled OLS estimators are inconsistent.

Based on the relative movement of the time-varying coefficients, we are able to fully characterize the prevailing investment regime at any point in time. The results suggest that investment was (a) *unconstrained* in the 1980 recession, (b) *demand constrained* in the 1981 recession, (c) *demand constrained* during the 1990 downturn, (d) *credit constrained* in the 1995 and 1998 investment stagnations, (e) *demand-and-credit constrained* in the 2001 recession, and (f) *demand constrained* in the recent 2009 recession. This view is consistent with the chronicles of US fiscal and monetary policy stance regarding the management of

aggregate demand and credit flow. Our policy evaluation implies that the policy attempts to stabilize demand were insufficient in order to stabilize investment in the recent economic crisis.

Traditionally, empirical studies of credit constraints of investment analyze how cash-flow coefficients differ between sub-groups of firms grouped according to their probability of facing liquidity constraints. The seminal study by Fazzari et al. (1988) classifies firms according to their dividend payout policy and find that firms with low dividend payout rates which can be expected to face stronger credit constraints tend to exhibit higher cash-flow elasticities of investment. Here, we do not follow this line of research because of two reasons: First, the interest of this paper is to study how investment regimes have changed over the cycles which is only possible by analyzing the elasticities of investment along the time dimension. Second, theoretical considerations and empirical evidence put forward, among others, in the discussion of the study attached to Fazzari et al. (1988) and by Schoder (2011) suggest that grouping on a priori grounds may have some downsides. Results may suffer from a selection bias as they are typically highly sensitive to the choice of the sample selection criterion as well as the sample of firms considered.[3]

Yet, one weakness of the time-varying approach pursued here is the fact that, in general, the sensitivity of investment with respect to cash flow does not necessarily reflect the firm's credit constraints. As shown by Kaplan and Zingales (1997) an inverse relationship may potentially arise. However, we argue that the assumptions of a concave average investment demand function and a convex average finance supply function are plausible and sufficient to establish a direct relationship between the cash-flow parameter and credit constraints.

The remainder of the paper proceeds as follows. Section 2 illustrates the underlying

[3]As argued in the discussion attached to the article by Fazzari et al. (1988), an endogeneity issue arises as firms with good investment prospects which may be partly reflected by a large cash flow are likely to choose a low dividend payout rate. Schoder (2011) finds empirical evidence that the differences between the cash-flow coefficients across groups formed based on a priori assumptions about liquidity constraints such as size and dividend payout policy are not robust.

theoretical investment model as well as the constellations of the investment demand and finance supply curves which give rise to different investment regimes. It also discusses the econometric model used for the empirical analysis. Section 3 outlines the System GMM estimator for panel data which is the econometric methodology applied to estimate the dynamic investment specification in a recursive manner. In section 4, the data set used and variables computed are discussed. Section 5 presents the estimation results as well as the empirical investment regimes identified. Section 6 interprets the investment regimes found in the context of the historical record of monetary and fiscal policy aimed at the management of aggregate demand and credit flows. Section 7 concludes the paper.

2 An econometric model of investment

2.1 A simple model of investment

To motivate our econometric specification we briefly conduct a graphical analysis of the links between capital investment, agency cost of external finance and investment opportunities based on a simple investment model with asymmetric information between lenders and borrowers in the vain of Gertler and Hubbard (1988) and Bernanke and Gertler (1989). Figure 1 illustrates a simple capital market characterizing the model outlined in Appendix A, i.e. an investment demand curve and two variants of the supply of funds curve which an average firm out of a large sample faces.

The demand curve DD reflects an inverse relationship between the marginal revenue of investment and the desired capital expansion. For reasons discussed below, we assume a production technology which implies the demand curve to be concave. The curve SS^p relates the supply of funds to its marginal cost on a perfect capital market. Since the firm is small, it can borrow whatever desired at the given risk-adjusted real interest rate. Given its demand curve, the firm's optimal investment I^p. SS^f describes the supply curve in an imperfect

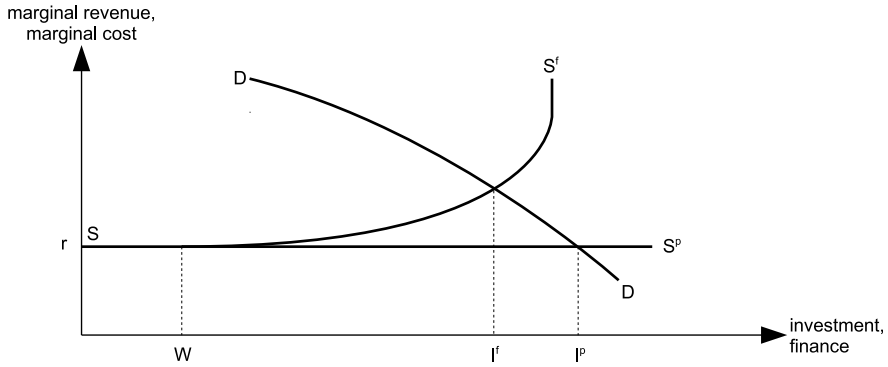

Figure 1: Capital market

capital market with information asymmetries between lenders and borrowers. It implies that firms prefer internal finance over external finance. As long as the desired investment is lower than the firm's internal funds W, the real interest rate applies as an opportunity cost. Beyond W, the firm has to acquire external funds. Due to asymmetric information additional costs of external funds arise which are increasing in the amount borrowed.[4] The supply curve on an imperfect capital market is increasing and convex as shown in Appendix A. The equilibrium in this market is $I^f < I^p$.

A rise (fall) in investment opportunities, i.e. the expected profitability of an additional unit of capital, shifts the demand curve to the right (left). A rise (fall) in the internal funds shifts the supply curve to the right (left). Note that a change in the internal funds affects the equilibrium investment only of firms facing credit constraints.

The investment-internal funds sensitivity, $\frac{\partial I^f}{\partial W}$, can be taken as a measure for the extent to which firms are *credit constrained* as long as it is monotonically decreasing in the level of W. Otherwise, a firm with little internal funds facing large borrowing costs may appear less

[4]One justification for the external finance premium is provided by information economics. As argued by Greenwald et al. (1984) and Myers and Majluf (1984), creditors are not perfectly aware of the debtors' investment intentions and the associated risks. Under asymmetric information debtors may have an incentive to engage in riskier investments once they have received the external funds (moral hazard). As managers are not fully liable, they may tend to risky investments which may reduce the value of the firm (Jensen and Meckling 1976). Marking up the interest rate by a premium to compensate for that risk and information costs displaces risk averse investors and leaves risk taking investors (adverse selection).

credit constrained than a firm with large internal funds facing low borrowing costs (cf. Kaplan and Zingales 1997; Fazzari et al. 2000). However, a concave demand curve combined with a convex supply curve which are implied by the investment model outlined in Appendix A are sufficient to ensure that the investment-internal funds sensitivity is monotonically decreasing in W.

2.2 The determinants of investment over the cycle

Studying the cyclical behavior of investment within the graphical framework outlined above yields valuable insights. In a boom, expectations improve and investment opportunities rise increasing the demand for investment at any interest rate. In the upswing, firms tend to rise their profit margins expanding their internal funds which shifts the supply curve to the left. Moreover, tightening of monetary policy causes the market interest rates to go up shifting the supply curve upwards. In the downswing, usually the opposite is observed.

Given the slopes and curvatures of the demand and supply curves different investment regimes can arise. Figure 2 illustrates different stylized demand and supply constellations which give rise to *credit-constrained*, *demand-constrained*, *demand-and-credit-constrained*, and *unconstrained* investment regimes.

Panel (a) represents a *credit-constrained* regime with firms willing to invest but facing liquidity constraints. The demand curve is fairly flat implying a high interest elasticity of investment demand. The demand schedule cuts the supply curve at A where the latter is fairly steep. This regime has the following implications: First, a rise in investment opportunities which may be triggered by expansionary fiscal policy is rather ineffective in raising investment which moves the new equilibrium to B. A severe credit crunch arises as the additional investment demand is not met by additional supply. Second, lowering the cost of capital by expansionary monetary policy is only moderately effective as the relevant supply segment is rather inelastic to changes in the interest rate. The equilibrium moves to C.

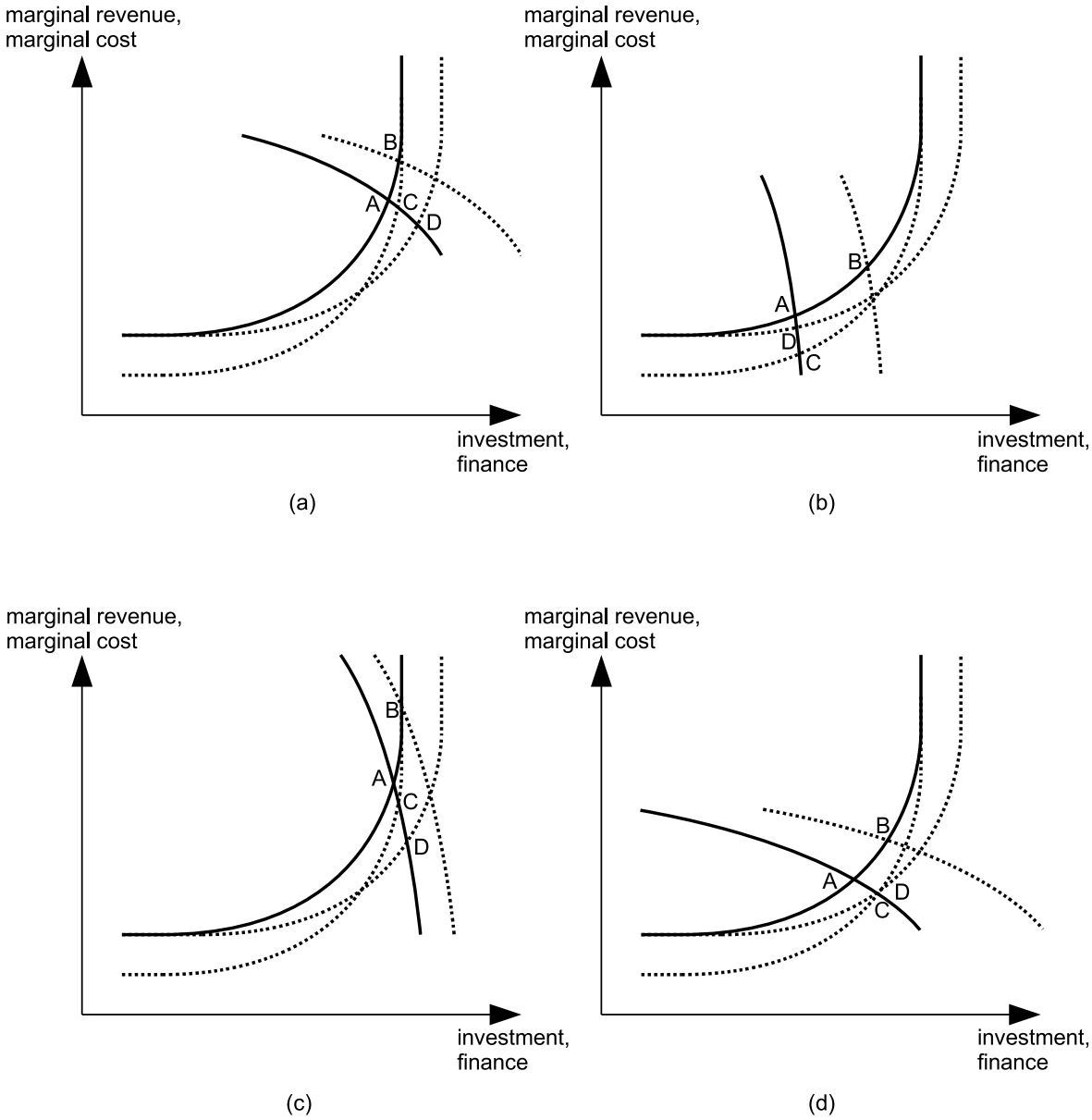

Figure 2: Different investment regimes: (a) *credit-constrained*, (b) *demand-constrained*, (c) *demand-and-credit-constrained*, and (d) *unconstrained* investment

Third, measures aimed at expanding the firms cash flow, for instance, through tax cuts move the supply curve to the left and are highly effective in fostering investment. The same holds for measures seeking to lower the agency cost of external finance through conventional and

unconventional monetary policy which stretches the increasing segment of the supply curve and makes it flatter. Together both measures imply a new equilibrium at D.

