Financial Frictions and Capital Misallocation

Jiaqi Li*
Bank of Canada

Nov 30, 2019

[Link to the Latest Version]

Abstract

While it is widely perceived that financial frictions have adverse impact on capital allocation, the importance of this impact is difficult to quantify. This paper provides a novel two-step approach to estimate the importance of financial frictions on capital misallocation, measured by the dispersion of the marginal revenue product of capital. First, based on the theoretical result that the capital investment of financially constrained firms is more sensitive to their internal financing than for unconstrained firms, I use a switching regression approach to jointly estimate the two different investment regimes and the probability of each firm being constrained. Firms are classified as financially constrained or unconstrained based on the estimated probabilities. Second, I provide a decomposition of capital misallocation and estimate the fraction that can be explained by the presence of financially constrained firms. Applying this method to large panels of manufacturing firms for 20 countries from the 1990s to 2015, this paper finds that for most countries and two-digit industries, more than a quarter of firms are classified as financially constrained. Furthermore, the presence of these constrained firms accounts for more than half of capital misallocation.

JEL classification: E44, L60, O16

*Bank of Canada. Email: jiaqili@bankofcanada.ca.
I am deeply indebted to my Ph.D. supervisor, Petra Geraats, and my Ph.D. advisor, Donald Robertson, for their valuable advice and continuous help. I am extremely grateful to my supervisor at the International Monetary Fund, Jérôme Vandenbussche, who provides me with the Orbis firm-level data to carry out the empirical analysis and gives me some valuable feedback. This paper was inspired from the project on the riskiness of credit allocation during my internship at the International Monetary Fund. I want to thank Thorsten Beck, Tiago Cavalcanti, Lu Han, and Pete Klenow for their insightful comments and suggestions.

1 Introduction

Capital misallocation has important implications on aggregate productivity (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2008) and understanding the causes of capital misallocation is one of the central topics in the literature (Gopinath et al., 2017; Midrigan and Xu, 2014; Bartelsman, Haltiwanger and Scarpetta, 2013).[1] Capital misallocation refers to the inefficient distribution of existing capital stock across producers given their productivity. As a contributing factor for capital misallocation, financial frictions or credit market imperfections have received a lot of attention.[2] More recently, there are a few papers aiming to quantify the impact of financial frictions on capital misallocation by estimating a structural model (Bai, Lu and Tian, 2018; David and Venkateswaran, 2017; Midrigan and Xu, 2014).

In this paper, I provide an alternative approach to quantify the impact of financial frictions on capital misallocation, which requires fewer restrictive assumptions and uses more information from large firm-level datasets. The approach consists of two steps. First, firms are classified as financially constrained or unconstrained using a switching regression approach. The idea is that the investment of the two types of firms follows two different processes, since the investment of constrained firms should be more responsive to cash flow than that of unconstrained firms. The probability of a firm being constrained is used to classify firms and is jointly estimated with the two different investment regimes.

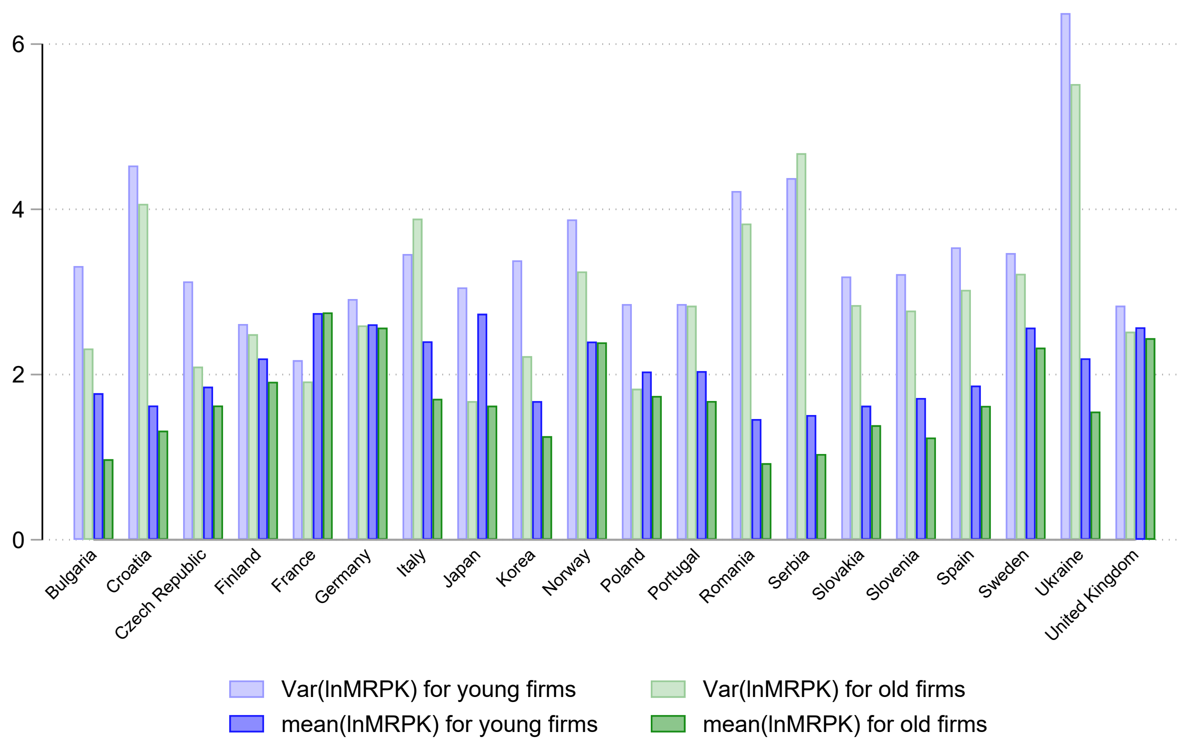
Second, assuming that the distribution of the observed marginal revenue product of capital (MRPK) is a mixture of two distributions, one for each type of firm, this paper provides a statistical decomposition of capital misallocation, which is measured by the dispersion (variance) of MRPK. Since the efficient allocation of capital in a neoclassical model indicates equalisation of the MRPK across firms, capital misallocation can be indirectly measured by the dispersion of the MRPK across firms within a given industry (Restuccia and Rogerson, 2017). This is motivated by the fact that younger firms who are more likely to be financially constrained have a higher dispersion and mean of MRPK than older firms, as can be seen in Figure 1. For most countries, it is highly statistically significant that both the means and the dispersions of MRPK for the young firms are larger than those for the old firms.[3]

[1] See also Busso, Madrigal and Pagés (2013), Restuccia and Rogerson (2013).

[2] The potential impact of financial frictions on resource misallocation and total factor productivity losses has received a lot of attention in the recent literature. See Gopinath et al. (2017), Gilchrist, Sim and Zakrajšek (2013), Banerjee and Duflo (2005), etc.

[3] One-sided t-tests and F-tests are used to test whether young firms have a higher mean and variance of log MRPK than old firms respectively in each country. The p-values from the t-test are smaller than 0.001 for all countries except for France, Germany, Norway, the UK, while the p-values from the F-test are smaller than 0.001 for all countries except for Croatia, Finland, Italy, Portugal, and Serbia. A small p-value for the t-test (F-test) rejects the null hypothesis that the means (variances) of ln MRPK are equal between the two types of firms.

Figure 1: Dispersions and Means of Marginal Revenue Product of Capital (MRPK) in 2015



Note: The bar chart shows the cross-section variances (or dispersions) and means of the logarithm of the MRPK for young firms (age $\leqslant$ 15 years) and old firms (age $>$ 15 years) in each of the 20 selected countries in 2015. Firm age is computed as the difference between the year (2015) and the incorporation year plus one. MRPK (output elasticity of capital multiplied by output over capital stock based on a Cobb-Douglas production function) is computed as the nominal revenue (proxy for output) divided by fixed tangible assets (proxy for capital stock). Since the output elasticity of capital does not affect the dispersion of the MRPK within a given industry, it is neglected in the computation of the MRPK.
Data source: Orbis

Using the decomposition and the classified types estimated using large panels of manufacturing firms for 20 countries from the 1990s to 2015, this paper finds that the dispersions and means of MRPK for the financially constrained types are much larger than those for the unconstrained firms. For most countries and two-digit industries, more than a quarter of firms are classified as financially constrained and the presence of constrained firms can account for more than half of the observed dispersion of MRPK across firms.

This paper contributes to two strands of literature, namely, the macro literature on capital misallocation and the empirical finance literature on the impact of financial frictions on firm investment.

Financial frictions are often regarded as one of the leading contributing factors for capital misallocation and there are a few papers attempting to quantify the impact of financial frictions on capital misallocation by estimating a structural model. The importance of financial frictions is either implied by the estimated parameters or the predictions of quantitative models. However, whether financial frictions would cause large aggregate productivity loss via the capital misallocation channel remains unclear (Wu, 2018), which is likely due to the different modelling assumptions and datasets used.

There is evidence that financial frictions play an important role in generating the dispersion of MRPK in Spain and China. For instance, Gopinath et al. (2017) find that a size-dependent borrowing constraint is essential in generating the large increase in the dispersion of MRPK among Spanish manufacturing firms during 1999 and 2007. David and Venkateswaran (2017) use a quantitative model to find that firm-specific factors that are correlated with productivity, including financial frictions, account for around 47% of the MRPK dispersion using data for Chinese manufacturing plants from 1998 to 2009. Similar evidence can be found in Bai, Lu and Tian (2018) who estimate their model using Chinese private manufacturing plants during 1998-2007. However, there are also papers showing that financial frictions only cause moderate efficiency losses through capital misallocation.[4] For instance, Midrigan and Xu (2014) calibrate their model using Korean manufacturing plants during 1991-1996 and find that financial frictions in the form of borrowing constraints do not lead to substantial aggregate productivity losses via resource misallocation.

My paper contributes to this strand of literature by proposing a new method to estimate the impact of financial frictions on capital misallocation, which relies on fewer restrictive assumptions and thus can be readily applied to a large number of countries. More specifically, I come up with a credit distortion measure using the decomposition of the dispersion of MRPK, which measures the fraction of the dispersion of MRPK that can be attributed to the presence of financially constrained firms. I then compute this credit distortion measure in each two-digit manufacturing industry from 20 countries during the period of the 1990s to 2015.

By building a simple model of firm dynamics with capital adjustment costs in the form of a one-period time to build for capital and costly debt enforcement as a financial friction that gives rise to a borrowing constraint *à la* Kiyotaki and Moore (1997), this paper shows that the MRPK for constrained and unconstrained firms is determined by different processes. While the investment of an unconstrained firm is driven by expected future productivity growth or

[4] Gilchrist, Sim and Zakrajšek (2013) adopt a slightly different approach by applying an accounting framework (in which firm-specific borrowing costs are mapped into measures of resource misallocation) to US listed manufacturing firms and find that financial frictions are unlikely to be a major factor for resource misallocation, which is not very surprising as large listed firms tend to have better access to credit than small unlisted firms.

expected future sales growth, that of a constrained firm is driven by the availability of its internal financing.

Given the distribution of the observed MRPK is a mixture of two distributions, one for each type of firms, I decompose the dispersion of MRPK across all firms into the dispersions and means of MRPK within the unconstrained and constrained groups of firms. This statistical decomposition gives new insights into the mechanisms through which the presence of financially constrained firms affects the extent of the observed dispersion of MRPK. While the usual mechanism emphasizes that the higher MRPK of constrained firms relative to the unconstrained firms would lead to a higher dispersion of MRPK, the decomposition in this paper shows that the dispersions of MRPK within different groups of firms also matter.[5] More importantly, this paper provides a new credit distortion measure using this decomposition, which measures the fraction of the dispersion of MRPK that is caused by the presence of constrained firms, but requires information on firms' financially constrained status.

The empirical finance literature on financial frictions and firm investment has proposed various ways to classify firms into constrained and unconstrained groups.[6] One common approach is to divide firms based on one indirect proxy for financial constraints, such as dividend payout, age, size or leverage (e.g., Moshiriana et al., 2017; Carpenter and Guariglia, 2008; Alti, 2003; Hubbard, Kashyap and Whited, 1995; Fazzari et al., 1988). As a direct extension of this approach, there are a lot of index-based measures of financial constraints that have been built on various combinations of firm characteristics (e.g., Mulier, Schoors and Merlevede, 2016; Hadlock and Pierce, 2010; Whited and Wu, 2006; Lamont, Polk and Saaá-Requejo, 2001).[7]

Alternatively, instead of identifying constrained firms based on some a priori criteria, a switching regression model could be used to simultaneously estimate the probability that firms are financially constrained and the two different investment regimes for constrained and

[5]For example, the mechanism in Gopinath et al. (2017) operates through the increasing gap in the MRPK between the constrained and unconstrained firms. In response to a decline in the real interest rate, unconstrained firms increase their capital demand and experience a decline in their MRPK, while the constrained firms are not able to invest more capital and their MRPK does not fall, leading to an increased dispersion of the MRPK in the sector.

[6]To test whether financial frictions affect firm investment, classification of firms into constrained and unconstrained groups is often a prerequisite step after which the differential investment behavior between the two groups of firms can be tested.

[7]Lamont, Polk and Saaá-Requejo (2001) construct the KZ index, which is a weighted sum of five accounting ratios, using the regression coefficients from Kaplan and Zingales (1997) as the weights. An alternative index measuring the degree of financial constraints (WW index) was constructed by Whited and Wu (2006), based on estimating an investment Euler equation from a structural model. Hadlock and Pierce (2010) question the validity of the KZ index and WW index and propose a new measure based on firm size and age only, arguing that these two firm characteristics are particularly useful in predicting the levels of financial constraints.

unconstrained firms (e.g., Almeida and Campello, 2007; Hovakimian and Titman, 2006; Hu and Schiantarelli, 1998). The two investment regimes differ in terms of the sensitivity to cash flow. That is, firm investment should be more sensitive to cash flow for constrained firms after controlling for the investment opportunity. The probability estimated using this maximum likelihood approach can be used to classify firms into constrained and unconstrained groups.

In fact, the switching regression model is closely related to the index-based approach. The index-based measure of financially constrained status can also give the probability of a firm being constrained via a logit or probit function. However, this probability does not use any model structure or data information of the two investment regimes, unlike in a switching regression. This paper builds on the switching regression model to classify firms into constrained and unconstrained firms.

This paper contributes to this strand of the empirical finance literature by providing evidence for unlisted firms and more countries, using new proxies for investment opportunity that are motivated by the theoretical model. Existing literature often uses Tobin's q as a proxy for investment opportunity in a sample of US listed firms.[8] However, q is not available for unlisted firms, so this paper uses sales growth, value added growth and productivity growth as different proxies for investment opportunity, in order to analyse unlisted firms that are more likely to suffer from financial constraints. Furthermore, instead of focusing on the US firms, this paper provides new evidence for manufacturing firms in 20 countries from the 1990s to 2015.

The remainder of the paper is structured as follows. Section 2 shows the theoretical framework and the decomposition of the dispersion of MRPK. Section 3 describes the data and the summary statistics. Section 4 shows two different empirical specifications and the corresponding empirical results. Section 5 concludes.

2 Theoretical Framework

This section builds a simple model of firm dynamics with one-period time to build for capital and a borrowing constraint. The model is used to show that the capital demand and hence the marginal revenue product of capital (MRPK) for unconstrained and constrained firms are driven by different processes. I then decompose the dispersion of MRPK across all firms into the dispersions and means within the two types of firms. The model is also used to derive the two different investment equations for empirical analysis in Section 4.

[8] In a model with convex capital adjustment costs, marginal q or the shadow value of one additional unit of capital is a sufficient statistic for investment. However, as pointed out by Schiantarelli (1995), stock markets may be inefficient and if stock prices are driven by fads, q may not be a good proxy for investment opportunity.