Panel (b) illustrates a regime which we refer to as *demand-constrained* with banks eager to lend but firms not willing to borrow.[5] The relevant segment of the demand curve is steep, i.e. firms do not increase investment to a great extent even if the cost of capital decreases a lot. The supply curve has a large curvature but is cut in a rather flat segment in point A. In this case, an increase in the investment opportunities shifting the demand curve rightwards has a large effect on investment (B). Lowering the interest rate (C) and lowering the agency cost of external funds (D) have low effects on investment.

Panel (c) represents the worst case scenario, a *credit-and-demand-constrained* regime, with high uncertainty in the financial and real sector and insufficient policy measures to lower agency costs. The demand and supply curve intersect in A where both curves are steep. Neither an improvement in investment opportunities (B) nor a reduction in the interest rate (C) nor a decline in the cost of external finance (D) have pronounced individual effects on investment. Only a policy mix has large effects on investment.

Finally, panel (d) represents the best case scenario in terms of policy options, an *unconstrained* regime. The flat demand curve cuts the flat supply curve in A. All previously discussed policy measures have fairly considerable effects on investment moving the equilibrium to B, C and D, respectively.

2.3 The econometric specification

The econometric challenge is to disentangle the effects of changes in internal funds, investment opportunities and the interest rate on equilibrium investment, i.e. to identify shifts of the supply curve due to changes in internal funds and due to changes in the cost of capital

[5] Note that the term "constraint" is used in a very loose manner here as a lack of investment opportunities is technically not a constraint to investment (as opposed to a credit constraint) but it is rather affecting the control variable itself.

as well as shifts of the demand curve due to changes in investment opportunities. Since none of the variables is directly observable, proxies need to be identified.

Introducing convex adjustment costs to the investment demand model implied by the graphical illustration above, one can show that the shadow price of a marginal unit of capital, Tobin's marginal q, fully describes expected profitability (Tobin 1969). Yet, as this variable is not observable either, the empirical literature usually approximates it with Tobin's average q, i.e. the ratio of capital stock's market value to its replacement cost, which equals marginal q only if additional assumptions on the production and adjustment cost functions are imposed (Hayashi 1982).[6]

As the data allows only for a rough approximation of the investment opportunities by q, its empirical performance has proven to be notoriously poor. To prevent shifts in the net worth from capturing the part of the variance in the residuals caused by the investment opportunities not captured by q, it is common in the empirical investment literature to include a sales related measure as another proxy for expected profitability.[7]

As a measure of a firm's internal funds, the firm's cash flow is most often used. The cash flow can be expected to be correlated with investment opportunities. Yet, the resulting distortion of the cash flow's parameter is minimal as both Tobin's q and sales growth control for expected profitability.

A change in the real market interest rate shifts the capital supply curve up- or downwards. To control for this effect on the equilibrium capital stock, a proxy for the user cost of capital needs to be included in the econometric analysis.

[6] An exception is Gugler et al. (2004) who estimated Tobin's marginal q. We followed their approach, but the estimated variable proved to be mostly insignificant in the succeeding regressions. Hence, we report the results with Tobin's average q only.

[7] Early accelerator theories of investment derive a positive relationship between capital expansion and changes in sales from a production function with decreasing returns to scale with fixed factor proportions (cf. Eisner 1960).Abel and Blanchard (1986) set up a general accelerator model which features a positive relationship between investment and sales in levels. Influential recent empirical studies of firm-level investment expenses using sales either in levels or in differences as a covariate are, among others, Fazzari et al. (1988) and Chirinko et al. (1999).

Taking into account these theoretical considerations, we assume the accumulation rate of firm i to be determined by the following data generating process:

$$g_{i,t} = \sum_{k=1}^{L} \beta_{g,k} g_{i,t-k} + \sum_{k=0}^{L} \beta_{s,k} s_{i,t-k} + \sum_{k=0}^{L} \beta_{q,k} q_{i,t-k} + \sum_{k=0}^{L} \beta_{r,k} r_{i,t-k} + \sum_{k=0}^{L} \beta_{j,k} j_{i,t-k} + \sum_{s=0}^{4} \beta_{d,s} d_s + \mu_i + \varepsilon_{i,t} \tag{1}$$

where $g_{i,t}$, $s_{i,t}$, $q_{i,t}$, $r_{i,t}$, and $j_{i,t}$ are the rate of capital accumulation, the growth rate of sales, Tobin's average q, the cash flow-capital ratio and the cost of capital for firm i in time t. d_s with $s = \{1, 2, 3, 4\}$ are seasonal dummies. μ_i are unobserved fixed firm effects and $\varepsilon_{i,t}$ are idiosyncratic random disturbances independent of the regressors. We define

$$\phi_x \equiv \frac{\sum_{k=0}^{L} \beta_{x,k}}{1 - \sum_{k=1}^{L} \beta_{g,k}} \tag{2}$$

as the average long-run response of g to a one-unit change in variable x.

Lags of the covariates have been included in (1) as investment expenses do not adjust instantaneously to changes in investment opportunities, cash flow and the cost of capital. This is because expectations of future sales may depend on past sales, costs of adjustment may slow down the adjustment process, and delivery lags may delay investment expenditures (cf. Abel and Blanchard 1986).

Pre-analysis of the firm-level data on investment strongly suggests that the accumulation rate exhibits strong autocorrelation. Therefore, we also include lags of the dependent variable as regressors.

3 Estimation strategy

The General Method of Moments (GMM) allows us to obtain consistent parameter estimates even though endogenous regressors remain in the specification.[8] The idea of the GMM estimator is to exploit moment conditions which are assumed to hold, in order to come up with a consistent estimate of the parameter vector. Parameter estimates are obtained by choosing the vector of coefficients such that a weighted quadratic form of the empirical moment conditions is minimized. This sum is usually larger than zero as there are, in general, more moment conditions to exploit than parameters to estimate. The choice of the weighting matrix is crucial for the efficiency of the estimator, yet irrelevant for its consistency. Moments with lower variance and covariance should be given more weight. Whereas the one-step GMM estimator simply chooses an optimal weighting matrix under the assumption of spheric disturbances, the two-step GMM estimator employs the residuals of an auxiliary regression to obtain a consistent estimate of the weighting matrix which implies the GMM estimator to be efficient.[9]

Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998) applied the GMM framework to dynamic panels and developed estimators which are able to cope with the correlation between the unobserved fixed group effects and the lags of the dependent variables as well as the endogeneity of other regressors without requiring the researcher to find suitable instruments outside the sample at hand.

Difference GMM and System GMM are two methods to deal with the endogeneity of the

[8] Since the lags of the dependent variable are positively correlated with the fixed firm effects included in the disturbance terms and none of the contemporaneous values of the independent variables can be plausibly assumed to be exogenous, i.e. they are most likely correlated with the idiosyncratic shocks, pooled estimation of (1) by OLS is not an option. Also estimating (1) by OLS after removing the fixed firm effects by applying a within transform to the data by which the group-mean is subtracted from each variable (fixed effects estimator) and dropping the contemporaneous values of the independent variables will yield biased estimates. Although this procedure gets rid of the most pronounced biases in the numerator of the long-run coefficients in (2), the estimates for the lagged dependent variables are now downward biased, i.e the denominator in (2) is upward biased and the long-run coefficient overall is downward biased.

[9] Appendix C formally derives of the one and two-step estimators.

lagged dependent variable and of other regressors. Difference GMM developed by Arellano and Bond (1991) applies a first-difference transform to the data in order to eliminate the fixed effects. The lags of the dependent variable on the right-hand side may still be endogenous after taking first differences. Yet, in contrast to a Within Group transform, lags longer than the lags included in the specification remain orthogonal to the disturbances and, therefore, available as instruments. To overcome the trade-off between lag length and sample size which is a common problem of the conventional 2SLS estimator and arises from the elimination of observations for which lagged values are missing, so-called "GMM-style" instruments in levels are constructed which imply a different set of instruments for each period and replace missing values by zeros.

Blundell and Bond (1998) build on Arellano and Bover (1995) seeking to further increase the efficiency of the GMM estimator under additional assumptions. They develop System GMM which exploits the moment conditions of Difference GMM for transformed data as well as an additional set of moment conditions for untransformed data derived from the assumption that the first differences of any variable used as an instrument are uncorrelated with the fixed group effects. In this case, the endogeneity of the untransformed lagged dependent variable arising from its correlation with the fixed effects can be resolved by instrumenting it with transformed, i.e. first-differenced, lags.[10]

4 Data

We use quarterly firm-level data ranging from 1975:1 to 2010:4 obtained from S&P's Compustat North America Fundamentals Quarterly database. It includes data on the balance sheets and income statements of publicly owned corporations. We excluded the finance, insurance and real estate sectors (SIC codes 6011 to 6799) as their investment dynamics can

[10]Difference GMM and System GMM are formally derived in Appendix D.

be expected to deviate substantially from the rest of the private business sector.

For the econometric analysis, the following variables have been computed: The accumulation rate, g, is the quarterly growth rate of the real net stock of capital in property, plant and equipment. The rate of sales growth, s, is the growth rate of real net sales. The cash flow-capital ratio, r, is defined as real after-tax income normalized by the beginning-of-period real net capital stock. Tobin's average q is approximated by the sum of the market value of equity and the book value of total debt, divided by the book value of total assets. To compute the cost of capital, j, we follow Fazzari and Athey (1987) by using the capital asset pricing model (CAPM) to estimate a firm-specific measure.[11]

A few remarks on the construction of the variables are in order. First, the change in the net capital stock is used as a proxy for investment as the coverage of capital expenses in the firms' cash flow account is insufficient for our purposes. Moreover, the quality of the quarterly year-to-date data on capital expenses as a quarterly measure is questionable as many firms appear to report the yearly capital expenses in the fourth quarter of the year. The correlation between the growth rate of the capital stock and the capital expenses-capital stock ratio is around 75%.

Second, in order to be consistent with our dependent variable, we relate the cash flow net of depreciation to the net capital stock. Surprisingly, related studies such as Gugler et al. (2004), Fazzari et al. (1988) as well as Gilchrist and Zakrajsek (2007) relate the cash flow including deprecation to the net capital stock. In the course of time, this, ceteris paribus, necessarily leads to a rising cash flow-capital ratio and may thus create distortions.

[11] The CAPM postulates that the rate of return on assets required by asset holders equals the risk-free rate plus the market price of risk weighted by the so-called beta-coefficient, i.e. the extent to which a return of an asset varies with the market. We estimate the required asset return using Moody's Aaa bond rate as the risk free rate, the difference between the Aaa and Baa bond rate as the market price of risk and, as the asset beta, the equity beta reported in the Compustat database, averaged over time for each firm and adjusted by the debt-asset ratio. The adjustment is required as we are interested in the asset beta which is the weighted sum of the debt beta and the (reported) equity beta with the debt-asset ratio and the equity-asset ratio being the respective weights. We follow Fazzari and Athey (1987) by assuming that the debt beta is zero. A description of all variables and the data sources is provided in Appendix B.

Third, we follow Chung and Pruitt (1994) and construct an approximation of Tobin's average q which has been found to explain at least 96.6% of q constructed according to Lindenberg and Ross' (1981) procedure. This is in line with the literature such as Gilchrist and Zakrajsek (2007) and Gugler et al. (2004).

After applying a standard screening procedure to the data, which is discussed in greater detail in Appendix B, with the aim of removing outliers and condensing the data, the final data set shrunk to 311,892 observations and 10,426 firms covering the period from 1975:1 to 2010:4. Note that observations with a non-positive cash flow have been removed. As documented by Schoder (2011), the marginal effect of cash flow on investment is lower with negative realizations than with positive ones as firms usually do not reduce their capital stock in the case of negative profits to the same extent as they raise their capital stock in the case of positive profits. Using unfiltered data, this asymmetry in the cash-flow elasticity of investment implies a strong cyclical behavior of the cash flow-capital ratio coefficient as profits decrease in the downturn implying a lower estimated coefficient. Since the main interest of the present paper is to study the cyclical behavior credit constraints approximated by the cash-flow elasticity of investment, we only consider observations with positive cash flows and exclude the possibility that a cyclical behavior of the cash flow-capital ratio coefficient is driven by an asymmetry in the cash-flow elasticity of investment as found by Schoder (2011).

5 Empirical investment regimes

Figure 3 depicts the aggregate rate of accumulation for the US non-farm non-financial corporate business sector since the 1970s.[12] A heat map indicates the business cycle. The more intense the color, the slower the economic expansion. As a business cycle measure we use an

[12]The accumulation rate is gross fixed investment in non-residential equipment, software and structures divided by fixed assets in non-residential equipment and software and non-residential structures of the non-farm non-financial corporate business sector. The data has been taken from the Flow of Funds Account of the US published by the Fed.

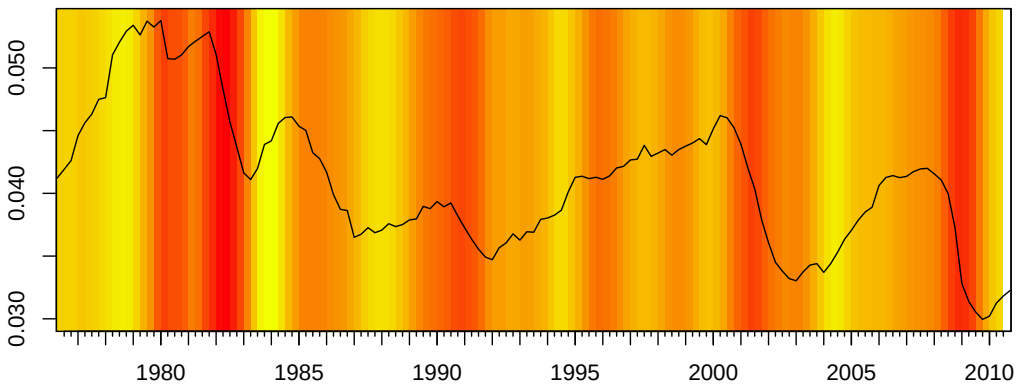

Figure 3: Aggregate accumulation rate for the US non-farm non-financial business sector (Source: Fed)

index for business confidence also known as the Purchasing Managers Index computed by the Institute of Supply Management. It is a composite index using information on production levels, new orders, supplier deliveries, inventories, employment levels of US manufacturing firms.[13] Note that the business cycle indicator has been smoothened to facilitate readability.

Unsurprisingly, the rate of capital accumulation exhibits a strong pro-cyclical pattern. The recessions in the early 1980s, the mid-1980s, the early 2000s and the recent crisis in the late 2000s featured the largest declines in the accumulation rate. The last recession saw a slow-down of the speed of capital accumulation to the lowest rate measured in the entire period considered. In the beginning of 2011 the accumulation rate appears to be still below both average and trend.

The remainder of this section, discusses by means of econometric analysis when demand and/or credit constraints to investment were binding historically.

[13] Alternatively, we experimented with the growth rate of GDP as a business cycle indicator which yielded a similar heat map.

5.1 Estimation results

For estimating (1), we choose a lag length of $L = 4$ as the standard specification.[14] To estimate this specification consistently we use the two-step System GMM estimator. Since the idiosyncratic shocks in (1) are likely to be correlated with the contemporaneous values of s, q, r and j, we take these variables as endogenous.[15]

Since we are interested in the cyclical behavior of the long-run coefficients, we apply a rolling window procedure in order to obtain time-varying coefficients. We recursively estimate the investment function for samples spanning over five consecutive quarters and moving from the beginning to the end of the period considered. For each of the samples, the respective long-run coefficients are then calculated according to (2). Since we are not primarily interested in quarter-to-quarter fluctuations which blur the overall picture, the series of long-run coefficients have been smoothened by applying 5-quarter two-sided moving average filters.

To analyze the cyclical behavior of the estimated long-run coefficients, the line plots are contrasted by the heat map indicating the business cycle. Figure 4 depicts the results for the two-step System GMM estimation of (1) with $L = 4$ for publicly traded North American firms. The first four panels of the figure show the smoothed time varying long-run coefficients obtained by rolling regressions applying a window of 5 quarters as well as the 95%-confidence intervals. The last panel reports the Arellano-Bond test for zero second-order autocorrelation in first-differenced errors (with the null hypothesis of no autocorrelation).[16]

The estimates for the sales growth's long-run effect on investment oscillate around 0.1. The confidence interval indicates that the estimates are mostly significant at the 5% level.

[14] This decision is based on the analysis of dynamic correlograms which indicate a decline of the correlation between accumulation and the lags of the explanatory variables when $L \geq 5$. However, we report the results obtained with $L = 3$ and $L = 5$ as robustness checks in Appendix E. The main results are robust to the lag length chosen.

[15] Obviously, the dummies for the quarter of the year are exogenous.

[16] The Sargan test of over-identifying restrictions (with the null hypothesis that the over-identifying restrictions are valid) cannot be applied to the two-step estimator.

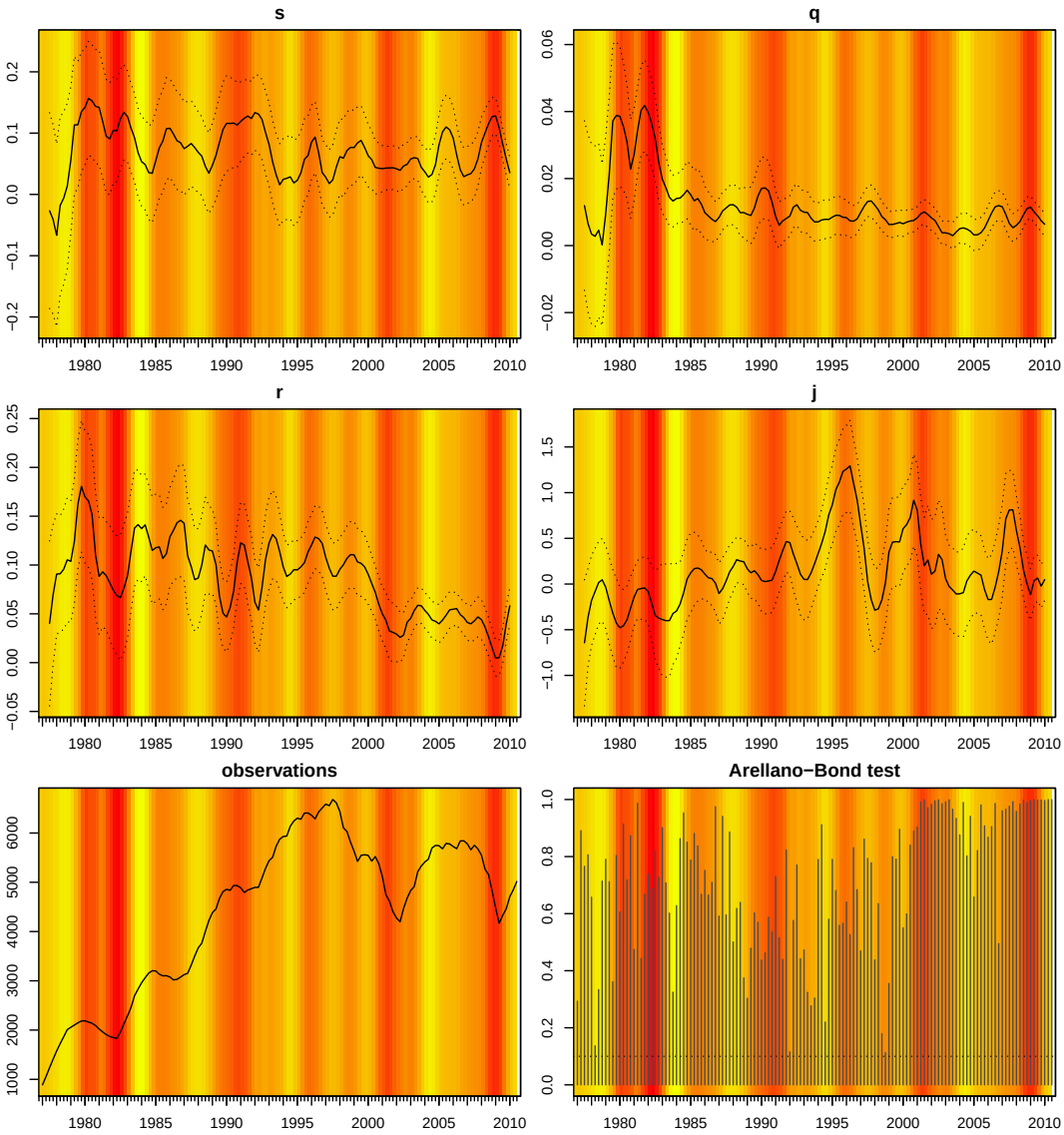

Figure 4: Regression results for the two-step System GMM estimation of (1) with $L = 4$

Apart from the early 1980s which featured high coefficients for Tobin's q, its long-run effects fluctuates around 0.01 and is mostly significant at the 5% level. The cash-flow's long-run coefficient exhibits a decreasing trend (from around 0.15 to around 0.05) which is consistent with the view that financial market integration and innovation alleviated the firms' access to credit over the last decades. The coefficient is also mostly significant. The coefficient for

the cost of capital fluctuates without significant trend around zero and is usual insignificant at the 5% level.

More interesting than the trends are the cyclical components of the estimated long-run coefficients as they, analyzed jointly, allow us to empirically assess and locate the investment regimes outlined in the theoretical section. Before analyzing each cycle individually, note the following: First, the effect of sales growth on investment exhibits in general a strong cyclical behavior. In downturns the coefficient tends to go up significantly. Further, the coefficient tends to be the higher the lower the business confidence. Notice the exception of the 2001 recession during which investment was not very sensitive to changes in sales growth. Second, the long-run coefficient for Tobin's q exhibits some cyclical behavior. The coefficient tends to go up during severe downturns such as the double-dip recession in the early 1980's and, to a lesser extent, during the S&L crisis in the early 1990s, the burst of the dot-com bubble in 2001 and the recent financial meltdown in 2009. Overall, demand expectations tend to be an important determinant of investment during times of economic distress. Third, the cash-flow coefficient does not exhibit a straightforward cyclical pattern. According to our results, strong cash-flow effects were present only in the recession following the interest rate shock in the late 1970s, weaker ones during the mild downturns in business confidence in 1995 and 1998. In neither the 2001 nor the 2009 recessions, investment was much driven by cash flow. Quite the contrary, the cash flow coefficient dropped considerably. Fourth, the fluctuations of the cost of capital's coefficient are difficult to interpret. Moreover, the coefficient is mostly insignificant.[17]

As illustrated in the bottom-left panel, the number of observations used for the recursive regressions with a window of 5 quarters increases until late 1990s from around 1,000 to more than 6,000 and fluctuates thereafter between 4,000 and 6,000. Note the obvious cyclical

[17] We estimated specifications excluding the cost of capital which, however, did not change the results. Also using a different measure for the cost of capital based on the ratio between interest payments and stock of debt as a proxy for the interest rate of the firm (cf. Dwenger 2010) did not yield different results.

pattern of the number of observations. In times of low business confidence, less observations are available. This is because we removed firm-years with negative cash flows which unsurprisingly arise especially in times of stagnation. (Note on selection bias.)

The Arellano-Bond test fails to reject the null hypothesis of second-order serial correlation in the residuals at any reasonable level of significance for all time windows considered.

5.2 Investment regimes in the US business sector

A demand regime is uniquely characterized by the relative size of the effect of demand and cash flow on investment. We can therefore use the scatter plot depicted in Figure 5 as a guidance to identifying investment regimes in the US. Figure 5 plots the demand effect against the cash-flow effect for each quarter with the color intensity indicating the lack of business confidence.[18] The first to fourth quadrants represent the *unconstrained*, *demand-constrained*, *demand-and-credit-constrained* and *credit-constrained* investment regimes, respectively.

Times of crisis and stagnation are concentrated in the right quadrants which implies that demand has typically large effects on investment during such times. Most of these quarters are located in the second quadrant meaning that investment was *demand constrained*. *Credit-and-demand-constrained* regimes seem to be rare events, as either demand expansion or cash-flow expansion or both typically push investment. Further, it seems that it is mostly times of expansion during which firms are credit constrained. Also note that the recession in the late 1970s/early 1980s is very peculiar in the sense that it featured both large demand and cash flow effects. This may indicate that policy options have not been utilized sufficiently.

Let us now study the US business cycles in more detail. The moderate downturn of

[18]The demand effect is the average of the effect of sales growth and the adjusted effect of Tobin's q. The time-varying coefficient for Tobin's q has been adjusted such that its variance is equal to the variance of the sales growth's coefficient. Since a considerable share of each of the time-varying coefficients is determined by factors independent of the business cycle, we use for the construction of the relative demand and cash-flow effects the cyclical components of the coefficients by applying an HP-filter ($\lambda = 1,600$) to the unfiltered coefficients estimated.