2.1 Modeling Preliminaries

Assume there are M monopolistically competitive firms in a specific subsector s of the manufacturing industry, which are infinitely lived, each producing a differentiated product. Firms are indexed by i, where $i = 1, ..., M$.[9] For notational simplicity, sector subscripts are suppressed in this theory section. The total industry output y_t is a constant elasticity of substitution (CES) aggregate of M differentiated products:

$$y_t = \left(\sum_{i=1}^{M} y_{i,t}^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}} \tag{1}$$

where $y_{i,t}$ is the real output produced by firm i in period t, and $\epsilon > 1$ is the elasticity of substitution between varieties. Each firm i in period t produces output $y_{i,t}$ using capital $k_{i,t-1}$, which is predetermined (i.e., purchased and installed in period $t-1$), materials $m_{i,t}$, and labor $l_{i,t}$ via an industry-specific Cobb-Douglas production function:

$$y_{i,t} = A_{i,t} k_{i,t-1}^{\alpha_k} m_{i,t}^{\alpha_m} l_{i,t}^{\alpha_l} \tag{2}$$

where $A_{i,t}$ is the firm-specific physical productivity or total factor productivity (TFP), and $\alpha_k \in (0,1)$, $\alpha_m \in (0,1)$ and $\alpha_l \in (0,1)$ are the industry-specific output elasticities of capital, materials and labor, respectively. Assume constant returns to scale such that $\alpha_k + \alpha_m + \alpha_l = 1$.

Firms engage in monopolistic competition and each of them charges a price $p_{i,t}$ for their differentiated product i. Given the aggregate output index y_t (1), it can be calculated from the cost minimization problem of the buyers of the industry output that each firm faces a downward-sloping demand with a constant elasticity $\epsilon > 1$ for their product:

$$y_{i,t} = \left(\frac{p_{i,t}}{p_t} \right)^{-\epsilon} y_t \tag{3}$$

where both the industry output y_t and the industry price p_t are normalized to one in this partial equilibrium model, following Gopinath et al. (2017). As a result, combining the production function (2) and the demand for the firm's product (3), the revenue-based production function is:

$$p_{i,t} y_{i,t} = Z_{i,t} k_{i,t-1}^{\beta_k} m_{i,t}^{\beta_m} l_{i,t}^{\beta_l} \tag{4}$$

where $Z_{i,t} \equiv A_{i,t}^{\frac{\epsilon-1}{\epsilon}}$ is the revenue-based productivity or TFPR, and $\beta_k \equiv \alpha_k \frac{\epsilon-1}{\epsilon}$, $\beta_m \equiv \alpha_m \frac{\epsilon-1}{\epsilon}$ and $\beta_l \equiv \alpha_l \frac{\epsilon-1}{\epsilon}$ are the industry-specific revenue elasticities of capital, materials and labor

[9] This partial equilibrium model can be used to describe firm dynamics within the manufacturing industry, as well as any subsector of it, as used in the empirical analysis in Section 4.

respectively.[10] The revenue-based production function is often used in the literature because firm-level prices $p_{i,t}$ and output $y_{i,t}$ are often unavailable, while $p_{i,t}y_{i,t}$ can be empirically measured by nominal revenue or sales.[11]

Assume that the firm-specific revenue-based productivity can be decomposed into the product of three independent components, so $Z_{i,t} \equiv Z_t z_i z_{i,t}$, with a common trend Z_t, a firm-specific component z_i and an idiosyncratic component $z_{i,t}$, where the latter follows an AR(1) process in logs:

$$\ln z_{i,t} = \rho \ln z_{i,t-1} + e_{i,t} \tag{5}$$

with $\rho \in (0,1)$ indicating the persistence of the process, and $e_{i,t} \sim N(0, \sigma_z^2)$ being an independent and identically normally distributed random variable with mean zero and variance σ_z^2.

2.2 Firm's Capital Choice and Financial Frictions

Assume firms own the capital, which depreciates at a rate $\delta \in [0,1]$. They also purchase and install new capital each period for production in the following period. Assuming they start with different levels of initial net worth $n_{i,0}$ at $t = 0$, firms with low initial net worth may need to borrow at an exogenous real gross interest rate R_0 to finance the purchase of physical capital $k_{i,0}$. Similarly, firms with enough net worth to finance the capital save at the same interest rate. Firms install the purchased capital $k_{i,0}$, and at the beginning of $t = 1$, the productivity shocks realize and output $y_{i,1}$ is produced using the installed capital $k_{i,0}$, labor $l_{i,1}$ and materials $m_{i,1}$. Assume firms hire labor and acquire materials in a competitive market at an exogenous real wage rate w_t and real price of materials $p_{m,t}$ in each period t. Let $n_{i,t}$ denote the firm's net worth before its choice of $k_{i,t}$ and any borrowing $b_{i,t}$ in period t.

Financial friction is modelled via a costly debt enforcement problem, based on Kiyotaki and Moore (1997). In other words, borrowers cannot be forced to repay unsecured debt. Since creditors recognize the possibility of default by borrowers, they would never lend more than what they can obtain in the case of default. Hence, each firm would face a borrowing constraint that is tied to the value of their collateral, which is the value of their undepreciated capital:

$$b_{i,t} \leqslant \phi(1-\delta)k_{i,t} \tag{6}$$

[10] Note that $\beta_k + \beta_m + \beta_l = \frac{\epsilon - 1}{\epsilon}$.

[11] The notion of TFPR was introduced by Foster, Haltiwanger and Syverson (2008). The heterogeneity in TFPR across firms can reflect a combination of productivity differences and monopolistic pricing distortions.

where $\phi \in (0,1)$ is the loan-to-value ratio.[12] Assume firms are risk-neutral and in each period t, after the production of output and the payments of wage, materials and debt, there is a constant probability $\varphi \in (0,1)$ that the firm exits, in which case the firm consumes its net worth $n_{i,t}$. The surviving firms choose how much capital to purchase in period t given their net worth $n_{i,t}$. The exiting firms are replaced by new firms, with random levels of initial net worth, such that the total number of firms in the industry stays unchanged.[13] Given the firm faces a borrowing constraint, it is reasonable to assume that the firm delays consumption until the period it exits. Let $\eta \in (0,1)$ denote the firm's discount factor. Each firm i in period t chooses its capital $k_{i,t}$, labor $l_{i,t}$ and materials $m_{i,t}$ to maximize the expected discounted terminal net worth:[14]

$$\mathrm{E}_t \sum_{\tau=0}^{\infty} \varphi(1-\varphi)^{\tau}\eta^{\tau} n_{i,t+1+\tau} \tag{7}$$

subject to the borrowing constraint (6). The net worth $n_{i,t+1}$ at the beginning of period $t+1$ equals the sum of the revenue $p_{i,t+1}y_{i,t+1}$ and the undepreciated capital stock $(1-\delta)k_{i,t}$, net of the real wage cost $w_{t+1}l_{i,t+1}$, the real materials cost $p_{m,t+1}m_{t+1}$, and the gross debt interest payment $R_t b_{i,t}$:

$$n_{i,t+1} \equiv p_{i,t+1}y_{i,t+1} - w_{t+1}l_{i,t+1} - p_{m,t+1}m_{i,t+1} - R_t b_{i,t} + (1-\delta)k_{i,t} \tag{8}$$

Using the assumption that firms do not consume until the period they exit and (8), it can be shown that the firm finances the purchase of capital using either the internal financing (net worth) or external financing (debt):[15]

$$k_{i,t} = n_{i,t} + b_{i,t} \tag{9}$$

Using (9) to rewrite the borrowing constraint (6) in terms of net worth:

$$k_{i,t} \leqslant \frac{n_{i,t}}{1-\phi(1-\delta)} \tag{10}$$

[12] Since $b_{i,t}$ is used to buy part of the capital stock $k_{i,t}$, $k_{i,t}$ is not observed at the time of borrowing and there is a possibility that the firm absconds with the borrowed fund. To avoid this possibility, assume that the loan is conditional on the firm using it to purchase capital.

[13] One example to justify this assumption is that in each period, a random fraction of households start new firms using their savings as initial net worth.

[14] In period 0, the firm only chooses capital $k_{i,0}$ because there is no capital yet to produce.

[15] Suppose firm i consumes $c_{i,t+1}$ in period $t+1$, then the firm faces the following budget constraint:

$$c_{i,t+1} + k_{i,t+1} = b_{i,t+1} + p_{i,t+1}y_{i,t+1} - w_{t+1}l_{i,t+1} - p_{m,t+1}m_{t+1} - R_t b_{i,t} + (1-\delta)k_{i,t} \equiv b_{i,t+1} + n_{i,t+1}$$

Given that the firm does not consume until the period of exit, $c_{i,t+1} = 0$ and hence $k_{i,t+1} = b_{i,t+1} + n_{i,t+1}$. In the terminal period T, $b_{i,T} = k_{i,T} = 0$, so $c_{i,T} = n_{i,T}$.

Let $\lambda_{i,t}$ denote the Lagrange multiplier associated with the borrowing constraint (10), and let $k_{i,t}^U$ and $k_{i,t}^C$ denote firm i's unconstrained capital demand and constrained capital demand, respectively. It is shown in Appendix A.3 and A.5 that the firm's capital demand $k_{i,t}$ is:

$$k_{i,t} = \begin{cases} k_{i,t}^U & \text{if} \quad \lambda_{i,t} = 0 \\ k_{i,t}^C & \text{if} \quad \lambda_{i,t} > 0 \end{cases} \tag{11}$$

where the log of the capital demand of a constrained firm i is:

$$\ln k_{i,t}^C = \ln n_{i,t} - \ln[1 - \phi(1-\delta)] \tag{12}$$

and the log of the capital demand of an unconstrained firm i is:

$$\begin{aligned} \ln k_{i,t}^U =& \epsilon\rho \ln z_{i,t} + (1+\epsilon\beta_k)\left\{ \ln\left(\beta_k \beta_l^{\frac{\epsilon\beta_l}{1+\epsilon\beta_k}} \beta_m^{\frac{\epsilon\beta_m}{1+\epsilon\beta_k}}\right) - \ln(r_t + \delta) \right. \\ & \left. + \ln \mathrm{E}_t\left[\left(\frac{Z_{t+1}}{w_{t+1}^{\beta_l} p_{m,t+1}^{\beta_m}}\right)^{\frac{\epsilon}{1+\epsilon\beta_k}}\right] + \frac{\epsilon}{1+\epsilon\beta_k}\ln z_i + \frac{\sigma_z^2\epsilon^2}{2(1+\epsilon\beta_k)^2} \right\} \end{aligned} \tag{13}$$

where $r_t \equiv R_t - 1$ is the net real interest rate. Since productivity is assumed to follow an AR(1) process, the expected future productivity can be written in terms of the current productivity $z_{i,t}$. As can be seen, the capital demand of the unconstrained firm is increasing in its idiosyncratic transitory productivity $z_{i,t}$, permanent productivity z_i, and the expected future common productivity Z_{t+1}, and decreasing in the net real interest rate r_t and the expected factor prices (w_{t+1} and $p_{m,t+1}$). Intuitively, higher (expected) firm productivity (both common, permanent and idiosyncratic) leads to a higher demand for physical capital. A higher net interest rate increases the marginal cost of capital and thus reduces the capital demand. A higher real wage or price of materials reduces the demand for labor and materials respectively, leading to a lower marginal revenue product of capital (MRPK). Hence, capital demand falls to ensure that the expected MRPK equals the user cost of capital ($r_t + \delta$). By contrast, the constrained firm cannot operate at an optimal scale and its capital demand is constrained by its net worth $n_{i,t}$, as shown in (12).

2.3 Dispersion in Marginal Revenue Product of Capital

Define firm i's period-t marginal revenue product of capital $\mathrm{MRPK}_{i,t}$ as:

$$\mathrm{MRPK}_{i,t} \equiv \frac{\partial p_{i,t} y_{i,t}}{\partial k_{i,t-1}} = \beta_k Z_{i,t} k_{i,t-1}^{\beta_k - 1} l_{i,t}^{\beta_l} m_{i,t}^{\beta_m} = \beta_k \frac{p_{i,t} y_{i,t}}{k_{i,t-1}} \tag{14}$$

where $p_{i,t}y_{i,t}$ denotes the nominal revenue. Based on the model, if all firms were financially unconstrained, their expected MRPK is identical, as they face the same interest rate. There is still dispersion in ex post MRPK due to the different realizations of the productivity shocks across firms, but this source of dispersion is not treated as misallocation in the literature, since the allocation of capital is efficient ex ante (e.g., Restuccia and Rogerson, 2017; Asker, Collard-Wexler and De Loecker, 2014). Hence, capital misallocation should be measured by the dispersion of the expected MRPK.

However, it is difficult to measure the expected MRPK, so this paper measures capital misallocation by the static dispersion of MRPK across firms within a given two-digit industry.[16] As a result, capital misallocation may be overestimated if production inputs (such as capital in this paper) are chosen before the shock realizes. However, this paper does not attempt to disentangle the efficient causes of the dispersion. Assuming capital adjustment costs affect constrained and unconstrained firms equally, this paper only aims to estimate the proportion of the dispersion caused by the financial friction.

The static dispersion of MRPK across all firms, as shown in Appendix A.2, can be written as:

$$\begin{aligned}\mathrm{Var}_i(\ln\mathrm{MRPK}_{i,t}) =&\psi_1\mathrm{Var}_i(\ln z_i) + \psi_1\mathrm{Var}_i(e_{i,t}) + \psi_1\rho^2\mathrm{Var}_i(\ln z_{i,t-1}) + \psi_2\mathrm{Var}_i(\ln k_{i,t-1}) \\ &- \psi_3\mathrm{Cov}_i(\ln z_i + \rho\ln z_{i,t-1}, \ln k_{i,t-1})\end{aligned} \tag{15}$$

where Var_i and Cov_i denote the cross-section variance and covariance across firms in a given time period, and $\psi_1 \equiv \left(\frac{\epsilon}{1+\epsilon\beta_k}\right)^2$, $\psi_2 \equiv \left(\frac{1}{1+\epsilon\beta_k}\right)^2$, and $\psi_3 \equiv 2\frac{\epsilon}{(1+\epsilon\beta_k)^2}$ are positive coefficients. This is a general decomposition that holds regardless of the types of firms. In general, capital misallocation measured by the static dispersion of MRPK depends on the cross-section dispersions of idiosyncratic permanent productivity $\mathrm{Var}_i(\ln z_i)$, the productivity innovation $\mathrm{Var}_i(e_{i,t})$, the past productivity $\mathrm{Var}_i(\ln z_{i,t-1})$, and installed capital stock $\mathrm{Var}_i(\ln k_{i,t-1})$, and the cross-section covariance between the firm's capital and different components of the firm's productivity, as shown in (15). There is no dispersion of the marginal revenue product of labor (MRPL) in this model, as $\mathrm{MRPL}_{i,t} = \beta_l\frac{p_{i,t}y_{i,t}}{l_{i,t}} = w_t$, using the first order condition with respect to labor from the model.

In the absence of the one-period time to build and financial frictions, it can be shown that there is no dispersion of MRPK in this model.[17] In other words, there are two causes

[16] Note that $k_{i,t-1}$ is used in the model to reflect that capital is chosen in period $t-1$, but only used in period t. In empirical analysis, fixed tangible asset in period t is used to measure $k_{i,t-1}$, so the dispersion of MRPK (14) is still the static dispersion of MRPK.

[17] If firms are unconstrained and capital adjusts immediately in response to productivity shocks, firms can always borrow to finance their optimal demand for capital and their MRPK will be equalised within a given industry, as they face the same interest rate: $\mathrm{MRPK}_{i,t} = (r_t + \delta)$.

for the dispersion of MRPK: time-to-build for capital and financial frictions. First, due to a one-period time to build, $k_{i,t}$ is chosen based on the expected future productivity $\mathrm{E}_t Z_{i,t+1}$. Hence, any realized productivity $Z_{i,t+1}$ that differs from the expectation would cause the MRPK to differ across firms ex post.[18] This explains why the cross-section dispersion in the productivity innovation $\mathrm{Var}_i(e_{i,t})$ causes dispersion of MRPK. In fact, if all firms are unconstrained (i.e., without the financial frictions), $\mathrm{Var}_i(e_{i,t})$ is the only source of dispersion of MRPK. Let $\mathrm{Var}_i(\mathrm{lnMRPK}^U_{i,t})$ denote the cross-section variance of the log MRPK across financially unconstrained (U) firms, then it is shown in Appendix A.4 that:

$$\mathrm{Var}_i(\mathrm{lnMRPK}^U_{i,t}) = \psi_1 \mathrm{Var}_i(e_{i,t}) \tag{16}$$

where the superscript U denotes unconstrained firms. As can be seen, the dispersion of MRPK for unconstrained firms is purely driven by the cross-section dispersion in the productivity innovation $\mathrm{Var}_i(e_{i,t})$.