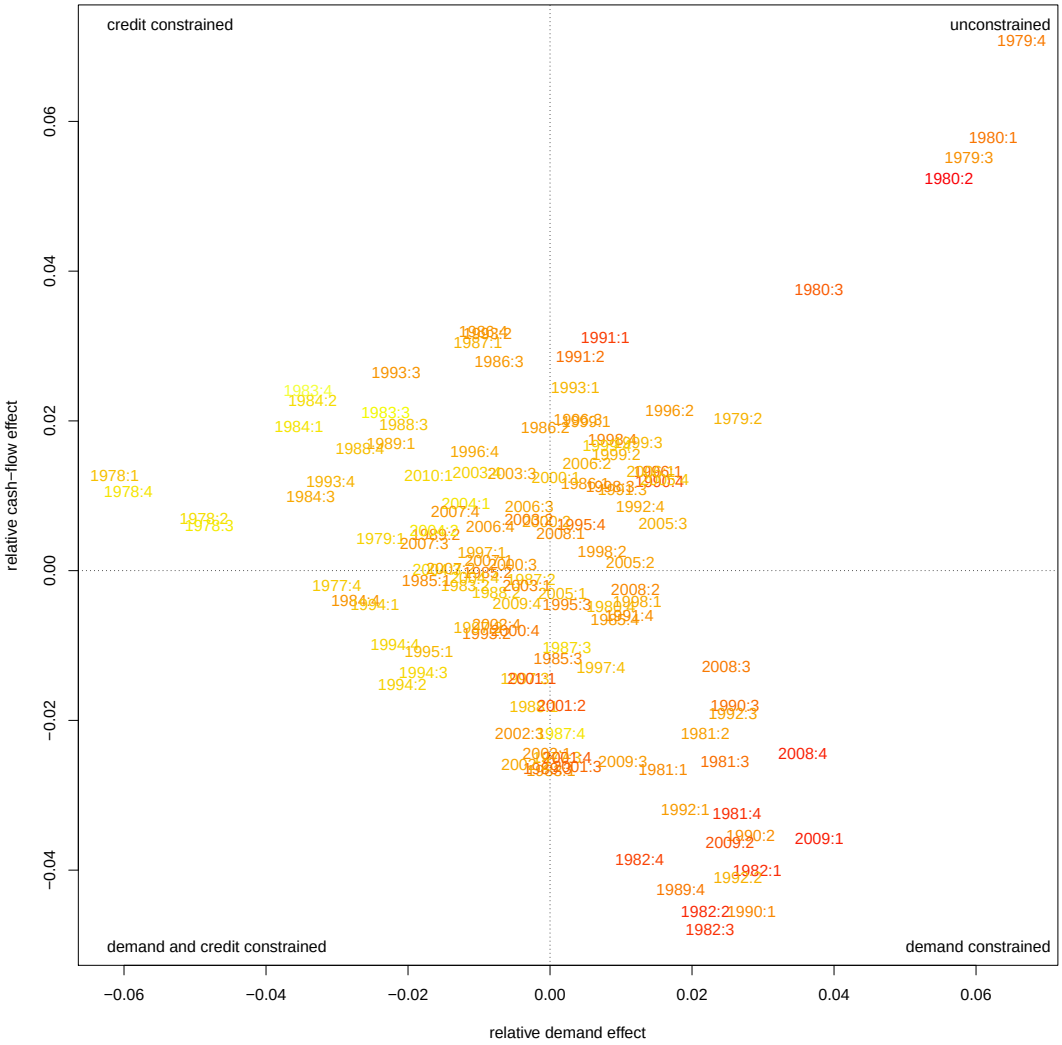

Figure 5: Investment regimes in the US: demand effects vs. cash-flow effects

investment in the recession in 1980 is associated with both a high sensitivity of investment to both demand, reflected by sales growth and Tobin's q, and cash flow which, in terms of the investment regimes discussed above, indicates an *unconstrained* regime. Hence, the decline in investment may have been caused by deteriorating demand expectations and limited access to finance as well as the failure of economic policy to stabilize demand and the flow of credit which would have been very effective. Further, the spike in the interest rate has increased

the cost of capital which contributed to the decline in investment. For the subsequent recession in 1982, the estimated coefficients indicate *demand-constrained* investment as the demand-related coefficients move upwards whereas the cash flow's coefficient drops. Credit constraints on capital expansion seem to become more important in the following upswing.

The drops in the accumulation rate observed in the mid 1980s and early 1990s are both associated with increases in the demand effects on investment (especially sales growth) and an ambiguous behavior of the cash flow effect. Again both periods of economic stagnation seem to have been caused by a lack of demand and deteriorating demand expectations. It seems that credit constraints become important at the end of each downturn.

The rest of the 1990s is, in general, a period of rising capital accumulation. Yet expansion slows down slightly in 1995 and 1998. Again, sales growth becomes relatively more important in predicting investment expenses. Interestingly, the coefficient of Tobin's q spikes upwards between the two periods of stagnation. The slowdown of accumulation may also have been enforced by tightening credit constraints as indicated by rising cash flow effects.

The decline of non-farm non-financial business investment after the burst of the dot-com bubble in 2001 is associated with low demand and cash-flow effects on investment which is consistent with a *demand-and-credit-constrained* investment regime.

The crisis following the financial meltdown in 2007 clearly features a *demand-constrained* investment regime with large demand and low cash-flow effects on investment.

6 Fiscal and monetary policy in the US in the context of the investment regimes identified

Using the development of business investment and the estimated elasticities of investment with respect to demand and access to credit, one can assess the historical monetary and fiscal policy measures conducted to manage demand and the flow of credit. Before doing so,

however, it is useful to briefly review the fiscal and monetary policy stance for the period under consideration.

6.1 Management of aggregate demand and credit flows

The first panel in Figure 6 depicts the federal funds rate which has traditionally been the main instrument used by the FED to stabilize demand, apart from the intermezzo in the early 1980s when the Reagan administration sought to target the money supply instead of the short term interest rate. Apart from the recession in the mid 1970s, when the nominal interest (but also inflation) was still high during the downturn, the monetary instrument usually spiked before the recession and dropped sharply during the downturn.

Since there is a feedback effect of fiscal policy to aggregate demand, a quantitative assessment of the fiscal effort relative to the severity of the economic downturn is a non-trivial endeavor. The challenge is to relate the indicator of the fiscal effort to a measure of the deepness of the recession which is independent of the fiscal indicator used. We consider three measures plotted in the second to fourth panel of Figure 6.

First, a rough approximation but still fairly insightful measure is the percentage deviation of actual GDP from potential GDP, i.e. the relative output gap. As long as there is no crowding out of private spending through public spending, the relative output gap indicates the extent by which the public sector fails to to compensate the downturn of private demand. Yet, the assumption of no crowding out may be too restrictive. The following two measures are highly robust to crowding out.

Second, we consider, for each period of consecutive years with negative output gaps since the 1960s, the ratio between the part of the primary fiscal deficit arising from discretionary policy over period of stagnation considered and the part arising from the automatic stabilizers. This measure of the relative fiscal policy stance may include some revealing information as the denominator of this ratio is affected by the numerator only to the second

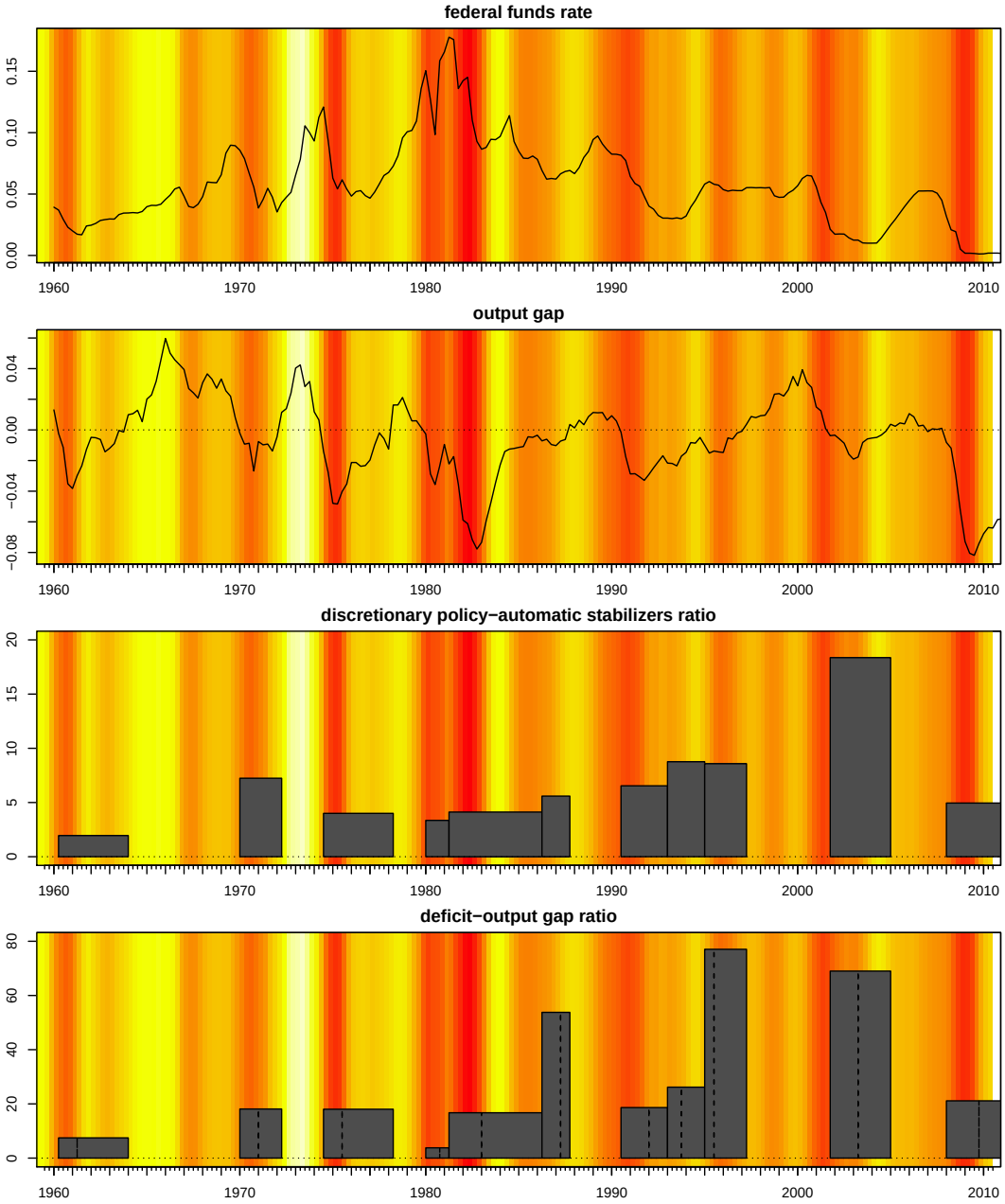

Figure 6: Indicators for monetary and fiscal efforts to stabilize aggregate demand

order (through GDP).

Finally, we consider the ratio between the real primary deficit cumulated over the period of economic stagnation and the real negative output gap as an average from the beginning of

the period to the trough, i.e. the maximum negative output gap. This measure approximates the relative fiscal effort as long as the deficits until the trough do not affect the economic downturn too much. Of course, this is a strong assumption but as the trough is usually reached within only a few quarters, the bulk of the fiscal policy measures, especially those with a strong delay, may affect the recovery rather than the downturn.[19]

As can be seen in the lower three panels of Figure 6, the fiscal effort measures are broadly consistent with each other. The trend of the discretionary policy-automatic stabilizers ratio reveals that fiscal stimulus packages became increasingly important in fighting recessions as compared to automatic stabilizers. Also the deficit-average downturn output gap ratio exhibits an increasing trend which, however, is not easy to interpret. It may be due to the way the ratios have been computed and the fact that the economy is growing.

Apart from the reserve-targeting experiment in the late 1970s and early 1980s, the Fed used the short-term interest rate as the primary target of monetary policy since the 1970s. To some extent, interest rate targeting in itself implies a mechanism of sustaining the flow of credit by providing the required short-term liquidity to financial institutions. A moderate positive demand shock for and/or negative supply shock of reserves put upwards pressure on the federal funds rate which is the rate on the market for reserves held at the Fed by depository institutions. Conducting outright or temporary open market operations, i.e. buying Treasury bills from its primary dealers, the Fed can increase the supply of reserves in the market to meet the increased relative demand.

As an additional safety valve, depository institutions may directly lend from the Fed to meet their reserve requirements through the discount window. This is a very effective way

[19]These measures have been computed for each *sub-cycle* which is the period between two local maxima of the output gap with all values being negative. The trough of a sub-cycle is defined as the minimum output gap in that period. Hence, the downturn is the part of the sub-cycle which features a decreasing output gap whereas the output gap is increasing in the recovery. Note that the downturn starts not necessarily when the output gap starts falling but when it becomes negative. Equivalently, the recovery ends when the output gap becomes positive and not necessarily when the next downturn starts, i.e. when the output gap starts falling again. This is in order to have the two measures as comparable as possible.

to provide short-term liquidity to the banking sector in case of sudden shifts in demand and supply as was the case, for instance, after the terrorist attacks in September 2001 and after the bankruptcy of Lehman Brothers in September 2008.

6.2 Assessing fiscal and monetary policy measures in the context of investment regimes

According to our results, the double-dip recession in the early 1980s featured *unconstrained* and *demand-constrained* investment regimes, respectively, with the relevant quarters being located deep in the first and second quadrant, respectively. At the same time the low discretionary policy-automatic stabilizers ratio suggests that fiscal policy under the Reagan administration relied primarily on automatic stabilizers rather than discretionary expansionary policy which may have contributed to the rather long duration of the economic stagnation. Overall, the fiscal effort was lower than during the succeeding recessions. The deficit-output gap ratio suggests that the second part of the recession involved more fiscal efforts than the first one. Given the high elasticity of investment with respect to demand and cash flow in the first dip (*unconstrained* investment), a more active demand as well as credit-flow management would have likely had considerable marginal effects and might have reduced the decline of investment. During the second dip, the potential stabilizing effects of demand management were still large whereas the option for credit-flow policies were exhausted. Interestingly the expansion following the double-dip recession is consistent with a *credit-constrained* investment regime.

The relative fiscal efforts during the recession in the early 1990s as a consequence of the S&L crisis were rather large compared to previous downturns as can be read from Figure 6. The mild recession in the early 1990s was mainly fought by increasing benefits for unemployeds and coincided with a fiscal expansion due to the first Gulf War. To maintain the

flow of credit during the S&L crisis in the late 1980s and early 1990s, the Fed conducted monetary policy known as quantitative policy by which corporate bonds as well as mortgage-backed securities were purchased by the Fed in order to inject liquidity to the banking sector. The latter measures might have been successful in providing businesses with credit, whereas the demand management might not have been timely enough as investment was *demand constrained* during the downturn.

The respective investment regimes during the slowdowns of expansion in the mid and late 1990s are difficult to categorize. According to the scatter plot in Figure 5, the prevalent regimes tend to be between *credit constrained* and *unconstrained*. Analyzing the respective coefficients in Figure 4 reveals that both the sales-growth effect and the cash-flow effect increase significantly whereas the coefficient of Tobin's q does not change in the expected anti-cyclical way. Hence, even though demand expectations were not particularly optimistic, firms with high sales and, therefore, high capacity utilization were able to expand their capital stock, given the level of cash flow.

The recession following the dot-com bubble constitutes a very interesting case as it features the purest form of a *demand-and-credit-constrained* investment regime observed in the data. Policy attempts to stimulate demand and the flow of credit in order to stabilize investment may have been either effective or ineffective as both scenarios are consistent with the coefficients observed. Note, however, that the fiscal policy measures have been enormous.[20] Nevertheless, business investment dropped tremendously indicating that these measures were not particularly successful. Again it is the early recovery which seems to be *credit constrained*.

We obtain an unambiguous result for the recent recession following the financial meltdown

[20]The Bush administration taking over in 2001 started off with major tax cuts such as the Economic Growth and Tax Relief Reconciliation Act of 2001, the Job Creation and Worker Assistance Act of 2002 and the Jobs and Growth Tax Relief Reconciliation Act of 2003. On the spending side, the wars in Afghanistan (started in 2001) and in Iraq (started in 2003) implied a significant fiscal expansion.

beginning in 2007: investment was *demand constrained.* Only in the early recovery, the economy moved into the *credit-constrained* quadrant. According to the policy measures depicted in Figure 6, the relative fiscal policy effort has been far below trend in the recent economic crisis. Taking into account the fact that the potential output has been revised downwards considerably after 2007, the figures for the current recession overstate the fiscal policy effort. The relative fiscal expansion is moderate despite two major stimulus packages launched by the US government.[21] Nevertheless, these measures were only moderate given the speed and extent of the downturn which has been argued by many Keynesian economists and is confirmed by the plots in Figure 6.

Since the federal funds rate hit the zero lower bound by the end of 2008 and the economy responded too slowly to the declining short-term interest rate, the Fed made excessive use of unconventional monetary policy through existing as well as newly created, temporary lending facilities.[22]

[21]The *Economic Stimulus Act of 2008* involved $168 billion in tax rebates for consumers and businesses. The second major stimulus package was implemented by the $787 billion *American Recovery and Reinvestment Act of 2009* which mainly included tax reliefs, funding for states and public investment in infrastructure, health and education.

[22]First, the Fed extended its balance sheet to raise the monetary base and provide liquidity (quantitative easing), growing from $869 billion in August 2008 to $2,882 billion in July 2011. The balance sheet expansion from August 2008 to March 2009 was driven by short-term loans to the banking sector as well as to institutions in key credit markets. Since depository institutions avoided borrowing through the discount window due to its stigma of bank failure, the Fed introduced the *Term Auction Facility* in December 2007 to provide short-term liquidity to banks without revealing their identity. In March 2008, the Fed announced the *Primary Dealer Credit Facility* through which the central bank lent out directly to primary dealers which was not permitted before. The *Term Asset-Backed Securities Loan Facility* announced in November 2008 aimed at stimulating consumer and business credit lending by granting loans against the collateral of asset-backed securities backed by recently issued consumer and business loans. After March 2009, the balance sheet expansion mainly reflected purchases of Treasury, agency and agency-guaranteed securities. Second, the central bank shifted its portfolio towards riskier assets (qualitative easing) to take out risk from financial institutions' balance sheets to foster private lending. The *Term Securities Lending Facility* implemented in March 2008 was aimed at increasing the liquidity of the banking sector by allowing primary dealers to switch illiquid debt for tradable government securities with the Fed. From November 2008 to November 2009, the Fed purchased $175 billion of government-sponsored enterprise debt as well as $1,250 billion of mortgage-backed securities. In March 2009, the Fed started buying longer-term Treasury securities for $300 billion, followed by a second round of purchasing Treasury securities for $600 billion. To keep credits flowing, both the Fed and the Treasury guaranteed private assets and loans. For instance, the *Temporary Liquidity Guarantee Program* involved a $1.4 billion guarantee on senior subordinated debt and poorly performing assets by banks as well as on non-interest bearing deposit accounts. One major attempt of the US government to sustain

It is not implausible to argue that the observed coefficients for the recent economic downturn are the result of the adverse weights given to the management of demand and the management of credit flows. It seems that the policy measures taken were sufficient to eliminate credit constraints but insufficient to stabilize investment through raising aggregate demand and stimulating demand expectations. Hence the decline of investment starting in 2008 is due to a lack of demand management rather than attempts to maintain the access to credit. Yet, as sales and utilization increased and expectations improved out of the recession, credit constraints became binding for an increasing number of firms.

7 Concluding remarks

The policy makers' emphasis of measures aimed at maintaining the flow of credit from the banking sector to the private sector in order to overcome times of economic depression is theoretically well founded in the conventional literature on investment theory and informational economics. At the same time, demand management is not given high priority...

This paper seeks to contribute empirically to the question what factors drive investment over the business cycle. It sets out with a simple model of investment which motivates the notion of demand constraints and credit constraints. We distinguish four stylized investment regimes: (1) *demand-constrained* investment with firms not willing to increase investment even though the conditions by which they can obtain external finance improves, (2) *credit-constrained* investment with firms not able to increase investment as a response to improved profit expectations due to the financial market's unwillingness to provide funds, (3) *demand-and-credit-constrained* investment with firms facing both demand and credit constraints, and (4) *unconstrained* investment with firms' capital accumulation being very responsive to changes in investment opportunities and in financial market conditions.

the flow of credit was the *Emergency Economic Stabilization Act of 2008* which enabled the government to spend up to $700 billion to purchase troubled bank assets in order to strengthen the banks' balance sheets.

Following the convention in the empirical investment literature, investment opportunities are approximated by the growth rate of sales and Tobin's q whereas the finance constraints are reflected by the internal flow of cash. The four investment regimes imply the following effects on investment: (a) high demand effect, low cash-flow effect, (b) low demand effect, high cash-flow effect, (c) low demand effect, low cash-flow effect, and (d) high demand effect, high cash-flow effect, respectively.

To characterize corporate non-financial business investment since 1977 for the US in terms of the investment regimes postulated, a dynamic investment function including sales growth, Tobin's q, cash flow and the cost of capital as covariates has been estimated using firm-level data obtained from Compustat. As we are interested in the cyclical behavior of the covariates' contributions to explaining investment, we estimate time-varying coefficients by applying a rolling window regression with a size of five quarters. Since the econometric model of investment is dynamic, we use the System GMM estimator developed by Arellano and Bover (1995) and Blundell and Bond (1998).

We find that the investment elasticities of the demand variables, i.e. sales growth and Tobin's q, are in general highly counter-cyclical. Overall, demand constraints seem to be crucial factors contributing to the slowdown of accumulation in times of economic distress. In contrast to the prediction of the financial accelerator literature that credit constraints tighten in the downturn as net worth deteriorates, the cash-flow coefficient does not exhibit a clear counter-cyclical pattern.

We further find that the most tremendous declines in business investment which occurred in the contexts of the recessions in 1982, 1990, 2001 and 2008/09 are to be associated with *demand-constrained* investment regimes as, during these times, an improvement of investment opportunities reflected by an expansion of sales growth and Tobin's q, on average, induced firms to raise investment to a disproportionately large extent whereas an easing of credit constraints, on average, provoked only a disproportionately small expansion of

investment.

These findings do not imply that credit constraints are irrelevant in practice. In fact, during the slight slow-downs of economic expansion in 1985, 1995 and 1998 firms seem to have faced credit constraints. Moreover, during the early recoveries after the 1982, 1990, 2001 and 2008/09 recessions, the cash-flow coefficients rise indicating a tightening of credit constraints. Given the large efforts of US monetary and fiscal policy to maintain the flow of credit during deep recessions as reviewed in the present paper, low cash-flow coefficients may indicate that these policy measures have been successful in preventing credit crunches.

Yet, the high elasticities of investment with respect to sales growth and Tobin's q may reflect an insufficient use of measures aimed at stabilizing aggregate demand and expectations during times of economic distress. Only during the 2001 recession which is associated with the largest relative fiscal effort that we measure, the demand effects on investment did not unambiguously rise.

Of particular interest is the recent recession in 2008/09 following the financial meltdown which has been answered by dramatic policy intervention such as the creation of special lending facilities as well as government purchases and guaranties of troubled assets all of which were mainly aimed at ensuring the liquidity of the banking sector. At the same time, the fiscal stimuli packages have been criticized by economists as being too small which is also confirmed by our measures of the relative fiscal effort. The adverse usage of demand management tools and credit flow tools is also reflected by the coefficients of the investment function during the economic downturn. Large demand and low cash-flow effects characterize the recession and indicate demand rather than credit constraints to investment. The policy conclusion is that more ambitious demand management might have reduced the size of the recession considerably.

References

Abel, A. B., Blanchard, O. J. (1986): Investment and sales: some empirical evidence, Working Paper 2050, National Bureau of Economic Research.

Ahn, S. C., Schmidt, P. (1995): Efficient estimation of models for dynamic panel data, *Journal of Econometrics*, 68(1), pp. 5–27.

Arellano, M., Bond, S. (1991): Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies*, 58(2), pp. 277–97.

Arellano, M., Bover, O. (1995): Another look at the instrumental variable estimation of error-components models, *Journal of Econometrics*, 68(1), pp. 29–51.

Bernanke, B. S., Gertler, M. (1989): Agency Costs, Net Worth and Business Fluctuations, *The American Economic Review*, 79(1), pp. 14–31.

Blundell, R., Bond, S. (1998): Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics*, 87(1), pp. 115–143.

Blundell, R., Bond, S., Windmeijer, F. (2000): Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator, IFS Working Papers W00/12, Institute for Fiscal Studies.

Chirinko, R. S., Fazzari, S. M., Meyer, A. P. (1999): How responsive is business capital formation to its user cost?: An exploration with micro data, *Journal of Public Economics*, 74(1), pp. 53–80.

Christiano, L., Eichenbaum, M., Rebelo, S. (2009): When is the government spending multiplier large?, NBER Working Papers 15394, National Bureau of Economic Research.

Chung, K. H., Pruitt, S. W. (1994): A Simple Approximation of Tobin's q, *Financial Management*, 23(3), pp. 70–4.

Cogan, J. F., Cwik, T., Taylor, J. B., Wieland, V. (2010): New Keynesian versus old Keynesian government spending multipliers, *Journal of Economic Dynamics and Control*, 34(3), pp. 281–295.

Dwenger, N. (2010): User cost elasticity of capital revisited.

Eisner, R. (1960): A Distributed Lag Investment Function, *Econometrica*, 28(1), pp. 1–29.

Fazzari, S. M., Athey, M. J. (1987): Asymmetric Information, Financing Constraints, and Investment, *The Review of Economics and Statistics*, 69(3), pp. 481–87.