Second, financial frictions in the form of a collateral constraint also cause MRPK to differ across firms. Using the first order condition for capital, financially constrained firms have higher expected MRPK than unconstrained firms as they cannot borrow enough to finance their optimal capital demand, as shown in Appendix A. As a result, the differences of MRPK between constrained and unconstrained firms contribute to the overall dispersion of MRPK. Furthermore, this paper finds that the dispersion of MRPK within constrained firms is also important for understanding the overall dispersion of MRPK caused by financial frictions. Let $\mathrm{Var}_i(\mathrm{lnMRPK}^C_{i,t})$ denote the cross-section variance of the log MRPK across financially constrained (C) firms, then it is shown in Appendix A.6 that:

$$\begin{aligned}\mathrm{Var}_i(\mathrm{lnMRPK}^C_{i,t}) = &\psi_1 \mathrm{Var}_i(\mathrm{ln}z_i) + \psi_1 \mathrm{Var}_i(e_{i,t}) + \psi_1 \rho^2 \mathrm{Var}_i(\mathrm{ln}z_{i,t-1}) + \psi_2 \mathrm{Var}_i(\mathrm{ln}n_{i,t-1}) \\ &- \psi_3 \mathrm{Cov}_i(\mathrm{ln}z_i + \rho \mathrm{ln}z_{i,t-1}, \mathrm{ln}n_{i,t-1})\end{aligned} \tag{17}$$

where the superscript C denotes constrained firms. Note that $\mathrm{Var}_i(\mathrm{lnMRPK}^C_{i,t})$ now also depends on the variances and covariances of the logs of z_i, $z_{i,t-1}$ and $n_{i,t-1}$. Using (16) and (17), it can be shown that constrained firms have a higher dispersion of MRPK than unconstrained firms.[19] As can be seen from (17), the dispersion of MRPK for constrained firms

[18] Restuccia and Rogerson (2017) note that one problem with measuring misallocation using the dispersion of marginal products is that when inputs are chosen before firm-specific shocks realize, the marginal products across firms may not equalize in every time period even under efficient allocation. Similarly, Asker, Collard-Wexler and De Loecker (2014) pointed out that in the presence of capital adjustment costs, the ex ante efficient choice of capital can be inefficient ex post.

[19] Rearranging (17) gives:

$$\mathrm{Var}_i(\mathrm{lnMRPK}^C_{i,t}) = \psi_1 \mathrm{Var}_i(e_{i,t}) + \mathrm{Var}_i(\psi_1^{\frac{1}{2}} \mathrm{ln}z_i + \psi_1^{\frac{1}{2}} \rho \mathrm{ln}z_{i,t-1} - \psi_2^{\frac{1}{2}} \mathrm{ln}n_{i,t-1})$$

is lower when firms' net worth and productivity are more positively correlated. The firm's productivity and net worth are expected to be positively correlated since more productive firms tend to be more profitable, and thus accumulate more net worth over time. Intuitively, firms with higher productivity (in period $t-1$ or permanently) anticipate higher future productivity and hence would want to demand more capital. Since more productive firms also tend to have more net worth, their borrowing capacity is higher, which means they are able to borrow more while being constrained, and their capital demand is closer to their optimal unconstrained capital demand, bringing down their MRPK and thereby also the dispersion of MRPK among constrained firms.

This paper analyzes the impact of financial frictions on the dispersion of MRPK by empirically estimating the percentage of the dispersion that can be attributed to the presence of constrained firms. Suppose there are N_t unconstrained firms in an industry in a given time period t, and the remaining $M_t - N_t$ firms are constrained, then it is shown in Appendix B that the cross-section variance of MRPK across all firms can be decomposed as follows:

$$\begin{aligned}\mathrm{Var}_i(\ln\mathrm{MRPK}_{i,t}) =& \frac{N_t}{M_t}\mathrm{Var}_i(\ln\mathrm{MRPK}^U_{i,t}) + \frac{M_t - N_t}{M_t}\mathrm{Var}_i(\ln\mathrm{MRPK}^C_{i,t}) \\ &+ \frac{N_t(M_t - N_t)}{M_t^2}\left[\mathrm{E}_i(\ln\mathrm{MRPK}^U_{i,t}) - \mathrm{E}_i(\ln\mathrm{MRPK}^C_{i,t})\right]^2\end{aligned} \tag{18}$$

where the cross-section variances and means on the right hand side of (18) are defined for the two subgroups of firms. For example, $\mathrm{E}_i(\ln\mathrm{MRPK}^U_{i,t})$ denotes the cross-section mean of MRPK across unconstrained firms only. As can be seen from (18), the overall dispersion of MRPK equals a weighted average of the dispersion for unconstrained and constrained firms plus a measure of distance between the mean for each group. It is shown in (16) and (17) that the dispersion of MRPK within the unconstrained group is driven by the dispersion in productivity innovation while that within the constrained group is also driven by the dispersions in firms' net worth $n_{i,t-1}$, firms' idiosyncratic permanent productivity z_i, the realized idiosyncratic transitory productivity $z_{i,t-1}$, and the covariance between their productivity and net worth.

Constrained firms have a higher MRPK than unconstrained firms because of the lower level of capital that can be financed.[20] As a result, the cross-section average of MRPK for constrained firms $\mathrm{E}_i(\ln\mathrm{MRPK}^C_{i,t})$ is larger than that for the unconstrained firms $\mathrm{E}_i(\ln\mathrm{MRPK}^U_{i,t})$.[21]

---

where the second term on the RHS, which is strictly positive, is the only difference from the dispersion of MRPK for unconstrained firms (16). Hence, $\mathrm{Var}_i(\ln\mathrm{MRPK}^C_{i,t}) > \mathrm{Var}_i(\ln\mathrm{MRPK}^U_{i,t})$.

[20] This follows from the first order condition for capital as shown in Appendix A.

[21] As shown in (45) in Appendix A.1, on top of the effect of lower capital for constrained firms on increasing their MRPK, the realizations of their productivity $Z_{i,t}$ are also different for constrained and unconstrained

According to the decomposition of the variance of MRPK (18), the larger the gap in mean MRPK between the two groups of firms $\left[\mathrm{E}_i(\ln\mathrm{MRPK}^U_{i,t}) - \mathrm{E}_i(\ln\mathrm{MRPK}^C_{i,t})\right]$, the higher the cross-section variance of MRPK at an industry-year level. This resembles the usual mechanism that the presence of borrowing constraints increases the dispersion of MRPK due to the differences in MRPK between unconstrained and constrained firms.

More importantly, based on the decomposition (18), if financially constrained firms can be identified empirically, then the proportion of the dispersion of MRPK that is caused by the financial friction can be estimated by:

$$\text{Credit Distortion} \equiv \frac{\mathrm{Var}_i(\ln\mathrm{MRPK}_{i,t}) - \frac{N_t}{M_t}\mathrm{Var}_i(\ln\mathrm{MRPK}^U_{i,t})}{\mathrm{Var}_i(\ln\mathrm{MRPK}_{i,t})} \in [0,1] \tag{19}$$

where $\mathrm{Var}_i(\ln\mathrm{MRPK}_{i,t})$ is observable in data, while $\mathrm{Var}_i(\ln\mathrm{MRPK}^U_{i,t})$ is the cross-section variance defined over the unconstrained group of firms, which can only be calculated once the unconstrained firms are identified. Section 4.2 uses a switching regression approach to identify the two groups of firms, following Hu and Schiantarelli (1998).

If the constrained firms were not present, so $N_t = M_t$ and $\mathrm{Var}_i(\ln\mathrm{MRPK}_{i,t})$ equals $\mathrm{Var}_i(\ln\mathrm{MRPK}^U_{i,t})$ (as implied by (19)), then the credit distortion measure equals zero. If all firms were constrained, so $N_t = 0$, then the credit distortion measure equals one. Hence, the fraction of the dispersion of MRPK caused by the presence of constrained firms is based on the difference between the observed $\mathrm{Var}_i(\ln\mathrm{MRPK}_{i,t})$ and the counterfactual variance of MRPK without constrained firms, $\mathrm{Var}_i(\ln\mathrm{MRPK}^U_{i,t})$, which is normalised by the overall dispersion in MRPK.[22]

The credit distortion measure (19) may also capture any other structural differences across the two types of firms that lead to different dispersions of MRPK within the unconstrained and constrained firms. For instance, although this paper assumes a constant markup (i.e., $\frac{\epsilon}{\epsilon-1} > 1$) for all firms, in practice, larger firms tend to have higher markups, while smaller firms tend to have lower markups as they are more likely to operate in an environment that is close to perfect competition. If the markup dispersion is higher for larger, unconstrained firms, then $\mathrm{Var}_i(\ln\mathrm{MRPK}^U_{i,t})$ is also higher, leading to a lower credit distortion measure. The econometric analysis in Section 4.2 therefore aims to control for other structural differences between the two types of firms, to mitigate this issue.

Furthermore, this paper only considers a quantity-based credit distortion measure and

firms. More specifically, when the realized $Z_{i,t}$ is sufficiently high, a firm is more likely to be constrained as the firm demands more capital. As a result, the cross-section average productivity for constrained firms is higher than that for unconstrained firms, and this further explains why $\mathrm{E}_i(\ln\mathrm{MRPK}^C_{i,t}) > \mathrm{E}_i(\ln\mathrm{MRPK}^U_{i,t})$.

[22] Note that $\mathrm{Var}_i(\ln\mathrm{MRPK}^U_{i,t})$ is multiplied by the fraction of unconstrained firms $\frac{N_t}{M_t}$ to ensure that the measure is always positive.

does not incorporate any distortions in the price of credit such as the interest rate wedge between the borrowing and saving rate caused by imperfect banking competition. Thus, the current measure of credit distortion is likely to underestimate the full impact of credit distortions by neglecting that constrained firms also tend to face greater price distortions.

In addition, this paper assumes that the revenue elasticity of capital β_k is the same for all firms within a subsector of the manufacturing industry. As a result, β_k does not affect the dispersion of MRPK and is neglected when computing the MRPK. However, if β_k differs across unconstrained and constrained firms, this also contributes to the dispersion of MRPK for the two types of firms and thus affecting the credit distortion measure. Using (14) and the first order condition for materials, MRPK can be equally measured using nominal revenue $p_{i,t}y_{i,t}$ or value added $\text{VA}_{i,t}$:

$$\text{MRPK}_{i,t} = \beta_k \frac{p_{i,t}y_{i,t}}{k_{i,t-1}} = \frac{\beta_k}{1-\beta_m}\frac{\text{VA}_{i,t}}{k_{i,t-1}} \tag{20}$$

where $\text{VA}_{i,t} \equiv p_{i,t}y_{i,t} - p_{m,t}m_{i,t} = (1-\beta_m)p_{i,t}y_{i,t}$. Using nominal value added over fixed tangible assets to measure MRPK requires both the revenue elasticities of materials β_m and capital β_k to be identical across firms, which is more restrictive. As a result, this paper measures MRPK using nominal revenue over fixed tangible assets in the baseline analysis and uses nominal value added to measure MRPK for robustness checks.

3 Data

The firm-level data for different countries used in this paper are from the Bureau van Dijk's Orbis historical financial database, which provides annual financial information from firms' balance sheets and income statements from early 1990s to 2015.[23] The financial variables extracted from the Orbis historical financial database are combined with some time-invariant variables extracted from the Orbis rolling 10 years database. After extracting the variables from these two databases, the dataset for each country is cleaned following similar procedures as outlined in Kalemli-Ozcan et al. (2015). The full cleaning procedure and the summary statistics for the sample used for empirical analysis are shown in Appendix D.

One major advantage of this database is that it contains both listed and unlisted firms, unlike Compustat or Worldscope which only cover listed firms. Since unlisted firms are more likely to suffer from financial constraints than listed firms which tend to be larger, it is useful to study financial constraints using datasets that contain those unlisted firms. The

[23] The time series for some European countries start in 1990, while for many countries, the time series are shorter.

empirical analysis in this paper focuses on the manufacturing sector in each country because the capital stock can be proxied by fixed tangible assets, whereas in other industries, it is more difficult to define the capital stock.[24] Within the manufacturing sector, an industry is defined by either a two-digit or four-digit NACE Rev.2 code in this paper.

Table 1: Data Description for Each Country in the Baseline Sample

Country	Period	Observations	Obs/Year	Industries	Unlisted Firms
Bulgaria	1995-2015	119,346	5,683	223	0.983
Croatia	1998-2015	124,184	6,899	220	0.981
Czech Republic	1994-2015	176,420	8,019	289	0.995
Finland	1995-2015	163,600	7,790	227	0.992
France	1995-2015	1,316,144	62,674	229	0.994
Germany	1990-2015	255,056	9,810	298	0.975
Italy	1995-2015	1,716,653	81,745	302	0.998
Japan	1989-2015	593,512	21,982	199	0.959
Korea	2001-2015	817,068	54,471	198	0.973
Norway	1996-2015	109,826	5,491	217	0.990
Poland	1994-2015	167,273	7,603	236	0.981
Portugal	1998-2015	372,214	20,679	227	0.999
Romania	1995-2015	558,739	26,607	231	0.984
Serbia	1999-2015	165,237	9,720	235	0.930
Slovakia	1995-2015	76,190	3,628	228	0.980
Slovenia	1997-2015	93,570	4,925	213	0.991
Spain	1994-2015	1,428,899	64,950	230	0.999
Sweden	1997-2015	299,408	15,758	229	0.988
Ukraine	2001-2015	422,144	28,143	227	0.985
United Kingdom	1994-2015	294,092	13,368	230	0.966

Note: The sample from each country consists of manufacturing firms only. Period shows the time period covered in each cleaned country-specific dataset. Observations and Obs/year show the total number of firm-year observations and the average number of firms respectively during the period covered in a given country. Industries shows the number of unique four-digit NACE Rev.2 industries over the period covered in each country. The last column shows the fraction of observations coming from unlisted firms.

Table 1 shows some basic information on the datasets of the manufacturing sector for each of the 20 countries used in the baseline analysis, including the time period covered, total number of observations, average number of firms per year, unique number of four-digit industries, and the fraction of observations coming from unlisted firms. Since value added is

[24]Based on the two-digit NACE Rev.2 code, the manufacturing sector is in the range of 10 to 33. The descriptions for each two-digit industry can be found in Table 11 in Appendix D.

used to calculate the marginal revenue product of capital (MRPK) for robustness checks and value added is often less available than sales, the countries in the baseline sample are selected based on the total number of observations and the availability of value added.[25] The statistics reported in Table 1 are after dropping the observations with missing operating revenue and missing or zero fixed tangible assets, but before dropping those with missing value added. As can be seen from Table 1, in most countries, more than 98% of the observations are from unlisted firms.

Table 2 shows the medians of some main variables of interest in this paper. As can be seen, the median number of employees is below 20 in 16 out of 20 countries, which shows that the dataset contains a lot of small firms. The year-on-year change in the log of fixed tangible assets FTA (proxy for capital stock) measures the firm net investment, which is used as the dependent variable in the empirical analysis shown in Section 4. The median value of ΔlnFTA is negative as small firms account for a large proportion of the sample. The average ΔlnFTA is positive in most countries, as can be seen in Table 8 in Appendix D.

According to the model, the year-on-year change in the log of sales, the log of value added, or the log of productivity can be used as different proxies for the investment opportunity, as shown in (21) and (22) in Section 4.1. Nominal value added is computed as the difference between operating revenue and material costs for most countries.[26] If firm productivity were estimated accurately, then this would be a more exogenous measure than sales growth. However, there is no perfect measure for firm productivity. This paper uses productivity growth as a robustness check, which is estimated using the Wooldridge (2009) approach.[27]

The last column of Table 2 shows the dispersion in log marginal revenue product of capital MRPK. According to (14), $\text{MRPK}_{i,t}$ is calculated as nominal revenue $p_{i,t}y_{i,t}$ over the capital stock $k_{i,t-1}$.[28] It can be seen that there is a large variation in the dispersion of

[25] Countries are ranked according to their total number of observations and also their availability of value added (in terms of the percentage of the total observations). The two ranks have equal weight and the top 19 countries according to the weighted rank are selected plus Japan. Japan has low availability for value added but it has a large number of observations and it is an important country to look at since this paper focuses on the manufacturing firms. US is not selected due to the low availability of value added and the relatively low number of observations.

[26] The original 'value added' variable in Orbis is the sum of taxation, profit/loss for the period (equivalent to profit/loss after taxation plus the extraordinary and other profit/loss), cost of employees, depreciation and interest paid. The computed value added is used if it has more observations. Except for Germany, Japan (no data for materials costs), Portugal, Spain, and UK, the computed value added is used in the other countries in the baseline sample.