Fazzari, S. M., Hubbard, R. G., Petersen, B. C. (1988): Financing Constraints and Corporate Investment, *Brookings Papers on Economic Activity*, 1988(1), pp. 141–206.

Fazzari, S. M., Hubbard, R. G., Petersen, B. C. (2000): Investment-cash flow sensitivities are useful: a comment on Kaplan and Zingales, *The Quarterly Journal of Economics*, pp. 695–705.

Gertler, M., Hubbard, R. G. (1988): Financial Factors in Business Fluctuations, *Financial market volatility: Causes, consequences, and policy recommendations. Federal Reserve Bank of Kansas City*, pp. 33–71.

Gilchrist, S., Zakrajsek, E. (2007): Investment and the Cost of Capital: New Evidence from the Corporate Bond Market, NBER Working Papers 13174, National Bureau of Economic Research, Inc.

Greenwald, B., Stiglitz, J. E., Weiss, A. (1984): Informational Imperfections in the Capital Market and Macroeconomic Fluctuations, *The American Economic Review*, 74(2), pp. 194–9.

Gugler, K., Mueller, D. C., Yurtoglu, B. B. (2004): Marginal q, Tobin's q, Cash Flow, and Investment, *Southern Economic Journal*, 70(3), pp. 512–31.

Hayashi, F. (1982): Tobin's Marginal q and Average q: A Neoclassical Interpretation, *Econometrica*, 50(1), pp. 213–24.

Jensen, M. C., Meckling, W. H. (1976): Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure, *Journal of Financial Economics*, 3(4), pp. 305–60.

Kaplan, S. N., Zingales, L. (1997): Do investment-cash flow sensitivities provide useful measures of financing constraints?, *The Quarterly Journal of Economics*, 112(1), pp. 169–215.

Lindenberg, E. B., Ross, S. A. (1981): Tobin's q Ratio and Industrial Organization, *Journal of Business*, 54(1), pp. 1–32.

Myers, S. C., Majluf, N. S. (1984): Corporate Financing and Investment Decisions when Firms Have Information That Investors Do Not Have, *Journal of Financial Economics*, 13(2), pp. 187–221.

Romer, C., Bernstein, J. (2009): The Job Impact of the American Recovery and Reinvestment Plan, Tech. rep.

Schoder, C. (2011): Demand, q, financial constraints and shareholder value: an econometric micro-analysis of US fixed investment, Tech. Rep. 2011-8, Schwartz Center for Economic Policy Analysis (SCEPA), The New School, New York.

Tobin, J. (1969): A General Equilibrium Approach to Monetary Theory, *Journal of Money, Credit and Banking*, 1(1), pp. 15–29.

A A simple investment model

Figure 1 is the graphical representation of a simple two period partial equilibrium investment model. In period zero, a risk neutral representative firm in a competitive market produces output from capital, K. Output becomes available for consumption in period one. Two states may arise: In state 1 (which is the "bad" state and occurs with probability π_1), K units of capital are transformed into $\alpha f(K)$ units of output with $0 < \alpha < 1$; in state two (which is the "good" state and occurs with probability π_2), $f(K)$ units of output are produced.

The production function, $f(K)$, is assumed to be at least once differentiable. It is assumed to characterize the following technology. Let there be an additional factor of production such as land which is not explicitly modeled. At low K's, $f'(K)$ is constant due to the abundant availability of land. At higher K's, decreasing returns set in. Not only does $f'(K)$ decrease, $f''(K)$ also decreases as it is assumed that the elasticity of substitution between capital and land decreases with increasing K. Such a production technology, $f(K)$, exhibits a decreasing and concave investment demand schedule.

To finance the desired investment, firms may borrow from competitive lenders an amount $K - W$ if their net worth, W, is insufficient. The interest rate on external funds is r. An information asymmetry arises from the assumption that only the firm can observe the state realized for free but not the lenders. Yet, the lenders may audit the firm's announcement which implies a cost of γK units of output. The cost of audit is assumed to rise with the capital investment, as, in practice, the state verification is more difficult with large investments than with small ones. Since for the case of only two possible states Bernanke and Gertler (1989) showed that lenders only audit in the bad state, we assume this a priori. The probability of an audit is p. The cost of borrowing in the bad state with auditing, in the bad state without auditing and in the good state are $P^a + \gamma K + rK$, $P_1 + rK$ and $P_2 + rK$, respectively, with P^a, P_1, P_2 being payments additional to interest rate.

The total expected revenues and the total expected costs are

$$R = \pi_1 \alpha f(K) + \pi_2 f(K) \tag{3}$$

and

$$C = \pi_1 p(P^a + \gamma K + r(K - W)) + \pi_1(1-p)(P_1 + r(K - W)) + \pi_2(P_2 + r(K - W)), \tag{4}$$

respectively.

The problem is to find the optimal contract characterized by p, P^a, P_1 and P_2 and the optimal investment, K. It is solved by maximizing expected profits, i.e.

$$\max R - C \tag{5}$$

subject to the following constraints: First, the expected payoff for the lenders has to be larger than the opportunity cost of lending. In fact, due to the assumption of competitive capital markets, this constraint is always binding, i.e.

$$\pi_1 p(P^a + r(K - W)) + \pi_1(1-p)(P_1 + r(K - W)) + \pi_2(P_2 + r(K - W)) = r(K - W). \tag{6}$$

Second, in order to prevent the firm from lying about the true state, the expected profit must not be lower than the expected profit given the firm announces the bad state regardless of the true state. Hence,

$$\begin{aligned} R - C \geq R - [&\pi_1 p(P^a + \gamma K + r(K - W)) + \pi_1(1-p)(P_1 + r(K - W)) \\ &+ \pi_2 p(P^b + \gamma K + r(K - W)) + \pi_2(1-p)(P_1 + r(K - W))], \end{aligned} \tag{7}$$

where it can be shown that $P^b = f(K) - \gamma K - r(K - W)$. The amount a firm can credible promise to be liable for in the case of insolvency is limited by its net worth. Hence,

$$P_1 \leq \alpha f(K) - r(K - W) \tag{8}$$

and

$$P^a \leq \alpha f(K) - \gamma K - r(K - W). \tag{9}$$

Obviously, it must hold for the audition probability that

$$0 \leq p \leq 1. \tag{10}$$

Two scenarios may arise. First, the firm's net worth is large enough to pay the lenders their required return even in the bad case, i.e. $\alpha f(K) \geq r(K - W)$, or auditions are costless, i.e $\gamma = 0$. In either case, no agency problem arise. Equations (7) to (9) do not bind and the optimal audition probability, p, is zero. Substituting (3), (4) and (6) into (5), the first oder condition with respect to K implies that the optimal investment is where marginal revenue equals marginal cost. The investment demand schedule is $(\pi_1\alpha + \pi_2)f'(K)$ and the supply of finance schedule is r.

As soon as the desired capital stock exceeds the net worth, agency costs of external finance rising with investment at a given net worth arise. Then, equations (7) to (9) start binding which yields a system of four equations in five unknowns. Adding the term $\pi p \gamma K$ to either side of (6) and substituting the resulting left-hand-side as well as (9) and (10) into (7), one can solve for the p conditional on K, i.e.

$$p(K) = \frac{r(K - W) - \alpha f(K)}{\pi_2(1 - \alpha)f(K) - \pi_1\gamma K} \tag{11}$$

Table 1: Variable definitions and data sources

g	net property, plant and equipment * 100/GDP deflater / lag[net property, plant and equipment * 100/GDP deflater] - 1
s	(sales - depreciation) * 100/GDP deflater / lag[(sales - depreciation) * 100/GDP deflater] - 1
q	lag[quarterly market price of shares * common shares outstanding + total liabilities] / lag[total assets]
r	income before extraordinary items * 100/GDP deflater / lag[net property, plant and equipment * 100/GDP deflater]
j	Baa bond rate + (equity beta * (1 - lag[total liabilities / total assets])) * (Baa bond rate - Aaa bond rate)

Using (3), (4), (8), (9) and (11), (5) can be maximized with respect to K. The investment demand schedule is the same as in the case without agency costs. The marginal cost schedule, i.e. supply of funds schedule, can be shown to be increasing and strictly convex on the relevant domain.

B Data

All firm-level data have been obtained from the Compustat North America Fundamentals (monthly updated) database. The quarterly GDP deflater has been taken from the BEA, NIPA Table 1.1.4. The variables used in the regressions have been constructed as reported in Table 1.

The following screening procedure has been applied to the data: First, we removed observations with implausible realizations in any variable. In particular, we eliminated observations with (a) a non-positive lagged capital stock, (b) non-positive total assets and (c) non-positive sales.

Second, from the remaining set we excluded outliers by trimming off the 1% and 99% quantiles of each variables' distribution which is common in the literature (cf. Chirinko et al. 1999; Gugler et al. 2004).

Third, observations with accumulation rates above 1 or below -1 have been eliminated in order to further reduce the influence of extreme values. Since we are interested in the question how credit constraints vary over time and over the cycle, observations with a non-positive cash-flow are removed.

Fourth, we excluded firms that were subject to major mergers and acquisitions. We eliminated all firms for whom M&A data are available and which had M&A's of more than 20% of their net capital stock. This is in compliance with the literature (cf. Fazzari et al. 1988).

Fifth, in order to condense the data set, any observation whose realization of the accumulation rate was not part of a at least 5-quarters long sequence of realizations of the accumulation rate has bee removed. Further, observations with a missing value for the accumulation rate which is the main dependent variable have been eliminated. From the remaining set, we only included firms with realizations in the accumulation rate for at least five consecutive quarters.

Finally, very large firms with a mean of real total assets exceeding $30,000 (which is roughly 1% of all firms) have been eliminated.

C One-step and two-step GMM estimator

This section outlines the one-step and two-step GMM estimators of the parameters of a linear model as well as estimators for parameter variances which are robust to heteroskedasticity and serial correlation in the disturbance terms.

Let us consider a general linear model of the form

$$\mathbf{y} = \mathbf{X}'\beta + \boldsymbol{\varepsilon} \tag{12}$$

where the $(N \times 1)$ dimensional vector $\mathbf{y}$, the $(k \times N)$ dimensional matrix $\mathbf{X}$ and the $(N \times 1)$ dimensional vector of disturbance terms $\boldsymbol{\varepsilon}$ are realizations of the random variables y, $x_1 \dots x_k$ and ε, respectively. β is a $(1 \times k)$ dimensional parameter-vector. To ensure identification of this model, we assume

$$\mathrm{E}[\boldsymbol{\varepsilon}|\mathbf{Z}] = \mathbf{0} \tag{13}$$

where $\mathbf{Z}$ is a $(j \times N)$ dimensional matrix of realizations of instruments $z_1 \dots z_j$ with $j \geq k$.

Minimizing

$$\frac{1}{N}\hat{\boldsymbol{\varepsilon}}'\mathbf{Z}\mathbf{A}\mathbf{Z}'\hat{\boldsymbol{\varepsilon}} \tag{14}$$

where $\mathbf{A}$ is a symmetric $(k \times k)$ matrix weighting the moment conditions yields an estimator for β,

$$\hat{\beta}_{\mathbf{A}} = (\mathbf{X}'\mathbf{Z}\mathbf{A}\mathbf{Z}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}\mathbf{A}\mathbf{Z}'\mathbf{y}. \tag{15}$$

One can show that the asymptotic variance of this estimator is

$$\mathrm{Avar}[\hat{\beta}_{\mathbf{A}}] = (\boldsymbol{\Sigma}'_{\mathbf{ZX}}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{ZX}})^{-1}\boldsymbol{\Sigma}'_{\mathbf{ZX}}\mathbf{A}\mathrm{Var}[\mathbf{z}\varepsilon]\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{ZX}}(\boldsymbol{\Sigma}'_{\mathbf{ZX}}\mathbf{A}\boldsymbol{\Sigma}_{\mathbf{ZX}})^{-1} \tag{16}$$

where $\boldsymbol{\Sigma}_{\mathbf{ZX}} = \mathrm{plim}_{N\to\infty}\frac{1}{N}\mathbf{Z}'\mathbf{X}$.

Efficiency of the GMM estimator in (15) depends on the right choice of $\mathbf{A}$. More weight should be given to moments with low variance and covariances. One can show that the estimator in (15) is efficient if $\mathbf{A} = \mathrm{Var}[\mathbf{z}\varepsilon]^{-1}$. Yet, $\mathrm{Var}[\mathbf{z}\varepsilon]$ is unknown in general which renders the estimator infeasible.