[27] The details of this estimation approach can be found in Appendix C.

[28] According to (14), $\text{MRPK}_{i,t}$ also depends on the revenue elasticity of capital β_k. However, since β_k is often estimated at a two-digit industry level to ensure enough number of observations and hence is the same across firms within the industry, it does not matter for the dispersion of log MRPK within a two-digit industry. As discussed in Section 4.1, when using value added to compute MRPK, β_m is assumed to be the same for each subsector as well, so it is better to use sales revenue to compute MRPK.

Table 2: Medians of Selected Variables for Each Country in the Baseline Sample

Country	Employees	Age	Δlog$k$	ΔlogSales	ΔlogVA	ΔlogTFPR	Var(logMRPK)
Bulgaria	14	9	-0.026	0.047	0.048	-0.002	2.86
Croatia	5	12	-0.038	0.020	0.031	-0.007	3.70
Czech Republic	23	12	-0.026	0.034	0.035	-0.001	2.80
Finland	8	15	-0.066	0.026	0.033	0.007	2.18
France	6	14	-0.102	0.024	0.026	0.004	1.76
Germany	27	21	-0.033	0.012	0.038	0.009	2.97
Italy	12	15	-0.033	0.020	0.026	0.001	3.23
Japan	17	32	-0.024	0.009	0.018		2.07
Korea	17	9	-0.000	0.084	0.101	0.005	2.79
Norway	9	12	-0.059	0.039	0.043	0.011	3.46
Poland	90	13	-0.019	0.048	0.052	0.003	2.34
Portugal	7	14	-0.074	0.012	0.019	0.001	2.47
Romania	6	9	-0.003	0.108	0.137	0.003	2.98
Serbia	5	11	-0.008	0.095	0.130	-0.008	3.94
Slovakia	15	10	-0.053	0.027	0.033	-0.001	2.92
Slovenia	5	15	-0.060	0.034	0.034	0.001	2.69
Spain	8	13	-0.038	0.027	0.034	-0.000	2.52
Sweden	5	17	-0.077	0.031	0.031	0.007	2.78
Ukraine	11	9	-0.040	0.076	0.109	-0.009	5.21
United Kingdom	76	17	-0.039	0.030	0.050	0.005	2.55

Note: The sample from each country consists of manufacturing firms only. All the statistics reported in the table except for the last column are medians. Employees shows the number of employees for each country. Age is computed as the difference between year and incorporation year plus one. ΔlnFTA, ΔlnSales, ΔlnVA, and ΔlnTFPR denote the year-on-year change in the log fixed tangible assets, log sales, log value added and log productivity respectively. TFPR is estimated using the Wooldridge (2009) approach. TFPR cannot be estimated for Japan due to the lack of data on material costs. Var(lnMRPK) is the cross-section variance of the log marginal revenue product of capital, where MRPK is computed as nominal revenue over fixed tangible assets. The variance reported here is unconditional on industries.

MRPK across countries. Ukraine has the highest dispersion of MRPK, followed by Serbia and Croatia, whereas France has the lowest dispersion of MRPK.

Table 3: Medians of Selected Variables for Listed and Unlisted Firms

	Employees		Age		$\Delta \log k$		ΔlogSales	
Country	Unlisted	Listed	Unlisted	Listed	Unlisted	Listed	Unlisted	Listed
Bulgaria	13	201	8	46	-0.026	-0.026	0.047	0.027
Croatia	5	261	11	56	-0.039	-0.020	0.020	0.014
Czech Republic	23	750	12	14	-0.026	-0.024	0.034	0.026
Finland	8	957	15	26	-0.067	-0.004	0.025	0.042
France	6	342	14	26	-0.103	0.009	0.024	0.049
Germany	26	752	21	44	-0.035	0.005	0.012	0.039
Italy	12	413	15	27	-0.033	-0.003	0.020	0.034
Japan	15	784	31	61	-0.025	-0.009	0.008	0.025
Korea	15	168	8	21	-0.000	0.014	0.085	0.067
Norway	9	406	12	14	-0.059	0.000	0.039	0.068
Poland	87	225	13	20	-0.019	0.022	0.047	0.074
Portugal	7	336	14	41	-0.074	-0.033	0.012	0.018
Romania	5	231	8	14	-0.003	-0.004	0.108	0.083
Serbia	4	111	10	20	-0.009	-0.000	0.098	0.067
Slovakia	15	225	10	14	-0.054	-0.034	0.027	0.008
Slovenia	5	228	15	23	-0.060	-0.011	0.034	0.028
Spain	8	558	13	40	-0.039	0.002	0.027	0.046
Sweden	5	81	17	18	-0.078	-0.010	0.030	0.068
Ukraine	10	328	9	21	-0.041	-0.009	0.075	0.111
United Kingdom	73	470	16	21	-0.040	0.008	0.029	0.045

Note: The sample from each country consists of manufacturing firms only. The table shows the median values of four variables computed using subsamples of unlisted and listed firms. Employees shows the median number of employees for each country. Age is computed as the difference between year and incorporation year plus one. ΔlnFTA and ΔlnSales denote the year-on-year change in the log fixed tangible assets and log sales respectively.

Table 3 shows the median values computed using subsamples of unlisted and listed firms. As can be seen, the median number of employees and age for listed firms are much larger than for unlisted firms. The year-on-year change in the log of fixed tangible assets ΔlnFTA is also higher for listed firms except for Bulgaria and Romania. By contrast, there is no clear pattern for the year-on-year change in the log of sales (i.e., sale growth) between the listed and unlisted firms. In 8 out of 20 countries, the sales growth of unlisted firms is higher than that of listed firms.

4 Empirical Analysis

Starting from Fazzari et al. (1988), there is a large literature on testing whether firm investment responds to cash flow fluctuations after controlling for the investment opportunity proxied by Tobin's q. By dividing firms into different groups according to some firm characteristics that can affect their constrained status, a higher investment-cash flow sensitivity within the 'constrained' group relative to the 'unconstrained' group after controlling for q would suggest the presence of financial frictions. Since the datasets contain mostly unlisted firms for which q is not available, this paper uses sales growth, value added growth and productivity growth as different proxies for the investment opportunity. Section 4.1 shows that these proxies are motivated by the model in Section 2.

Using ex ante divisions of firms into constrained and unconstrained groups based on the marginal revenue product of capital (MRPK), Section 4.1 shows some evidence that the investment of constrained firms is more sensitive to their internal financing. Section 4.2 uses a switching regression model to classify firms into constrained and unconstrained firms, where the probability of a firm being constrained depends on multiple firm characteristics and is estimated jointly with two different investment regimes depending on the firm's constrained status.

4.1 Firm Investment and Financial Frictions

Following (13), the capital demand of an unconstrained firm is determined by its productivity and the factor prices (r_t, w_{t+1} and $p_{m,t+1}$). Let $\Delta\ln k^U_{i,t} \equiv \ln k^U_{i,t} - \ln k^U_{i,t-1}$ denote the net capital investment of an unconstrained firm.[29] It can be shown that:

$$\Delta\ln k^U_{i,t} = \epsilon\rho\Delta\ln z_{i,t} - (1+\epsilon\beta_k)\Delta\ln(r_t+\delta) + (1+\epsilon\beta_k)\Delta\ln \mathrm{E}_t\left[\left(\frac{Z_{t+1}}{w_{t+1}^{\beta_l}p_{m,t+1}^{\beta_m}}\right)^{\frac{\epsilon}{1+\epsilon\beta_k}}\right] \tag{21}$$

where $\Delta\ln z_{i,t}$ is the idiosyncratic revenue-based total factor productivity growth that can proxy for the firm's investment opportunity. Note that $k_{i,t-1}$ is the capital decided in period $t-1$ but used in period t, so it is measured by contemporaneous fixed tangible assets $\mathrm{FTA}_{i,t}$ in the data. As a result, (21) implies that $\Delta\ln\mathrm{FTA}_{i,t+1}$ depends positively on productivity growth $\Delta\ln z_{i,t}$ and change in log expected future common productivity, and negatively on change in log user cost of capital $\Delta\ln(r_t+\delta)$ and change in log expected future real wage and

[29] Note that from the model, gross investment is $k_{i,t} - (1-\delta)k_{i,t-1}$. Since $\Delta\ln k_{i,t} = \ln\frac{k_{i,t}}{k_{i,t-1}} = \ln\left(\frac{k_{i,t}-k_{i,t-1}}{k_{i,t-1}}+1\right) \approx \frac{k_{i,t}-k_{i,t-1}}{k_{i,t-1}}$, $\Delta\ln k_{i,t}$ is a measure for growth rate of capital stock or net investment normalised by the capital stock $k_{i,t-1}$.

real price of materials. The last two terms in (21) are common to all firms in each subsector, so they are absorbed by (four-digit) industry-year fixed effects.

The investment of unconstrained firms can also be written in terms of the expected growth of the nominal revenue $p_{i,t}y_{i,t}$. Using the first order condition for capital, the investment of an unconstrained firm is:

$$\Delta \ln k^U_{i,t} = \Delta \ln \mathrm{E}_t \left[p_{i,t+1} y_{i,t+1} \right] + \ln \frac{r_{t-1} + \delta}{r_t + \delta} \tag{22}$$

Since $\Delta \ln \mathrm{E}_t \left[p_{i,t+1} y_{i,t+1} \right] = \Delta \ln \mathrm{E}_t \left[(1-\beta_m) p_{i,t+1} y_{i,t+1} \right]$, both the sales growth and the value added growth can be used to proxy for $\Delta \ln \mathrm{E}_t \left[p_{i,t+1} y_{i,t+1} \right]$. If the firm's productivity can be estimated accurately, then it is a more exogenous measure for investment opportunity. However, given the existing empirical methods may not perfectly back out the unobserved firm's revenue-based total factor productivity, I also use lagged sales growth and value added growth to proxy for the investment opportunity in the regression analysis.[30] Since expected future sales or value added growth is unavailable in data and current growth is likely to cause endogeneity problems, lagged growth is used instead under the assumption that lagged growth can predict future growth.

Following (12), the capital demand of a constrained firm is only determined by its net worth and hence its investment is determined by the growth in its net worth.

$$\Delta \ln k^C_{i,t} = \Delta \ln n_{i,t} \tag{23}$$

Alternatively, it can be expressed in terms of cash flow $\mathrm{CF}_{i,t}$, which is defined as the revenue net of the wage payments, material costs and net interest payments on debt, $\mathrm{CF}_{i,t} \equiv p_{i,t} y_{i,t} - w_t l_{i,t} - p_{m,t} m_{i,t} - r_{t-1} b_{i,t-1}$, assuming that all debt is rolled over in each period with no repayment of principal until the terminal period. It is shown in Appendix A.5 that:

$$\Delta \ln k^C_{i,t} \approx \frac{k^C_{i,t} - k^C_{i,t-1}}{k^C_{i,t-1}} = \frac{1}{1-\phi(1-\delta)} \frac{\mathrm{CF}_{i,t}}{k_{i,t-1}} - \frac{\delta}{1-\phi(1-\delta)} \tag{24}$$

where the firm's cash flow is the sum of its net income (equivalent to the change in net worth $\Delta n_{i,t}$ in this model setup) and depreciation of the capital stock. Since the capital $k_{i,t-1}$ decided in period $t-1$ but used in period t is measured by contemporaneous fixed tangible assets $\mathrm{FTA}_{i,t}$ in the data, (24) implies that $\Delta \ln \mathrm{FTA}_{i,t+1}$ depends positively on $\frac{\mathrm{CF}_{i,t}}{\mathrm{FTA}_{i,t}}$. Regressing $\Delta \ln \mathrm{FTA}_{i,t}$ on $\frac{\mathrm{CF}_{i,t-1}}{\mathrm{FTA}_{i,t-1}}$ would cause simultaneity bias as $\mathrm{FTA}_{i,t-1}$ appears on both

[30] There is no need to deflate the nominal sales or nominal value added because four-digit industry*year dummies are included in the regressions, which will absorb the deflators varying at the two-digit industry-year level (as deflators are often only available at the two-digit industry level in the data).

sides of the equation. Hence, I use lagged cash flow over twice lagged fixed tangible assets $\frac{\mathrm{CF}_{i,t-1}}{\mathrm{FTA}_{i,t-2}}$ in empirical analysis to prevent the simultaneity bias. This paper uses cash flow to measure the availability of internal financing because it is likely to be a more exogenous measure than net worth as capital stock forms part of the net worth. As can be seen from (24), whenever the firm is constrained by external financing, its investment is completely determined by the availability of internal financing.

According to the model, in the absence of any financial frictions, the investment in capital should not be affected by the availability of internal financing, after controlling for the investment opportunity. By contrast, when financial frictions exist so that some firms are financially constrained due to low net worth, these firms' investment is expected to be sensitive to the availability of their internal financing.

However, there are two caveats to the above implication. First, a firm's constrained status can change over time. If a firm was constrained in the previous period but becomes unconstrained this period, then the firm's investment will still depend on the net worth even after controlling for the investment opportunity.[31] More importantly, since there is no perfect empirical measure for the investment opportunity, the availability of internal financing also captures current profitability and is thus related to expected future profitability (due to persistence in $z_{i,t}$). Hence, when running a regression with the change in fixed tangible asset on investment opportunity and cash flow, it is likely to get a significant effect on cash flow even for unconstrained firms. Nevertheless, if the investment-cash flow sensitivity decreases with the factors that alleviate financial frictions, then this sensitivity is likely to be linked to financial frictions (Ağca and Mozumdar, 2008).

To test this hypothesis, I use the model to identify firms that are likely to be constrained. As explained in Section 2.3, with the financial frictions, a firm with a higher MRPK is more likely to be constrained due to a lower level of capital. An indicator variable d is created to indicate whether the firm belongs to the constrained group. I then run the following regression for each country separately:

$$\begin{aligned}\Delta\mathrm{lnFTA}_{i,s,t} =&\gamma_0 + \gamma_1\mathrm{Opp}_{i,s,t-1} + \gamma_2\mathrm{Opp}_{i,s,t-1} * d_{i,s,t} + \gamma_3\frac{\mathrm{CF}_{i,s,t-1}}{\mathrm{FTA}_{i,s,t-2}} + \gamma_4\frac{\mathrm{CF}_{i,s,t-1}}{\mathrm{FTA}_{i,s,t-2}} * d_{i,s,t} \\ &+ d_{i,s,t} + \gamma_i + \gamma_{s,t} + \varepsilon_{i,s,t}\end{aligned} \tag{25}$$

where i, s, and t denote firm, (four-digit) industry and year, respectively, and $\mathrm{Opp}_{i,s,t-1}$ denotes the investment opportunity, which is proxied by lagged sales growth, value added growth or productivity growth. The availability of internal financing $\frac{\mathrm{CF}_{i,s,t-1}}{\mathrm{FTA}_{i,s,t-2}}$ is measured

[31] Note that in this case, $\Delta\mathrm{ln}k_{i,t} \equiv \mathrm{ln}k^U_{i,t} - \mathrm{ln}k^C_{i,t}$, so the firm's investment depends on $\mathrm{ln}z_{i,t}$ and $\mathrm{ln}n_{i,t-1}$.

by lagged cash flow over twice lagged fixed tangible assets $\text{FTA}_{i,s,t-2}$, where cash flow is the sum of profit for the period and depreciation. γ_i and $\gamma_{s,t}$ represent firm and four-digit industry*year fixed effects respectively.

Firms with higher MRPK are more likely to be financially constrained and hence their investment should be more dependent on the availability of their internal financing and thus more sensitive to cash flow, as shown in (24). As a result, γ_4 is expected to be significantly positive if $d_{i,s,t}$ indicates these firms. Similarly, γ_2 is expected to be significantly negative as the investment of the 'constrained' firms should be less affected by investment opportunity than their unconstrained counterparts. However, lagged sales or value added growth are noisy proxies for investment opportunity, so the results on investment opportunity may not be reliable.

Table 4 shows the results from regressing firm investment $\Delta\text{lnFTA}_{i,t}$ on lagged sales growth $\Delta\text{lnSales}_{i,t-1}$ and lagged cash flow $\text{CF}_{i,t-1}$ over twice lagged fixed tangible assets $\text{FTA}_{i,t-2}$ for each country separately, where both explanatory variables are interacted with an indicator variable that equals one if the firm's lagged log of MRPK is in the top 30% and zero if otherwise, following the specification (25). A higher MRPK tends to indicate a more constrained status. The effect of cash flow on investment is expected to be larger for firms with higher MRPK. The effect of investment opportunity would only be smaller for these firms if there were a perfect measure for investment opportunity.

As can be seen from Table 4, in all countries except for Japan, the coefficient for cash flow interacted with the indicator variable for MRPK is highly significant and positive, indicating that the investment of firms with higher MRPK is more sensitive to their cash flow, as expected. In 9 out of 20 countries, these firms' investment also responds less to the investment opportunity, as shown by the significantly negative coefficient on lagged sales growth interacted with the indicator variable. However, in Ukraine, the coefficient for the interaction term between lagged sales growth and the indicator variable is significantly positive, which is inconsistent with the hypothesis. Besides, in three countries, the coefficient on lagged sales growth is significantly negative. These anomalous findings on lagged sales growth are likely because lagged sales growth is a noisy proxy for investment opportunity.

To interpret the coefficients in terms of their economic significance, I use France as an example. More specifically, as shown in Table 4, the coefficient of cash flow is 0.017 for France, which means for older firms, when cash flow increases by 0.1 (or 6.3% from its unconditional mean of 1.6), the capital growth increases by 0.0017 (or 8.1% from its unconditional mean of -0.021). The coefficient of the cash flow interacted with the age dummy of 0.006 means that for firms with MRPK in the top 30%, when cash flow increases by 0.1, their investment increases by 0.0023 or 11% from its unconditional mean. Similarly, the investment of older

Table 4: Capital Investment-Cash Flow Sensitivity and Marginal Revenue Product of Capital (MRPK)

Country	ΔlnSales	ΔlnSales $* d$	$\frac{\text{CF}}{\text{FTA}}$	$\frac{\text{CF}}{\text{FTA}} * d$	$d(\text{MRPK} > p70)$	Within R^2	Observations
Bulgaria	0.001	-0.014	-0.001	0.022***	0.400***	0.0668	67,519
	(0.0045)	(0.0131)	(0.0019)	(0.0024)	(0.0118)		
Croatia	0.034***	0.010	0.002	0.022***	0.414***	0.0687	82,909
	(0.0047)	(0.0144)	(0.0016)	(0.0020)	(0.0103)		
Czech Republic	0.012**	-0.050***	0.004***	0.019***	0.386***	0.0657	119,733
	(0.0048)	(0.0142)	(0.0015)	(0.0019)	(0.0081)		
Finland	0.018***	-0.018*	0.007***	0.006***	0.282***	0.0535	118,999
	(0.0041)	(0.0108)	(0.0013)	(0.0015)	(0.0060)		
France	0.074***	-0.039***	0.017***	0.006***	0.306***	0.0693	1,011,014
	(0.0030)	(0.0076)	(0.0004)	(0.0005)	(0.0021)		
Germany	0.026***	-0.008	0.001	0.008***	0.289***	0.0516	68,887
	(0.0087)	(0.0187)	(0.0014)	(0.0017)	(0.0094)		
Italy	0.035***	0.006	0.011***	0.012***	0.295***	0.0514	1,282,096
	(0.0016)	(0.0039)	(0.0006)	(0.0007)	(0.0020)		
Japan	-0.007	-0.053***	0.042***	-0.007	0.235***	0.0391	59,519
	(0.0071)	(0.0189)	(0.0099)	(0.0110)	(0.0098)		
Korea	0.018***	-0.060***	0.004***	0.028***	0.588***	0.1045	392,415
	(0.0025)	(0.0073)	(0.0009)	(0.0011)	(0.0057)		
Norway	0.037***	-0.017	0.006***	0.010***	0.363***	0.0624	81,944
	(0.0064)	(0.0200)	(0.0011)	(0.0013)	(0.0092)		
Poland	0.018***	-0.027*	0.008***	0.008***	0.284***	0.0505	87,532
	(0.0059)	(0.0150)	(0.0020)	(0.0023)	(0.0081)		
Portugal	0.039***	0.008	0.005***	0.014***	0.305***	0.0454	273,761
	(0.0030)	(0.0092)	(0.0012)	(0.0014)	(0.0044)		
Romania	-0.006***	-0.009*	0.005***	0.014***	0.396***	0.0541	367,514
	(0.0020)	(0.0049)	(0.0006)	(0.0008)	(0.0046)		
Serbia	0.023***	-0.003	0.005***	0.019***	0.384***	0.0663	105,557
	(0.0032)	(0.0091)	(0.0015)	(0.0020)	(0.0086)		
Slovakia	-0.019**	-0.035*	0.005	0.026***	0.457***	0.0614	47,523
	(0.0077)	(0.0189)	(0.0032)	(0.0041)	(0.0143)		
Slovenia	-0.118***	0.014	0.001	0.031***	0.422***	0.0513	62,121
	(0.0104)	(0.0161)	(0.0028)	(0.0036)	(0.0140)		
Spain	0.021***	0.002	0.012***	0.017***	0.276***	0.0521	1,111,449
	(0.0016)	(0.0050)	(0.0007)	(0.0008)	(0.0021)		
Sweden	0.044***	-0.001	0.010***	0.005***	0.304***	0.0488	237,819
	(0.0038)	(0.0110)	(0.0007)	(0.0009)	(0.0051)		
Ukraine	0.017***	0.009**	-0.001	0.008***	0.313***	0.0343	286,337
	(0.0017)	(0.0047)	(0.0008)	(0.0009)	(0.0053)		
United Kingdom	0.032***	-0.030***	0.003***	0.004***	0.266***	0.0432	200,281
	(0.0047)	(0.0106)	(0.0007)	(0.0008)	(0.0051)		

Note: The table shows the coefficients from regressing $\Delta\text{lnFTA}_{i,t}$ on lagged sales growth $\Delta\text{lnSales}_{i,t-1}$ and lagged cash flow over twice lagged fixed tangible assets $\frac{\text{CF}_{i,t-1}}{\text{FTA}_{i,t-2}}$, and each of which interacted with a dummy that equals one if lagged log MRPK is in the top 30% and zero if otherwise. The last column shows the number of firm-year observations used in each regression. Firm and four-digit industry*year fixed effects are included in all regressions. Firm-level clustered standard errors are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

firms increases by 0.000074 or 0.4% from its unconditional mean of -0.021 when the lagged sales growth rate increases by 0.001 (or 3.6% from its unconditional mean of 0.028). By contrast, for the same increase in the lagged sales growth, the investment of firms with a higher MRPK increases by only 0.000035 or around 0.17% from its unconditional mean. The summary statistics of the key variables can be found in Table 8 and 9 in Appendix D.

For robustness checks, I use different proxies for investment opportunity. Results using lagged value added growth are very similar to the baseline results, so they are not shown in the Appendix due to space limitations. Using lagged productivity growth gives more insignificant or anomalous findings on the coefficient for the interaction term between the lagged productivity growth and the indicator variable, but the coefficient on the interaction term between cash flow and the indicator is still highly significant and positive in all countries except for the UK, as shown in 12 in Appendix F.

Instead of using MRPK, other variables could be used to try to identify constrained firms, such as size or age.[32] However, using $d_{i,s,t}$ to indicate either small or young firms yields inconclusive results.[33] This is likely because firm size and age are imperfectly related to MRPK which is the theoretical determinant for the firm's constrained status.

4.2 Switching Regression Model

One problem with the specification (25) in Section 4.1 is that classifying firms as constrained or unconstrained based on just one variable may not be sufficient. As a result, more recent papers studying the investment-cash flow sensitivities tend to use switching regression models of Maddala (1986) and jointly estimate the probability of a firm being constrained and two different investment regimes depending on whether the firm is constrained or not (e.g., Almeida and Campello, 2007; Hovakimian and Titman, 2006; Hu and Schiantarelli, 1998). The objective of these papers is to show that financial frictions matter for firm investment, avoiding the use of ex ante classification of firms based on one firm characteristic each time.

The main objective of this paper is to estimate the percentage of the dispersion of MRPK that is caused by financial frictions or the presence of constrained firms based on the decomposition (18), which requires classifying firms as constrained or unconstrained. Consequently, this paper uses the switching regression model not only to show the importance of financial frictions in affecting firm investment, but also to classify firms by estimating the probability

[32] Beck, Demirgüç-Kunt and Maksimovic (2005) find that the smallest (largest) firms are affected the most (least) by financing obstacles, using survey data from 54 countries in the late 1990s.

[33] For instance, using an indicator variable that equals one if age (or total assets) is below its 30th or 50th percentile, or if age is below an absolute value (e.g., age $<$ 15 years), to indicate constrained firms yields inconclusive results. The results are only consistent with the hypothesis that constrained firm's investment is more sensitive to cash flow in around ten or fewer countries.

of each firm being constrained. The classification based on the switching regression is likely to be more reliable than the simple ex ante division based on one firm characteristic. Based on the firm classification, the percentage of the dispersion of MRPK that can be explained by the presence of constrained firms can be estimated.

In the switching regression model shown below, the probability of firm i being financially constrained or unconstrained is determined by the following selection equation:

$$s^*_{i,t} = \boldsymbol{x_{S,i,t}\gamma^S} + \varepsilon_{S,i,t} \tag{26}$$

where $s^*_{i,t}$ is a latent variable and $\boldsymbol{x_{S,i,t}}$ is a row vector that contains variables affecting a firm's constrained status, including $\text{Age}_{i,t}$, size measured by $\ln\text{Assets}_{i,t-1}$, $\ln\text{MRPK}_{i,t-1}$, inverse leverage measured by the net-worth-to-assets ratio $\left(\frac{\text{Shareholders' Funds}}{\text{Assets}}\right)_{i,t-1}$, and liquidity measured by cash-to-assets ratio $\left(\frac{\text{Cash}}{\text{Assets}}\right)_{i,t-1}$. In addition, vector $x_{S,i,t}$ also includes year dummies and four-digit industry dummies. The parameter $\boldsymbol{\gamma^S}$ is a column vector of the corresponding coefficients for the variables in $\boldsymbol{x_{S,i,t}}$. The error term $\varepsilon_{S,i,t}$ has a logistic distribution with mean zero and variance σ^2_S.

Constrained and unconstrained firms follow two different investment regimes, as the investment of the constrained firms should be more sensitive to fluctuations in internal financing. Which investment regime a firm follows depends on the selection equation (26). More specifically, the investment of a firm i follows the constrained (C) regime, i.e., $\Delta\ln\text{FTA}_{i,t} = \Delta\ln\text{FTA}^C_{i,t}$, if the latent variable $s^*_{i,t}$ is positive:

$$\Delta\ln\text{FTA}^C_{i,t} = \boldsymbol{x_{C,i,t}\gamma^C} + \varepsilon_{C,i,t} \quad \text{if} \quad \boldsymbol{x_{S,i,t}\gamma^S} + \varepsilon_{S,i,t} > 0 \tag{27}$$

But it follows the unconstrained (U) regime, i.e., $\Delta\ln k_{i,t} = \Delta\ln k^U_{i,t}$, if the latent variable $s^*_{i,t}$ is nonpositive:

$$\Delta\ln\text{FTA}^U_{i,t} = \boldsymbol{x_{U,i,t}\gamma^U} + \varepsilon_{U,i,t} \quad \text{if} \quad \boldsymbol{x_{S,i,t}\gamma^S} + \varepsilon_{S,i,t} \leqslant 0 \tag{28}$$

where $\boldsymbol{x_{C,i,t}}$ and $\boldsymbol{x_{U,i,t}}$ both contain the investment opportunity $\text{Opp}_{i,t-1}$, the availability of internal financing, year dummies, and four-digit industry dummies. As discussed in Section 4.1, $\text{Opp}_{i,t-1}$ is proxied by lagged sales growth, lagged value added growth, or lagged productivity growth, and the availability of internal financing is proxied by lagged cash flow over twice lagged capital stock. $\boldsymbol{\gamma^C}$ and $\boldsymbol{\gamma^U}$ are the corresponding coefficients for the variables in vectors $\boldsymbol{x_{C,i,t}}$ and $\boldsymbol{x_{U,i,t}}$ respectively. The differences between the parameters $\boldsymbol{\gamma^C}$ and $\boldsymbol{\gamma^U}$ reflect the differential investment behavior of firms in the two regimes. The error terms are normally distributed with mean zero and standard deviations of σ_C and σ_U, i.e.,

$\varepsilon_{C,i,t} \sim N(0, \sigma_C^2)$ and $\varepsilon_{U,i,t} \sim N(0, \sigma_U^2)$, where $\varepsilon_{C,i,t}$ and $\varepsilon_{U,i,t}$ are independent of $\varepsilon_{S,i,t}$. This paper uses an exogenous switching model as in Garcia, Lusardi and Ng (1997), because firms do not choose to become constrained or unconstrained. However, if shocks to firms' investment are correlated with shocks to the financial variables in the selection equation (26), then an endogenous switching regression model, where $\varepsilon_{C,i,t}$ and $\varepsilon_{U,i,t}$ are each allowed to be correlated with the error term in the selection equation $\varepsilon_{S,i,t}$, may be more appropriate, as studied in Almeida and Campello (2007), Hovakimian and Titman (2006), and Hu and Schiantarelli (1998), for example.

Although theoretically, the availability of internal financing only matters for the constrained firms, as shown in (24), it is included in both investment regimes since (21) and (24) only hold if a firm's constrained status stays the same in two consecutive periods.[34] Furthermore, cash flow (as a proxy for the availability in internal financing) captures current profitability which may capture information about the investment opportunity that is not captured by the proxies such as lagged sales growth or productivity growth, so it can be significant in both regimes. However, assuming that the proxies such as lagged sales growth or productivity growth capture the investment opportunity equally well for the two types of firms, the presence of financial frictions would be expected to lead to a larger coefficient on cash flow and a lower coefficient on sales growth in the constrained regime.

The switching regression itself does not automatically identify which investment regime is associated with constrained firms. The identification of the constrained investment regime requires theoretical priors on how certain firm characteristics indicate firms' constrained status. Since both $\boldsymbol{x_{C,i,t}}$ and $\boldsymbol{x_{U,i,t}}$ contain the same variables, the constrained regime is identified using some of the variables included in $\boldsymbol{x_{S,i,t}}$. In this paper, the identification relies on the signs and the significance of the coefficients for age, size and MRPK, since it is relatively unambiguous that firms with younger age, smaller size and higher MRPK are more likely to be constrained. More specifically, if the coefficients for these three variables are each significant at 10% level at least, then the regime is classified as constrained if the probability for being in this regime increases in MRPK and decreases in age and size at the same time. If not all three variables are each significant at 10% level at least, then the regime classification relies only on the signs of the significant variables.

Although firm leverage and liquidity ratios are also included in the selection equation, they are not used for the identification of the constrained regime because a priori, it is ambiguous how firm leverage and liquidity indicate a firm's constrained status. One the one hand, firms with higher leverage could be more constrained as they have lower net worth

[34] For example, if a firm is constrained in period $t-1$ but becomes unconstrained in period t, then its investment in period t still depends on its lagged net worth.

and thus lower borrowing capacity. On the other hand, they may be unconstrained because the fact that they have higher leverage could mean they are able to borrow a lot in the first place. Similarly, firms with high liquidity measured by cash over total assets can be unconstrained if it indicates the firms are profitable and mature. However, it can indicate that firms are constrained if firms cannot easily borrow from the credit market and thus hold more cash as precautionary savings.