To derive a feasible efficient estimator note first that $\mathrm{Var}[\mathbf{z}\varepsilon] = \mathrm{plim}_{N\to\infty}\frac{1}{N}\mathrm{E}[\mathbf{Z}'\boldsymbol{\Omega}\mathbf{Z}]$

where $\mathbf{\Omega} \equiv \mathrm{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'|\mathbf{Z}]$. Now a (not necessarily consistent) estimate for $\mathbf{\Omega}$, i.e. $\hat{\mathbf{\Omega}} = \hat{\boldsymbol{\varepsilon}}\hat{\boldsymbol{\varepsilon}}'$, has to be computed with the property that $\frac{1}{N}\mathbf{Z}'\hat{\mathbf{\Omega}}\mathbf{Z}$ is a consistent estimate for $\mathrm{Var}[\mathbf{z}\varepsilon]$. The residuals which meet this requirement can be obtained from any consistent estimator of β. In practice, the weighting matrix $\mathbf{A} = (\mathbf{Z}'\mathbf{H}\mathbf{Z})^{-1}$ is chosen for an initial estimate where $\mathbf{H}$ is an approximation of $\mathbf{\Omega}$ based on the assumption of spheric disturbances, which gives rise to the one-step GMM estimator,

$$\hat{\beta}_{\text{one-step}} = (\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{H}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{H}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}. \tag{17}$$

The residuals of this estimator are then used to construct $\hat{\mathbf{\Omega}}$ which is used for the two-step GMM estimator which is feasible and efficient,

$$\hat{\beta}_{\text{two-step}} = (\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\hat{\mathbf{\Omega}}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\hat{\mathbf{\Omega}}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{y}. \tag{18}$$

While the parameter estimates of the one-step and two-step procedure are consistent for any choice of H based on the assumption of spherical disturbances, this does not hold for the variance estimator specified in (16) as, in general, $\mathbf{Z}'\mathbf{H}\mathbf{Z}$ is not a consistent estimator for $\mathrm{Var}[\mathbf{z}\varepsilon]$. Hence, the standard errors derived from this estimate would not be robust to heteroskedasticity and auto-correlation in the disturbances.

For the one-step estimator, a robust variance estimator can be obtained by approximating $\mathrm{Var}[\mathbf{z}\varepsilon]$ in (16) by $\frac{1}{N}\mathbf{Z}'\hat{\mathbf{\Omega}}\mathbf{Z}$ which is a consistent estimate of the former. The residuals for computing $\hat{\mathbf{\Omega}}$ are obtained from the one-step estimator using $\mathbf{A} = (\mathbf{Z}'\mathbf{H}\mathbf{Z})^{-1}$. The robust variance estimator is

$$\widehat{\mathrm{Avar}}[\hat{\beta}_{\text{one-step}}] = \begin{pmatrix} (\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{H}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{H}\mathbf{Z})^{-1}\mathbf{Z}'\hat{\mathbf{\Omega}}\times \\ \times\mathbf{Z}(\mathbf{Z}'\mathbf{H}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}(\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{H}\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X})^{-1} \end{pmatrix} \tag{19}$$

D System GMM estimator

This section derives the System GMM estimator as proposed by Arellano and Bover (1995) and Blundell and Bond (1998). Let us consider a special case of the linear error components model in (12) which has the following form:

$$y_{i,t} = \sum_{l=1}^{m} \alpha_{1,l} y_{i,t-l} + \sum_{l=0}^{m} \boldsymbol{\beta}'_{2,l} \mathbf{x}_{i,t-l} + \sum_{l=0}^{m} \boldsymbol{\gamma}'_{2,l} \mathbf{w}_{i,t-l} + \sum_{l=0}^{m} \boldsymbol{\delta}'_{2,l} \mathbf{v}_{i,t-l} + u_{i,t} \tag{20}$$

$$u_{i,t} = \mu_i + \varepsilon_{i,t} \tag{21}$$

where the $k \times 1$ vector $\mathbf{x}_{i,t}$, the $p \times 1$ vector $\mathbf{w}_{i,t}$ and the $q \times 1$ $\mathbf{v}_{i,t}$ contain k strictly exogenous, p predetermined (but not strictly exogenous) and q endogenous variables, respectively. μ_i are unobserved group specific fixed effects. Similar to Ahn and Schmitz's (1995) and Blundell and Bond's (1998) discussion of a dynamic AR(1) model, we assume that all observations are independently distributed across individuals as well as the following for $i = 1, \ldots, N$:

$$\mathrm{E}[\mu_i] = \mathrm{E}[\varepsilon_{i,t}] = 0 \ \text{ for } t = m+1, \ldots, T \tag{22}$$

$$\mathrm{E}[\mu_i \varepsilon_{i,t}] = 0 \ \text{ for } t = m+1, \ldots, T \tag{23}$$

$$\mathrm{E}[\varepsilon_{i,t} \varepsilon_{i,s}] = 0 \ \ \forall t \neq s \tag{24}$$

$$\mathrm{E}[y_{i,1} \varepsilon_{i,t}] = \ldots = \mathrm{E}[y_{i,m} \varepsilon_{i,t}] = 0 \ \text{ for } t = m+1, \ldots, T \tag{25}$$

$$\mathrm{E}[\mathbf{x}_{i,1} \varepsilon_{i,t}] = \ldots = \mathrm{E}[\mathbf{x}_{i,m} \varepsilon_{i,t}] = \mathbf{0} \ \text{ for } t = m+1, \ldots, T \tag{26}$$

$$\mathrm{E}[\mathbf{w}_{i,1} \varepsilon_{i,t}] = \ldots = \mathrm{E}[\mathbf{w}_{i,m} \varepsilon_{i,t}] = \mathbf{0} \ \text{ for } t = m+1, \ldots, T \tag{27}$$

$$\mathrm{E}[\mathbf{v}_{i,1} \varepsilon_{i,t}] = \ldots = \mathrm{E}[\mathbf{v}_{i,m} \varepsilon_{i,t}] = \mathbf{0} \ \text{ for } t = m+1, \ldots, T \tag{28}$$

Note that the AR process considered here which is of order m and includes covariates other than the lagged dependent variable requires more initial conditions specified in (25)-(28) than are imposed by Ahn and Schmidt (1995). Under the assumptions made in (22)-(28)

the idiosyncratic error terms $\Delta\varepsilon_{i,t}$ of (20) in first differences obviously satisfy the following set of orthogonality conditions for all $i = 1, \ldots, N$:

$$\mathrm{E}[y_{i,s}\Delta u_{i,t}] = 0 \;\; \forall s \leq t-2 \text{ with } t = m+2, \ldots, T_i \tag{29}$$

$$\mathrm{E}[\mathbf{x}_{i,s}\Delta u_{i,t}] = \mathbf{0} \;\; \text{for } s = 1, \ldots, T_i \text{ and } t = m+2, \ldots, T_i \tag{30}$$

$$\mathrm{E}[\mathbf{w}_{i,s}\Delta u_{i,t}] = \mathbf{0} \;\; \forall s \leq t-1 \text{ with } t = m+2, \ldots, T_i \tag{31}$$

$$\mathrm{E}[\mathbf{v}_{i,s}\Delta u_{i,t}] = \mathbf{0} \;\; \forall s \leq t-2 \text{ with } t = m+2, \ldots, T_i \tag{32}$$

These moment conditions can be rewritten in compact form as

$$\mathrm{E}[\mathbf{Z}_i^{dif\prime}\boldsymbol{\Delta}\mathbf{u}_i] = 0 \tag{33}$$

where

$$\boldsymbol{\Delta}\mathbf{u}_i = (\Delta\varepsilon_{i,m+2}, \Delta\varepsilon_{i,m+3}, \ldots, \Delta\varepsilon_{i,T_i})' \tag{34}$$

is a $(T_i - m - 1)$ vector comprising the idiosyncratic error terms of the first-differenced model and $\mathbf{Z}_i^{dif}$ is a block diagonal matrix, $\mathbf{Z}_i^{dif} = \mathrm{diag}[\mathbf{Z}_i^{dif,(m+2)}, \mathbf{Z}_i^{dif,(m+3)}, \ldots, \mathbf{Z}_i^{dif,(T_i)}]$, with

$$\mathbf{Z}_i^{dif,(t)} = (y_{i,1}, \ldots, y_{i,t-2}, \mathbf{x}_{i,1}', \ldots, \mathbf{x}_{i,T_i}', \mathbf{w}_{i,1}', \ldots, \mathbf{x}_{i,t-1}', \mathbf{v}_{i,1}', \ldots, \mathbf{x}_{i,t-2}') \tag{35}$$

Using this instrument matrix, the one-step and two-step Difference GMM estimators can be derived as outlined in section B.

As argued by Blundell and Bond (1998), however, the level instruments used in the Difference GMM estimator are weak in two cases: first, if the processes driving the variables in $\mathbf{x}_{i,t}$, $\mathbf{w}_{i,t}$ and $\mathbf{v}_{i,t}$ are highly persistent; second, if the variance of the fixed effects compared to the variance of the idiosyncratic error term is large. This is because the information about

any instrumented variable $\Delta\omega_{i,t}$ contained in the respective lags of $\omega_{i,t}$ used as instruments converges to zero as $\omega_{i,t}$ approaches a random walk process or as the relative variance of the fixed effects approaches infinity.

To raise the efficiency of the GMM estimator, Blundell and Bond (1998) and Blundell et al. (2000) derive the System GMM estimator by identifying an additional set of orthogonality conditions obtained from the equation in levels using as instruments the right-hand side variables in first differences. For the multivariate case considered here the moment conditions are, for all $i = 1, \ldots, N$,

$$\mathrm{E}[\Delta y_{i,t-1} u_{i,t}] = 0 \text{ for } t = m+2, \ldots, T_i \tag{36}$$

$$\mathrm{E}[\boldsymbol{\Delta}\mathbf{x}_{i,t} u_{i,t}] = \mathbf{0} \text{ for } t = m+1, \ldots, T_i \tag{37}$$

$$\mathrm{E}[\boldsymbol{\Delta}\mathbf{w}_{i,t} u_{i,t}] = \mathbf{0} \text{ for } t = m+1, \ldots, T_i \tag{38}$$

$$\mathrm{E}[\boldsymbol{\Delta}\mathbf{v}_{i,t-1} u_{i,t}] = \mathbf{0} \text{ for } t = m+2, \ldots, T_i \tag{39}$$

which require that each of the variables contained in $\Delta y_{i,t}$, $\boldsymbol{\Delta}\mathbf{x}_{i,t}$, $\boldsymbol{\Delta}\mathbf{w}_{i,t}$ and $\boldsymbol{\Delta}\mathbf{v}_{i,t}$ are uncorrelated with μ_i. Blundell et al. (2000) show for the multivariate case that a sufficient (but not necessary) condition for $\mathrm{E}[\omega_{i,t}\mu_i] = 0$ with $\omega_{i,t}$ being any covariate is that $\omega_{i,t}$ is mean-stationary.

The moment conditions in (36)-(39) read in compact matrix notation

$$\mathrm{E}[\mathbf{Z}_i^{lev\prime} \mathbf{u}_i] = 0 \tag{40}$$

where

$$\mathbf{u}_i^{lev} = (\varepsilon_{i,m+1}, \varepsilon_{i,m+2}, \ldots, \varepsilon_{i,T_i})' \tag{41}$$

is a $(T_i - m)$ vector comprising the idiosyncratic error terms of the level equation and $\mathbf{Z}_i$ is

a block diagonal matrix, $\mathbf{Z}_i^{lev} = \text{diag}[\mathbf{Z}_i^{lev,(m+1)}, \mathbf{Z}_i^{lev,(m+2)}, \ldots, \mathbf{Z}_i^{lev,(T_i)}]$, with

$$\mathbf{Z}_i^{lev,(t)} = \begin{cases} (0, \boldsymbol{\Delta}\mathbf{x}_{i,t}', \boldsymbol{\Delta}\mathbf{w}_{i,t}', \mathbf{0}') & \text{for } t = m+1 \\ (\Delta y_{i,t-1}, \boldsymbol{\Delta}\mathbf{x}_{i,t}', \boldsymbol{\Delta}\mathbf{w}_{i,t}', \boldsymbol{\Delta}\mathbf{v}_{i,t-1}') & \text{for } t = m+2, \ldots, T_i \end{cases} \tag{42}$$

The System GMM estimator can be obtained by combining the moment conditions of the difference equation and the moment conditions of the level equation, i.e.

$$\text{E}[\mathbf{Z}_i^{sys\prime}\mathbf{u}_i^{sys}] = 0 \tag{43}$$

where

$$\mathbf{u}_i^{sys} = (\boldsymbol{\Delta}\mathbf{u}_i', \mathbf{u}_i')' \tag{44}$$

and

$$\mathbf{Z}_i^{sys} = \begin{bmatrix} \mathbf{Z}_i^{dif} & 0 \\ 0 & \mathbf{Z}_i^{lev} \end{bmatrix}. \tag{45}$$

is a stacked matrix of the difference and level instruments. After stacking the observation matrices $\mathbf{y}$ and $\mathbf{X}$ correspondingly, the one-step and two-step System GMM estimators can be derived as described in section B.