Although the financially constrained status of a firm is unobservable, according to (26) and (27), the probability of a firm i being constrained in period t can be specified as:

$$P(\boldsymbol{x_{S,i,t}\gamma^S} + \varepsilon_{S,i,t} > 0) = P(\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^S}) \tag{29}$$

Assuming the error term $\varepsilon_{S,i,t}$ has a logistic distribution with mean zero and standard deviation of σ_S, i.e., $\varepsilon_{S,i,t} \sim Logit(0, \sigma_S^2)$, then the probability of firm i being constrained in period t is determined by a logit function:[35]

$$P(\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^S}) = P(\varepsilon_{S,i,t} < \boldsymbol{x_{S,i,t}\gamma^S}) = \frac{\exp(\boldsymbol{x_{S,i,t}\gamma^S})}{1+\exp(\boldsymbol{x_{S,i,t}\gamma^S})} \tag{30}$$

The likelihood function of an observation $L_{i,t}$ is the weighted sum of the likelihoods of being in each latent class (i.e., the constrained and unconstrained groups of firms), where the weights are the associated latent class probabilities, $P(\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^S})$ and $P(\varepsilon_{S,i,t} \leqslant -\boldsymbol{x_{S,i,t}\gamma^S})$. It is shown in Appendix E that:

$$L_{i,t} = f(\varepsilon_{C,i,t})P(\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^S}) + f(\varepsilon_{U,i,t})P(\varepsilon_{S,i,t} \leqslant -\boldsymbol{x_{S,i,t}\gamma^S}) \tag{31}$$

where $f(.)$ is the marginal normal density. It follows from (31) that the log-likelihood function for all observations is:

$$L = \sum_{i=1}^{M}\sum_{t=1}^{T_i} \ln(L_{i,t}) \tag{32}$$

where M is the number of firms in each industry and T_i is the number of observations for each firm i. By maximizing the log-likelihood function (32), the parameters γ^S, γ^C, γ^U, $\ln\sigma_C$, and $\ln\sigma_U$ can be estimated. With the estimated parameters, it is possible to calculate the posterior probability of each firm being in each of the two regimes. Once the regimes

[35]Similarly, according to (26) and (28), the probability of firm i being unconstrained is:

$$P(\boldsymbol{x_{S,i,t}\gamma^S} + \varepsilon_{S,i,t} \leqslant 0) = P(\varepsilon_{S,i,t} \leqslant -\boldsymbol{x_{S,i,t}\gamma^S}) = 1 - P(\varepsilon_{S,i,t} \leqslant \boldsymbol{x_{S,i,t}\gamma^S}) \quad = \frac{1}{1+\exp(\boldsymbol{x_{S,i,t}\gamma^S})}$$

where the last step uses (30).

are identified as constrained or unconstrained, the posterior probability of a firm being in the constrained regime can be used to classify firms as constrained or unconstrained for each period t. In this paper, if the posterior probability of a firm being constrained is greater than 0.5 in any period, then this firm is classified as constrained in that period. The posterior probability of a firm being constrained $P(\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^{S}}|\Delta\text{lnFTA}_{i,t})$ takes into account the information about investment by updating the prior probability based on Bayes' rule:

$$\frac{f(\varepsilon_{C,i,t}|\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^{S}})P(\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^{S}})}{f(\varepsilon_{C,i,t}|\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^{S}})P(\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^{S}}) + f(\varepsilon_{U,i,t}|\varepsilon_{S,i,t} \leqslant -\boldsymbol{x_{S,i,t}\gamma^{S}})P(\varepsilon_{S,i,t} \leqslant -\boldsymbol{x_{S,i,t}\gamma^{S}})} \tag{33}$$

where $P(\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^{S}})$ is the prior probability for being constrained.

To control for unobserved firm heterogeneity, it would be desirable to add firm fixed effects, but there are two difficulties with doing so in a switching regression. Although adding firm dummies in all three equations (26)-(28) would control for the firm fixed effects, this is computationally very costly. For instance, if there are 1000 firms, then adding 1000 dummies into each of the three equations would result in 3000 additional parameters to be estimated. Demeaning the two investment equations would impose a rather strong assumption on the nature of the firm fixed effects, i.e., firm heterogeneity in the two different regimes has the same impact on firm investment.

The second difficulty is that even if it is computationally feasible to include thousands of firm dummies, the 'incidental parameters problem' (Neyman and Scott, 1948) still exists in the nonlinear selection equation (26). This is because there are only T_i observations to estimate each firm i's dummy and the estimate of the firm dummy remains random even as the number of firms N grows. This randomness cannot be averaged out due to the nonlinearity, unlike in a linear model. Hence, in a nonlinear model with firm fixed effects and a fixed time dimension, the maximum likelihood estimators of the firm dummies and the explanatory variables are inconsistent (Greene, 2004; Chamberlain, 1980). Within transformation or first-differencing will not eliminate the individual firm heterogeneity in a nonlinear model either.[36]

Hu and Schiantarelli (1998) deal with the firm-fixed effects by modelling them as a linear function of the means of the firm-specific variables in each investment equation and the selection equation. They control for the means of these variables in each equation. Hovakimian and Titman (2006) adopt a different approach to partially control for firm fixed

[36] In a few cases, including the logistic regression, the incidental parameters problem can be solved by conditioning on a sufficient statistic for the incidental parameters. For instance, the sufficient statistic is $\sum_{t=1}^{T_i} s_{i,t}$ for a logit model, where $s_{i,t}$ is the dependent variable that takes a value of zero or one. However, this conditional maximum likelihood approach cannot be used here, because whether a firm is constrained or not is unobserved (i.e., the value of $s_{i,t}$ is unknown).

effects. They include the firm-specific variables and their lags in each equation and also the lagged dependent variable in each investment equation before estimating the switching regression. However, first differencing will not eliminate the unobserved firm heterogeneity in a nonlinear model. So this paper follows the approach used in Hu and Schiantarelli (1998) to control for firm fixed effects.

To further reduce the problem of unobserved firm heterogeneity, this paper applies the switching regression model (26)-(28) to each two-digit industry in each country. The existing literature that adopts the switching regression often uses US data only and runs the switching regression on a country level after controlling for two-digit industry fixed effects (e.g., Hu and Schiantarelli, 1998). Since this paper uses much larger datasets, where countries have a large number of firms even at a two-digit industry level, it is possible to run the switching regression at a more disaggregated level and then control for four-digit industry fixed effects in order to mitigate the problem of unobserved firm heterogeneity.[37] Restricting the sample to a two-digit industry also improves the reliability of the proxies used for MRPK (i.e., nominal revenue or nominal value added over fixed tangible assets, treating the revenue elasticity as a constant within a subsector), as it overcomes the problem that the revenue elasticities β_k and β_m are likely to differ significantly across two-digit industries.

Using lagged sales growth to proxy for investment opportunity and lagged cash flow over twice lagged fixed tangible assets $\frac{\mathrm{CF}_{i,t-1}}{\mathrm{FTA}_{i,t-2}}$ to proxy for the availability of internal financing, Table 5 and 6 show the results from fitting the switching regression model (26)-(28) to the fabricated metal products industry (two-digit NACE Rev.2 code = 25) for each country separately. This industry has the largest number of observations.[38] Four-digit sector dummies and year dummies are included in each investment equation and the selection equation. To address the firm fixed effects, mean lagged sales growth and mean cash flow for each firm over time are included in the investment equations. The means of firm-specific variables (apart from age due to collinearity) for each firm over time in the selection equation are added as additional variables in the selection equation.

The switching regression itself does not automatically identify which investment regime is associated with constrained firms. The identification of the constrained regime relies on the

[37] As can be seen in Table 1, there are more than 200 four-digit industries in nearly all of the countries in the baseline sample. If the switching regression model were run on the country level, then it would be infeasible to control for four-digit industry fixed effects for the reasons discussed above.

[38] When summing or taking the mean of the number of observations across countries for a given industry, industry 25 has the highest data availability. On average, industry 25 accounts for around 17% of the total number of firm-year observations in the manufacturing sector in each country, while this percentage for the other industries is below 11%. The full category of the two-digit manufacturing industries and their descriptions can be found in Table 11 in Appendix D. Industries such as the manufacture of tobacco products, coke and refined petroleum products, and basic pharmaceutical products are quite concentrated in the sample of 20 countries, so they do not have enough observations for the switching regression.

theoretical priors that firms with higher MRPK, smaller size and younger age are more likely to be constrained. Although more variables are included in the selection equation, ex ante, variables such as leverage and liquidity ratios are ambiguous indicators for the constrained status, as discussed before in this section. Let P^C denote the probability of a firm being constrained, then rearranging (30) gives the log odds ratio $\ln\frac{P^C}{1-P^C} = \boldsymbol{x_{S,i,t}\gamma^S}$. As can be seen from this odds ratio, if the coefficient is positive, it means that when its corresponding variable increases, the probability of being constrained P^C also increases. Hence, if the probability of being in regime C increases in MRPK, but decreases in size and age, then regime C is the constrained regime.

Table 5 shows the coefficients for the key variables in the selection equation that determine whether a given regime is for the constrained firms (29). As can be seen, lagged log of MRPK is positive and significant in all countries, meaning that a higher MRPK increases the probability of a firm being in the constrained regime. Similarly, coefficients for age and size (proxied by log of total assets) are negative and highly significant in all countries except for one, meaning that an older age and a larger size will reduce (raise) the probability of a firm being constrained (unconstrained). This result is in line with the findings in Beck, Demirgüç-Kunt and Maksimovic (2005) that the smallest (largest) firms are affected the most (least) by financing obstacles. The coefficients for inverse leverage ratio measured by the shareholders' funds (net worth) to assets ratio have mixed signs across countries, implying that leverage ratio is an ambiguous indicator for a firm's constrained status, as expected. Higher liquidity proxied by cash-to-assets ratio increases the probability of being constrained, reflecting that firms hold more cash for precautionary reasons due to the lack of easy access to external financing.

Table 5 also shows the average proportion of constrained firms over the sample period in the last column, where firms are classified as constrained based on the estimated posterior probability. Using the posterior probability as shown in (33), a firm is classified as constrained if the posterior probability of the firm being in the constrained regime is greater than 0.5 and otherwise, the firm is classified as unconstrained. The probabilities of being constrained and unconstrained sum to one for each firm. This average proportion of constrained firms across time for each country can also be seen in graph (a) in Figure 3.

In Table 6, the coefficients for the lagged sales growth and the cash flow from two different investment regimes, the constrained regime (27) and the unconstrained regime (28), are reported. As can be seen, the coefficient for cash flow is significant and much larger for constrained firms in all countries, whereas it is not significant for unconstrained firms in 12 out of 20 countries. As discussed before, since cash flow captures the current profitability and is positively related to expected future profitability due to persistence of profitability, it can

Table 5: The Selection Equation of the Switching Regression in Fabricated Metal Products Industry

Country	Age	ln(Assets)	ln(MRPK)	$\frac{\text{Net worth}}{\text{Assets}}$	$\frac{\text{Cash}}{\text{Assets}}$	Fraction constrained
Bulgaria	-0.024***	-0.846***	0.656***	0.581	1.949***	0.39
	(0.0063)	(0.2097)	(0.1116)	(0.3773)	(0.7238)	
Croatia	-0.021***	-0.871***	1.027***	-0.117	0.981**	0.41
	(0.0049)	(0.0901)	(0.0576)	(0.2039)	(0.4581)	
Czech Republic	-0.080***	-0.842***	0.921***	0.268*	1.912***	0.41
	(0.0047)	(0.0770)	(0.0479)	(0.1601)	(0.2661)	
Finland	-0.019***	-0.959***	1.375***	0.692***	1.155***	0.23
	(0.0022)	(0.0626)	(0.0582)	(0.1434)	(0.2074)	
France	-0.014***	-1.094***	2.089***	1.043***	1.884***	0.24
	(0.0009)	(0.0412)	(0.0297)	(0.0936)	(0.1150)	
Germany	-0.006***	-0.942***	1.585***	1.619***	1.466**	0.26
	(0.0013)	(0.2625)	(0.1432)	(0.4629)	(0.6148)	
Italy	-0.013***	-0.845***	1.084***	-0.354***	1.179***	0.34
	(0.0007)	(0.0254)	(0.0156)	(0.0768)	(0.1046)	
Japan	-0.016***	-0.579	1.388***	-0.179	1.607	0.25
	(0.0032)	(0.3534)	(0.2184)	(0.7774)	(1.0320)	
Korea	-0.026***	-0.729***	1.163***	0.571***	1.389***	0.36
	(0.0024)	(0.0465)	(0.0370)	(0.1379)	(0.2789)	
Norway	-0.017***	-0.608***	1.016***	0.093	0.626**	0.34
	(0.0050)	(0.0945)	(0.0572)	(0.2121)	(0.2795)	
Poland	-0.031***	-1.074***	1.129***	0.310	2.135***	0.27
	(0.0062)	(0.1130)	(0.0856)	(0.2258)	(0.4190)	
Portugal	-0.031***	-0.931***	1.315***	0.506***	1.329***	0.37
	(0.0019)	(0.0626)	(0.0400)	(0.1223)	(0.2196)	
Romania	-0.020***	-0.507***	0.820***	-0.114**	1.370***	0.44
	(0.0036)	(0.0322)	(0.0299)	(0.0542)	(0.1720)	
Serbia	-0.005	-0.517***	0.833***	-0.507***	2.362***	0.31
	(0.0040)	(0.0625)	(0.0512)	(0.1870)	(0.5014)	
Slovakia	-0.066***	-1.290***	1.092***	0.394**	1.780***	0.40
	(0.0076)	(0.1388)	(0.0851)	(0.1859)	(0.4017)	
Slovenia	-0.039***	-1.565***	1.350***	0.989***	3.177***	0.32
	(0.0055)	(0.1349)	(0.0949)	(0.3471)	(0.5918)	
Spain	-0.016***	-0.667***	1.289***	0.080	0.984***	0.30
	(0.0010)	(0.0241)	(0.0167)	(0.0544)	(0.0921)	
Sweden	-0.013***	-0.947***	1.331***	0.780***	1.103***	0.32
	(0.0013)	(0.0488)	(0.0353)	(0.1310)	(0.1456)	
Ukraine	-0.017***	-0.448***	0.605***	-0.103	0.575**	0.33
	(0.0033)	(0.0419)	(0.0325)	(0.0731)	(0.2805)	
United Kingdom	-0.005***	-0.650***	1.029***	0.229	1.008***	0.24
	(0.0012)	(0.0799)	(0.0617)	(0.1547)	(0.2658)	

Note: The table shows the coefficients for the key variables in the selection equation that determines the probability of a firm being constrained, including age, log of assets, log of MRPK, net worth-to assets ratio, and cash-to-assets ratio, and the average proportion of constrained firms over the sample period. All variables apart from age are lagged. Four-digit industry and year fixed effects are included. To address firm fixed effects, the means of firm-specific variables (apart from age) are controlled in the selection equation. The last column shows the average proportion of constrained firms over the sample period, where firms are classified as constrained based on the estimated posterior probabilities. Robust standard errors are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Switching Regression Model of Firm Investment in Fabricated Metal Products Industry

	Unconstrained Regime		Constrained Regime				
Country	$\Delta\text{lnSales}_{i,t-1}$	$\frac{\text{CF}_{i,t-1}}{k_{i,t-2}}$	$\Delta\text{lnSales}_{i,t-1}$	$\frac{\text{CF}_{i,t-1}}{k_{i,t-2}}$	Observations	Prob > Chi2	df
Bulgaria	0.010	0.007*	0.050	0.025***	4,243	0.0000	76
	(0.0060)	(0.0042)	(0.0410)	(0.0066)			
Croatia	0.014**	-0.002	0.156***	0.025***	12,652	0.0000	76
	(0.0057)	(0.0015)	(0.0277)	(0.0038)			
Czech Republic	0.021***	-0.002	0.023	0.031***	25,421	0.0000	96
	(0.0043)	(0.0015)	(0.0249)	(0.0028)			
Finland	0.024***	-0.000	0.007	0.033***	27,429	0.0000	80
	(0.0039)	(0.0011)	(0.0257)	(0.0047)			
France	0.111***	0.010***	0.124***	0.039***	170,850	0.0000	82
	(0.0048)	(0.0006)	(0.0217)	(0.0015)			
Germany	0.067***	0.003***	0.073	0.018***	12,100	0.0000	98
	(0.0099)	(0.0011)	(0.0501)	(0.0041)			
Italy	0.020***	0.002***	0.123***	0.045***	246,989	0.0000	94
	(0.0015)	(0.0007)	(0.0083)	(0.0017)			
Japan	0.013**	0.016	0.024	0.061*	6,830	0.0000	86
	(0.0063)	(0.0103)	(0.0546)	(0.0332)			
Korea	0.009***	0.000	0.064***	0.059***	55,900	0.0000	60
	(0.0026)	(0.0011)	(0.0194)	(0.0039)			
Norway	0.028***	0.001	0.136***	0.028***	12,676	0.0000	78
	(0.0078)	(0.0007)	(0.0434)	(0.0028)			
Poland	0.036***	0.004***	0.023	0.022***	13,237	0.0000	82
	(0.0061)	(0.0016)	(0.0416)	(0.0038)			
Portugal	0.025***	-0.001	0.151***	0.022***	47,373	0.0000	76
	(0.0037)	(0.0017)	(0.0191)	(0.0023)			
Romania	0.015***	-0.001*	0.046***	0.021***	44,863	0.0000	82
	(0.0023)	(0.0005)	(0.0106)	(0.0014)			
Serbia	0.021***	0.003***	0.099***	0.047***	12,866	0.0000	74
	(0.0035)	(0.0012)	(0.0235)	(0.0056)			
Slovakia	0.030***	-0.004	-0.041	0.040***	10,806	0.0000	82
	(0.0069)	(0.0032)	(0.0328)	(0.0064)			
Slovenia	0.020***	0.002	0.075*	0.047***	12,476	0.0000	66
	(0.0071)	(0.0025)	(0.0419)	(0.0063)			
Spain	0.023***	0.001	0.131***	0.042***	193,141	0.0000	84
	(0.0016)	(0.0005)	(0.0124)	(0.0017)			
Sweden	0.054***	0.003***	0.110***	0.031***	56,662	0.0000	78
	(0.0040)	(0.0007)	(0.0206)	(0.0020)			
Ukraine	0.011***	0.002	0.100***	0.009***	20,782	0.0000	70
	(0.0023)	(0.0011)	(0.0148)	(0.0023)			
United Kingdom	0.038***	0.001	0.071**	0.013***	26,117	0.0000	82
	(0.0059)	(0.0008)	(0.0326)	(0.0020)			

Note: The dependent variable is firm investment $\Delta\text{lnFTA}_{i,t}$. The coefficients for lagged sales growth and lagged cash flow in two different investment regimes are reported. Four-digit industry and year fixed effects are included in the switching regression. Firm fixed effects are partially controlled by adding the means of the firm-specific variables in each equation, whose coefficients are not reported here. The last two columns show the p-value for the likelihood ratio test and the degrees of freedom for the χ^2 distribution respectively. A small p-value suggests that the switching regression (less restrictive model) fits the data significantly better than an OLS regression. Robust standard errors are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

be significant even for unconstrained firms. This explains why in 6 out of 20 countries, the coefficient on cash flow is highly significant for unconstrained firms. However, the coefficient on cash flow is much larger for constrained firms in all countries, so the results are consistent with the hypothesis that constrained firms' investment should be more sensitive to cash flow. In Romania, the coefficient on cash flow for unconstrained firms is negative and significant at 10% level, which is likely due to the unobserved firm heterogeneity not being fully controlled. The results without the firm fixed effects can be found in Table 17 and 18 in Appendix F, which show that the coefficient on cash flow for unconstrained firms can be negative without firm fixed effects.

The coefficient for lagged sales growth is significantly positive for unconstrained firms in all countries except for Bulgaria, whereas it is not significant for constrained firms in 7 out of 20 countries. However, in 12 countries, the coefficient on lagged sales growth for constrained firms is highly significant and larger than that for unconstrained firms, which differs from the theoretical prediction. This is likely because lagged sales growth is a poor proxy for investment opportunity, so the results on lagged sales growth are not very reliable. If there were a perfect measure for investment opportunity, then testing whether unconstrained firms' investment is more responsive to investment opportunity would be more meaningful.

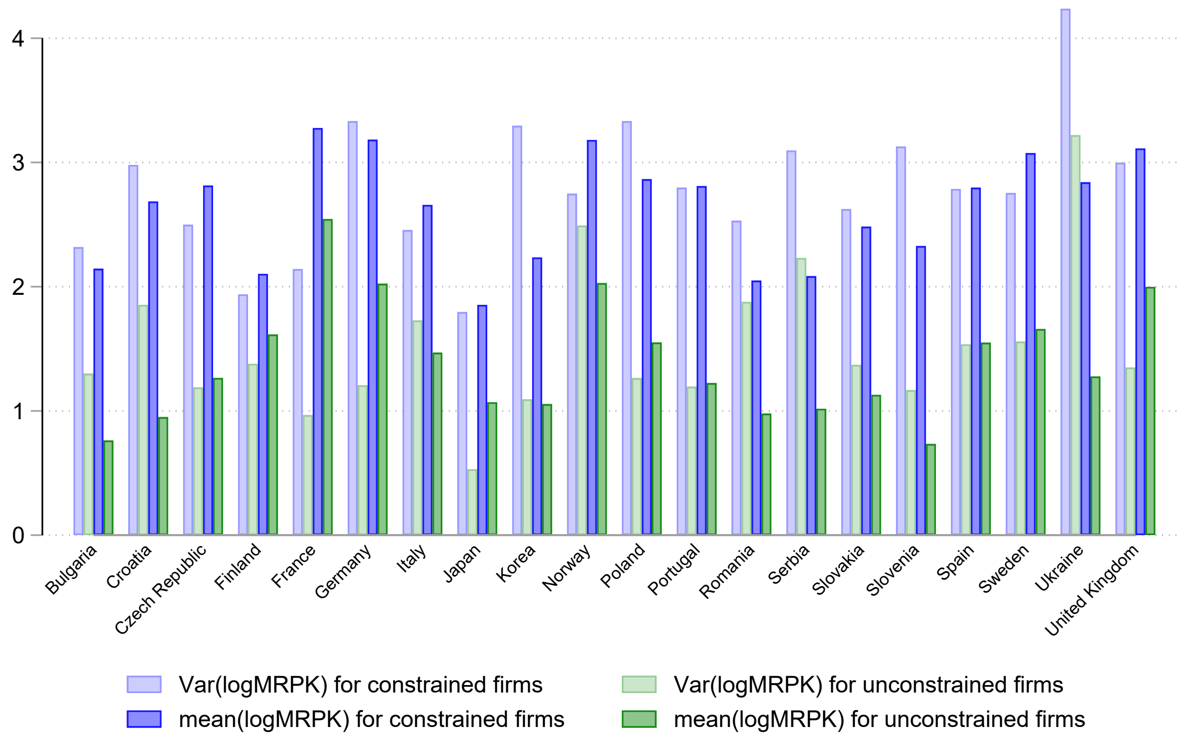
I use a likelihood ratio test to test whether the switching regression model (less restrictive) fits the data better than a single regime model estimated by OLS (more restrictive). Following Hu and Schiantarelli (1998), the degrees of freedom equal the number of constraints (that the coefficients in the two investment equations are equal) plus the number of parameters in the selection equation, which are shown in the last column of Table 6. The right-tail p-values of the chi-squared statistic are also reported in Table 6.[39] The small p-values reject the null and suggest that the switching regression model fits the data significantly better than the single regime model.

Figure 2 plots the cross-section variances (or dispersions) and means of ln(MRPK) for each type of firms in the fabricated metal products industry in each country, where the dispersions and means of ln(MRPK) are averaged over the sample period for each country. It can be seen that the dispersions and means of MRPK for constrained firms are much larger than those for unconstrained firms. These patterns are similar to those shown in Figure 1, although the contrast between the two types of firms is much more notable here. It is discussed in Section 2.3 that according to the model, constrained firms are expected to have a higher cross-section variance and mean of MRPK than unconstrained firms.

Figure 3 shows the proportion of constrained firms and credit distortion in the fabricated

[39]The chi-squared statistic equals 2*(log likelihood of less restrictive model - log likelihood of more restrictive model).

Figure 2: Dispersions and Means of Marginal Revenue Product of Capital in Fabricated Metal Products Industry



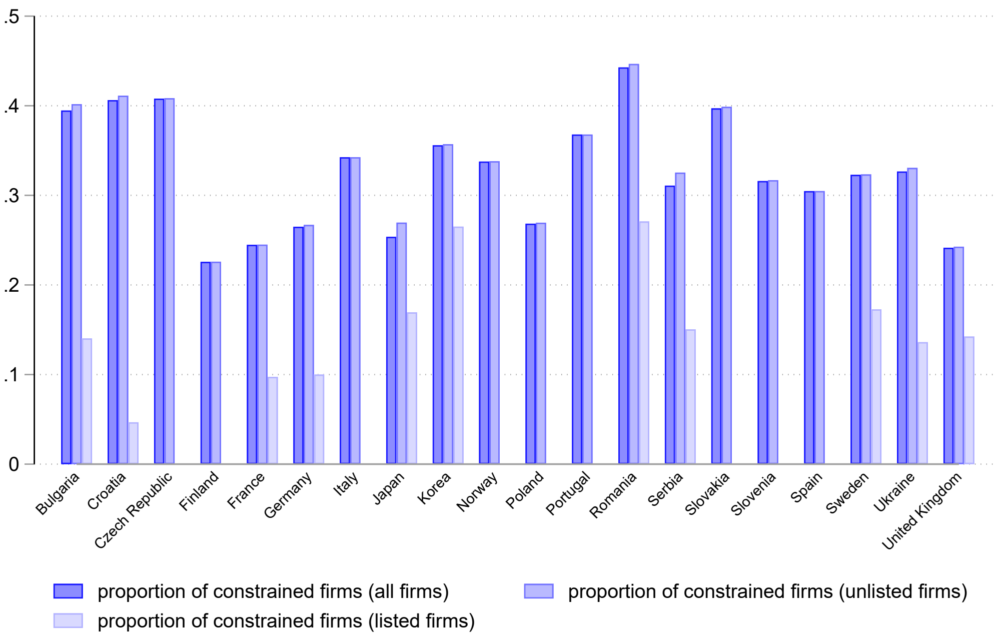
Note: The bar chart shows the cross-section variances (or dispersions) and means of the ln(MRPK) for constrained firms and unconstrained firms in industry 25 (manufacture of fabricated metal products) by NACE Rev.2 Code across 20 countries. The dispersions and means of MRPK are averaged over the sample period. Constrained and unconstrained firms are identified using the results from the switching regression model. MRPK is computed as the nominal revenue divided by tangible fixed assets.
Data source: Orbis

metal products industry (industry 25 by NACE Rev.2 Code), which are averaged over the sample period across all firms in a given sample (i.e., entire sample, listed firms only, or unlisted firms only). Graph (a) shows that in most countries, the average proportion of constrained firms across all firms and years is above 0.25 and this proportion is much larger than the average proportion computed using the subsample of listed firms over the sample period. For example, in the UK, there are around 25% of firms classified as financially constrained across all firms and years, but less than 15% of listed firms are classified as constrained on average across all listed firms and years. This is consistent with the expectation that large listed firms are less likely to be financially constrained.

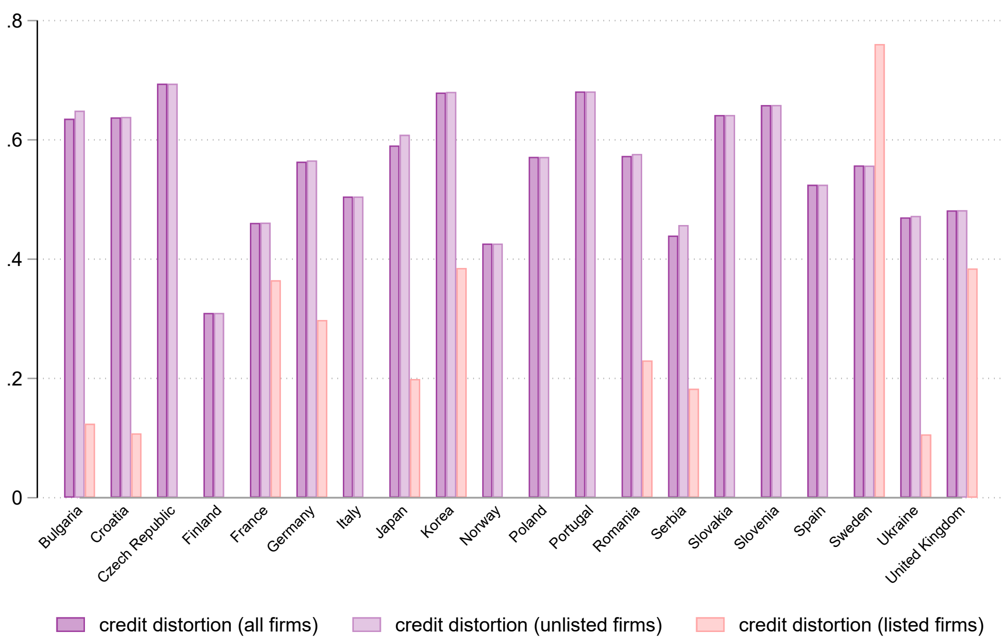
The missing bars for the listed firms in some countries in Figure 3 are because the number of observations for listed firms is below 100 over all years, in which case the sample

Figure 3: Proportion of Constrained Firms and Credit Distortion in Fabricated Metal Products Industry

(a) Proportion of Constrained Firms



(b) Credit Distortion



Note: In each graph, the corresponding measure is computed across all firms in a given sample and all years. Three different samples of firms are used: all firms, unlisted firms and listed firms. The missing bars for the listed firms are due to the number of observations being below 100 over the sample period. Graph (a) plots the fraction of constrained firms in industry 25 (manufacture of fabricated metal products) by NACE Rev.2 Code across 20 countries. Graph (b) plots the credit distortion in percent points (i.e., the proportion of the observed dispersion (cross-section variance) of MRPK that is caused by the presence of constrained firms) in industry 25, which is computed based on (19). MRPK is computed as the nominal revenue divided by tangible fixed assets.
Data source: Orbis

of listed firms is not representative and the two measures are not computed. Fewer than 100 observations for the listed firms over the sample period implies that there are around ten listed firms per year on average, given that most countries have more than ten years of data, as can be seen from Table 1. Furthermore, graph (a) in Figure 3 shows that the proportion of constrained firms is slightly above 0.1 in many countries, implying that only one out of the ten listed firms is classified as constrained, so this sample is not representative.

Classifying firms into the constrained and unconstrained groups is only the first step. The main objective of the paper is to estimate the impact of the financial constraints on the dispersion of MRPK. Based on (19), the fraction of the dispersion of MRPK that is caused by the presence of constrained firms can be computed. Graph (b) in Figure 3 plots the average credit distortion (in percent points) in the fabricated metal products industry, which are averaged over the sample period across all firms in a given sample. As shown in graph (b), the credit distortion ranges from 0.3 in Finland to around 0.7 in Czech Republic, Korea and Portugal, which means the presence of constrained firms in this industry can explain around 30-70% of the dispersion of MRPK, depending on different countries. Note that the average credit distortion computed using the subsample of listed firms tends to be lower except for Sweden.

I apply the switching regression model in the baseline analysis to all the other industries that have enough observations to run the switching regression.[40] The results for 14 different industries (including industry 25 in the baseline analysis) are summarized in Table 13 and 14 in Appendix F. Each column in Table 13 summarizes the average proportion of constrained firms across all firms and years in a given industry for different countries. The column for industry 25 in Table 13 corresponds to the height of the first bar in graph (a) in Figure 3. In addition, the last two rows of Table 13 show the mean difference between the proportion of constrained firms in a sample of unlisted firms and that in a sample of listed firms, and the number of countries used to calculate this mean difference, respectively. For example, this mean difference across 11 countries is 0.18 for industry 25 (fabricated metal products industry), as shown in Table 13.

Similarly, each column in Table 14 summarizes the credit distortion across all firms and years in a given industry for different countries. The column for industry 25 in Table 14 corresponds to the height of the first bar in graph (b) in Figure 3. The last two rows of Table 14 show the mean difference between the credit distortion in a sample of unlisted firms

[40] There are nine two-digit industries that do not have enough observations for most countries (i.e., industry 11, 12, 15, 17, 19, 21, 24, 29, and 30 by NACE Rev.2 code). Industry 33 (repair and installation of machinery and equipment) is neglected since it is not a typical manufacturing industry. Industry 33 only accounts for around 4.5% of the total number of firm-year observations in the manufacturing sector in each country on average.

and the credit distortion in a sample of listed firms, and the number of countries used to compute this mean difference, respectively.

As can be seen from Table 13 and 14, a general pattern is that for most two-digit industries and countries, there is at least a quarter of firms being classified as financially constrained and the presence of these firms explain more than half of the dispersion of MRPK across all firms. In addition, for each reported industry, the proportion of constrained firms within unlisted firms is larger than that within listed firms by around 0.15 except for industry 32 and the credit distortion within unlisted firms is higher by around 0.2 except for industries 26 and 32 on average across countries.

For robustness checks, I also use nominal value added to compute MRPK and the results for the proportion of constrained firms and credit distortion are summarized in Table 15 and 16 in Appendix F, which are very similar to the baseline results. In addition, I compare the baseline results with the case without trying to use the Hu and Schiantarelli (1998) approach to control for firm fixed effects, so with only four-digit industry and year dummies included. The coefficients for the investment equation and the selection equation can be found in Table 17 and 18 in Appendix F. The negative coefficients on cash flow for unconstrained firms are likely caused by the between variation in cash flow when firm fixed effects are ignored. Despite the differences in the coefficients, the results for the proportion of constrained firms and credit distortion are robust to whether firm fixed effects are included, as shown in Figure 4. Finally, I use different proxies for investment opportunity, i.e., lagged value added growth and productivity growth. The results for the proportion of constrained firms and credit distortion are robust to whether lagged sales growth, value added growth or productivity growth is used, except for Korea, as shown in Figure 5.