E Robustness check of the estimation results

To check the robustness of the results, we consider some variations of the model estimated above. In particular, we consider the following modifications: First, we vary the lag length and additionally consider models with $L = 3$ and $L = 5$. Second, we use gross values of

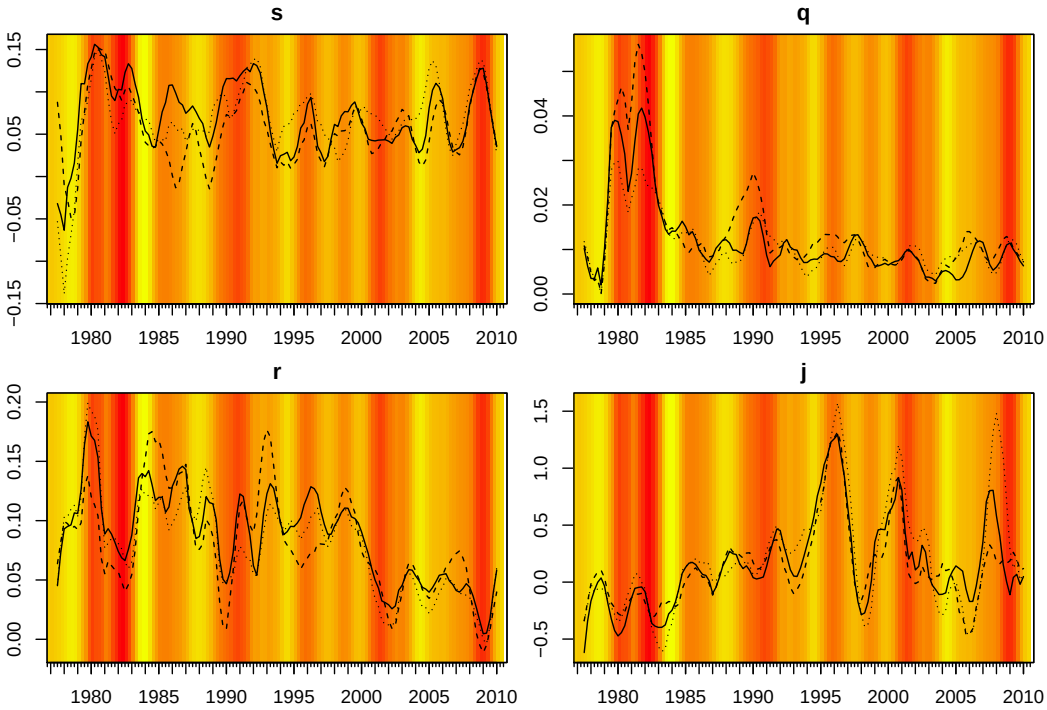

Figure 7: Regression results for the two-step System GMM estimation of (1) with $L = 3$ (dashed line), $L = 4$ (solid line) and $L = 5$ (dotted line)

the variables in the investment function rather then values net of depreciation. Third, we estimate the investment model for the manufacturing sector only.

Figure 7 plots the results of the recursive two-step System GMM estimation of (1) for $L = 3, 4, 5$. Overall, the results derived from the baseline specification with $L = 4$ are fairly robust to the lag length chosen. Some notable deviations from the standard specification occur around 1986 for the sales-growth coefficient and in 1995 for the cash-flow coefficient.

To further check the robustness of our results we consider an investment function with

variables gross of depreciation, i.e

$$g^g_{i,t} = \sum_{k=1}^{L} \beta_{g,k} g^g_{i,t-k} + \sum_{k=0}^{L} \beta_{s,k} s^g_{i,t-k} + \sum_{k=0}^{L} \beta_{r,k} r^g_{i,t-k} + \sum_{k=0}^{L} \beta_{q,k} q_{i,t-k} \\ + \sum_{k=0}^{L} \beta_{j,k} j_{i,t-k} + \sum_{s=0}^{4} \beta_{d,s} d_s + \mu_i + \varepsilon_{i,t} \tag{46}$$

where $g^g_{i,t}$ is the gross accumulation rate, $s^g_{i,t}$ is the growth rate of gross sales and $r^g_{i,t}$ is the ratio between cash flow including depreciation and the net capital stock. Again, $L = 4$.

The smoothened long-run coefficients estimated using two-step System GMM on a 5-quarter rolling window are plotted in Figure 8 for the investment specification in net and in gross values. The coefficients for sales growth and Tobin's q are fairly consistent with each other which is not surprising as the growth rates of net and gross sales are very similar and Tobin's q is the same in both specifications. The cash-flow and cost-of-capital coefficients are, on average, higher than the ones in the baseline specification, but exhibit the same cyclical pattern.

Finally, the estimates for the total corporate business sector excluding financial, insurance and real estate services are compared to estimates for different sub-sectors, in particular for all industries including agriculture, forestry, fishing, mining, construction, manufacturing, transportation, communications, electric, gas, and sanitary services (SIC codes smaller than 5000) as well as the manufacturing sector alone (SIC codes between 3000 and 4000).

The estimation results are plotted in Figure 9. Note that total industries and manufacturing imply only 3/4 and 1/2 of the observations available for the total business sector, respectively. This may contribute to the higher volatility of the coefficients of the smaller sectors. The results for the total industries are very similar to the baseline results. Exceptions are the sales-growth coefficient which does not spike in the late 1990s, the cash-flow coefficient spiking in the late 1970s which might be a consequence of a lack of observations

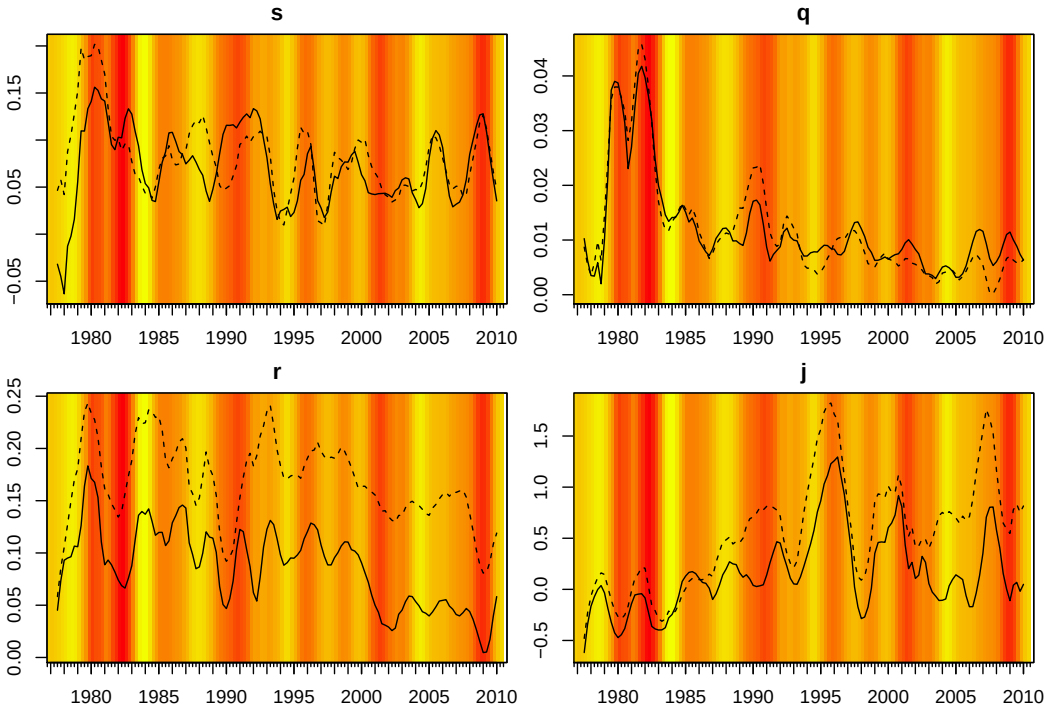

Figure 8: Regression results for the two-step System GMM estimation of (1) in net variables (solid line) and of (46) in gross variables (dashed line)

and in the mid 1980s as well as the coefficient of Tobin's q deviating from the baseline result in the mid 1980s.

The coefficients obtained for the manufacturing sector alone exhibit more significant deviations. It is worth to note that the manufacturing sector seems to be affected by demand constraints prior to the rest of the economy as the sales-growth coefficient seem to spike slightly before the average over the total business sector. Further, demand constraints seem to be more pronounced for the manufacturing sector.

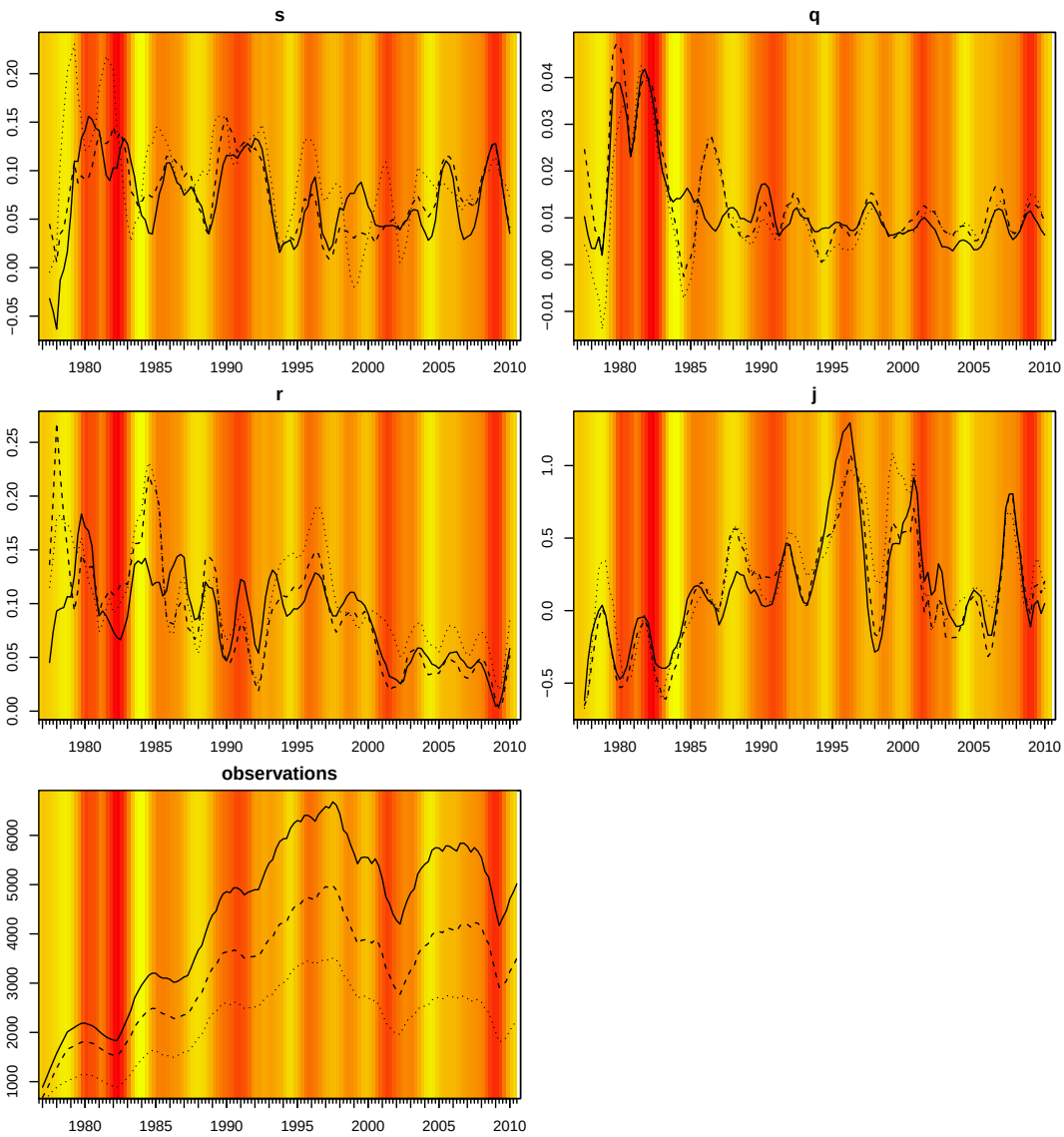

Figure 9: Regression results for the two-step System GMM estimation of (1) for the the total business sector (solid line), for the industries (dashed line), and for the manufacturing sector (dotted line)