5 Conclusions

This paper proposes a novel method for estimating the impact of financial frictions on capital misallocation measured by the dispersion of the marginal revenue product of capital (MRPK), which uses large firm-level datasets and requires fewer restrictive assumptions. The key idea is that the observed dispersion of MRPK can be viewed in terms of the dispersions and means for the financially constrained and unconstrained firms. Based on the decomposition of the dispersion of MRPK, this paper provides a credit distortion measure, which measures the proportion of the observed dispersion of MRPK that can be attributed to the presence of constrained firms.

A simple model of firm dynamics with a one-period time to build for capital and borrowing constraints shows that the capital decisions and thus the MRPK for constrained and unconstrained firms are driven by two different processes. While the capital investment of an unconstrained firm is driven by the expected future investment opportunity, that of a constrained firm is determined by the availability of its internal financing. As a result, the distribution of MRPK across all firms is a mixture of two distributions, one for the constrained firms and the other for the unconstrained firms.

By decomposing the dispersion of MRPK across all firms into the dispersions and means of MRPK for the two types of firms, this paper provides new insights into the mechanisms through which the borrowing constraint increases the dispersion of MRPK. While the usual mechanism is through a higher MRPK of constrained firms relative to the unconstrained firms, an often neglected mechanism is through the dispersions within the constrained and unconstrained firms.

More importantly, this statistical decomposition provides a way to estimate the impact of financial frictions on the dispersion of MRPK once the constrained firms are identified. Using a switching regression model to identify the constrained firms in the manufacturing industry for 20 countries from the 1990s to 2015, this paper finds that the dispersions and means of MRPK within the group of constrained firms are much larger than those within the group of unconstrained firms, which are consistent with the model predictions. Furthermore, for most two-digit industries and countries, more than a quarter of firms are classified as financially constrained and the presence of these constrained firms accounts for more than half of capital misallocation.

Therefore, this paper has quantified the impact of financially constrained firms on the allocation of physical capital.

Appendices

A Solving Firm's Problem

Let $\lambda_{i,t}$ denote the Lagrange multiplier associated with the borrowing constraint (10) of firm i in period t. Using (7), (8), (9) and (10), form the Lagrangian:

$$\begin{aligned}\mathcal{L} =& \mathrm{E}_t \sum_{\tau=0}^{\infty} \varphi(1-\varphi)^{\tau} \eta^{\tau} \{p_{i,t+1+\tau} y_{i,t+1+\tau} - w_{t+1+\tau} l_{i,t+1+\tau} - p_{m,t+1+\tau} m_{i,t+1+\tau} \\ & - R_{t+\tau}(k_{i,t+\tau} - n_{i,t+\tau}) + (1-\delta) k_{i,t+\tau} + \lambda_{i,t+\tau} \left[\frac{n_{i,t+\tau}}{1-\phi(1-\delta)} - k_{i,t+\tau} \right] \right\}\end{aligned} \tag{34}$$

Using (4) and (34), the first-order conditions with respect to the firm's capital demand $k_{i,t}$, labor demand $l_{i,t}$ and materials demand $m_{i,t}$ are respectively:

$$\mathrm{E}_t \left[\beta_k \frac{p_{i,t+1} y_{i,t+1}}{k_{i,t}} - R_t + (1-\delta) \right] = \lambda_{i,t} \tag{35}$$

$$\beta_l \frac{p_{i,t} y_{i,t}}{l_{i,t}} = w_t \tag{36}$$

$$\beta_m \frac{p_{i,t} y_{i,t}}{m_{i,t}} = p_{m,t} \tag{37}$$

The nominal value added $\mathrm{VA}_{i,t}$ is equivalent to a fraction of the nominal revenue $p_{i,t} y_{i,t}$, using (4) and (37):

$$\mathrm{VA}_{i,t} \equiv p_{i,t} y_{i,t} - p_{m,t} m_{i,t} = (1-\beta_m) p_{i,t} y_{i,t} \tag{38}$$

A.1 Marginal Revenue Product of Capital

Divide (36) by (37) to get the materials-to-labor ratio:

$$\frac{m_{i,t}}{l_{i,t}} = \frac{\beta_m w_t}{\beta_l p_{m,t}} \tag{39}$$

Use (4) and the first order condition with respect to labor (36) to write the optimal labor demand $l_{i,t+1}$ in terms of w_{t+1}, $Z_{i,t+1}$, $k_{i,t}$ and $m_{i,t+1}$:

$$l_{i,t+1} = \left[\frac{w_{t+1}}{\beta_l Z_{i,t+1} k_{i,t}^{\beta_k} m_{i,t+1}^{\beta_m}} \right]^{\frac{1}{\beta_l - 1}} = \left(\frac{\beta_l}{w_{t+1}} \right)^{\frac{1}{1-\beta_l}} Z_{i,t+1}^{\frac{1}{1-\beta_l}} k_{i,t}^{\frac{\beta_k}{1-\beta_l}} m_{i,t+1}^{\frac{\beta_m}{1-\beta_l}} \tag{40}$$

Use (39) and (40) to get:

$$l_{i,t+1} = \left(\frac{\beta_l}{w_{t+1}}\right)^{\frac{1}{1-\beta_l}} Z_{i,t+1}^{\frac{1}{1-\beta_l}} k_{i,t}^{\frac{\beta_k}{1-\beta_l}} \left(\frac{\beta_m w_{t+1}}{\beta_l p_{m,t+1}} l_{i,t+1}\right)^{\frac{\beta_m}{1-\beta_l}} \tag{41}$$

Rearrange to get the optimal labor demand in terms of w_{t+1}, $p_{m,t+1}$, $Z_{i,t+1}$, and $k_{i,t}$:

$$l_{i,t+1} = \left(\frac{\beta_l}{w_{t+1}}\right)^{\frac{1-\beta_m}{1-\beta_l-\beta_m}} \left(\frac{\beta_m}{p_{m,t+1}}\right)^{\frac{\beta_m}{1-\beta_l-\beta_m}} Z_{i,t+1}^{\frac{1}{1-\beta_l-\beta_m}} k_{i,t}^{\frac{\beta_k}{1-\beta_l-\beta_m}} \tag{42}$$

Use (39) and (42) to write the optimal materials demand in terms of w_{t+1}, $p_{m,t+1}$, $Z_{i,t+1}$, and $k_{i,t}$:

$$m_{i,t+1} = \left(\frac{\beta_l}{w_{t+1}}\right)^{\frac{\beta_l}{1-\beta_l-\beta_m}} \left(\frac{\beta_m}{p_{m,t+1}}\right)^{\frac{1-\beta_l}{1-\beta_l-\beta_m}} Z_{i,t+1}^{\frac{1}{1-\beta_l-\beta_m}} k_{i,t}^{\frac{\beta_k}{1-\beta_l-\beta_m}} \tag{43}$$

The expressions for labor and materials demand hold for both unconstrained and constrained firms. Substitute (42) and (43) into the production function (4) to write $p_{i,t+1}y_{i,t+1}$ in terms of w_{t+1}, $p_{m,t+1}$, $Z_{i,t+1}$, and $k_{i,t}$:

$$p_{i,t+1}y_{i,t+1} = \left(\frac{\beta_l}{w_{t+1}}\right)^{\frac{\beta_l}{1-\beta_l-\beta_m}} \left(\frac{\beta_m}{p_{m,t+1}}\right)^{\frac{\beta_m}{1-\beta_l-\beta_m}} Z_{i,t+1}^{\frac{1}{1-\beta_l-\beta_m}} k_{i,t}^{\frac{\beta_k}{1-\beta_l-\beta_m}} \tag{44}$$

Use $\text{MRPK}_{i,t} = \beta_k \frac{p_{i,t}y_{i,t}}{k_{i,t-1}}$ (14), (44), and $\beta_k + \beta_l + \beta_m = \frac{\epsilon-1}{\epsilon}$ to write the log of the marginal revenue product of capital in terms of w_t, $p_{m,t}$, $Z_{i,t}$, and $k_{i,t-1}$:

$$\begin{aligned}\ln\text{MRPK}_{i,t} =& \frac{\epsilon}{1+\epsilon\beta_k}\ln Z_{i,t} - \frac{1}{1+\epsilon\beta_k}\ln k_{i,t-1} - \frac{\epsilon\beta_l}{1+\epsilon\beta_k}\ln w_t - \frac{\epsilon\beta_m}{1+\epsilon\beta_k}\ln p_{m,t} \\ &+ \ln\left(\beta_k \beta_l^{\frac{\epsilon\beta_l}{1+\epsilon\beta_k}} \beta_m^{\frac{\epsilon\beta_m}{1+\epsilon\beta_k}}\right)\end{aligned} \tag{45}$$

A.2 Dispersion of MRPK across All Firms

Using (45), the cross-section dispersion of the marginal revenue product of capital $\text{Var}_i(\ln\text{MRPK}_{i,t})$ within a given industry in period t can be written as:

$$\begin{aligned}\text{Var}_i(\ln\text{MRPK}_{i,t}) =& \left(\frac{\epsilon}{1+\epsilon\beta_k}\right)^2 \text{Var}_i(\ln Z_{i,t}) + \left(\frac{1}{1+\epsilon\beta_k}\right)^2 \text{Var}_i(\ln k_{i,t-1}) \\ &- 2\frac{\epsilon}{(1+\epsilon\beta_k)^2}\text{Cov}_i(\ln Z_{i,t}, \ln k_{i,t-1}) \\ =& \psi_1 \text{Var}_i(\ln Z_{i,t}) + \psi_2 \text{Var}_i(\ln k_{i,t-1}) - \psi_3 \text{Cov}_i(\ln Z_{i,t}, \ln k_{i,t-1})\end{aligned} \tag{46}$$

where $\psi_1 \equiv \left(\frac{\epsilon}{1+\epsilon\beta_k}\right)^2$, $\psi_2 \equiv \left(\frac{\epsilon}{1+\epsilon\beta_k}\right)^2$, and $\psi_3 \equiv 2\frac{\epsilon}{(1+\epsilon\beta_k)^2}$. Rewrite the dispersion of MRPK in terms of the exogenous or predetermined variables by using $Z_{i,t} \equiv Z_t z_i z_{i,t}$ and the AR(1) process for $\ln z_{i,t}$ (5):

$$\begin{aligned}\mathrm{Var}_i(\ln\mathrm{MRPK}_{i,t}) =&\psi_1\mathrm{Var}_i(\ln Z_t + \ln z_i + \rho\ln z_{i,t-1} + e_{i,t}) + \psi_2\mathrm{Var}_i(\ln k_{i,t-1})\\ &- \psi_3\mathrm{Cov}_i(\ln Z_t + \ln z_i + \rho\ln z_{i,t-1} + e_{i,t}, \ln k_{i,t-1})\\ =&\psi_1\mathrm{Var}_i(\ln z_i) + \psi_1\mathrm{Var}_i(e_{i,t}) + \psi_1\rho^2\mathrm{Var}_i(\ln z_{i,t-1}) + \psi_2\mathrm{Var}_i(\ln k_{i,t-1})\\ &- \psi_3\mathrm{Cov}_i(\ln z_i + \rho\ln z_{i,t-1}, \ln k_{i,t-1})\end{aligned} \tag{15}$$

This is a general decomposition and holds even in the presence of constrained firms. Since $k_{i,t}$ is driven by different processes for unconstrained and constrained firms, by replacing $\ln k_{i,t}$ by $\ln k^U_{i,t}$ (13) or $\ln k^C_{i,t}$ (12), it is possible to find the dispersion of MRPK within the two subgroups of firms, i.e., $\mathrm{Var}_i(\ln\mathrm{MRPK}^U_{i,t})$ and $\mathrm{Var}_i(\ln\mathrm{MRPK}^C_{i,t})$.

A.3 Capital Demand of Financially Unconstrained Firms

If firm i is unconstrained in period t (i.e., $\lambda_{i,t} = 0$), then the first order condition (35) can be simplified to:

$$\mathrm{E}_t\left[\beta_k\frac{p_{i,t+1}y_{i,t+1}}{k_{i,t}}\right] = R_t - (1-\delta) = r_t + \delta \tag{47}$$

where $r_t \equiv R_t - 1$ is the net real interest rate. Rearrange (47) to get the unconstrained capital demand in terms of the expected revenue:

$$k^U_{i,t} = \frac{\beta_k}{r_t+\delta}\mathrm{E}_t\left[p_{i,t+1}y_{i,t+1}\right] \tag{48}$$

Hence,

$$\Delta\ln k^U_{i,t} \equiv \ln k^U_{i,t} - \ln k^U_{i,t-1} = \ln\frac{r_{t-1}+\delta}{r_t+\delta} + \Delta\ln\mathrm{E}_t\left[p_{i,t+1}y_{i,t+1}\right] \tag{49}$$

As can be seen above, the investment of an unconstrained firm is driven by the change in the net real interest rate and the growth in expected sales $p_{i,t+1}y_{i,t+1}$ or value added $(1-\beta_m)p_{i,t+1}y_{i,t+1}$. Alternatively, substitute (44) into (48) to get:

$$k^U_{i,t} = \frac{\beta_k}{r_t+\delta}\mathrm{E}_t\left[\left(\frac{\beta_l}{w_{t+1}}\right)^{\frac{\beta_l}{1-\beta_l-\beta_m}}\left(\frac{\beta_m}{p_{m,t+1}}\right)^{\frac{\beta_m}{1-\beta_l-\beta_m}} Z_{i,t+1}^{\frac{1}{1-\beta_l-\beta_m}} k_{i,t}^{\frac{\beta_k}{1-\beta_l-\beta_m}}\right] \tag{50}$$

Rearrange to solve for the following optimal unconstrained capital demand chosen in period t:

$$(k_{i,t}^U)^{\frac{1-\beta_l-\beta_m-\beta_k}{1-\beta_l-\beta_m}} = \frac{\beta_k}{r_t+\delta}\mathrm{E}_t\left[\left(\frac{\beta_l}{w_{t+1}}\right)^{\frac{\beta_l}{1-\beta_l-\beta_m}}\left(\frac{\beta_m}{p_{m,t+1}}\right)^{\frac{\beta_m}{1-\beta_l-\beta_m}} Z_{i,t+1}^{\frac{1}{1-\beta_l-\beta_m}}\right] \tag{51}$$

Use $Z_{i,t+1} \equiv Z_{t+1}z_i z_{i,t+1}$ and the assumption that the idiosyncratic transitory productivity $z_{i,t+1}$ is independent from the trend Z_{t+1} or the idiosyncratic permanent productivity z_i, to get:

$$\begin{aligned}\ln k_{i,t}^U =&(1+\epsilon\beta_k)\left\{\ln\left(\beta_k\beta_l^{\frac{\epsilon\beta_l}{1+\epsilon\beta_k}}\beta_m^{\frac{\epsilon\beta_m}{1+\epsilon\beta_k}}\right)-\ln(r_t+\delta)\right.\\ &\left.+\ln\mathrm{E}_t\left[\left(\frac{Z_{t+1}}{w_{t+1}^{\beta_l}p_{m,t+1}^{\beta_m}}\right)^{\frac{\epsilon}{1+\epsilon\beta_k}}\right]+\ln\mathrm{E}_t\left[z_{i,t+1}^{\frac{\epsilon}{1+\epsilon\beta_k}}\right]+\frac{\epsilon}{1+\epsilon\beta_k}\ln z_i\right\}\end{aligned} \tag{52}$$

According to (5), the firm's productivity $z_{i,t+1}$ follows the AR(1) process:

$$z_{i,t+1} = z_{i,t}^{\rho}\exp(e_{i,t+1}) \tag{53}$$

where $e_{i,t+1} \overset{i.i.d.}{\sim} N(0,\sigma_z^2)$, so

$$\mathrm{E}_t\left[z_{i,t+1}^{\frac{\epsilon}{1+\epsilon\beta_k}}\right] = \mathrm{E}_t\left[z_{i,t}^{\frac{\rho\epsilon}{1+\epsilon\beta_k}}\exp\left(\frac{e_{i,t+1}\epsilon}{1+\epsilon\beta_k}\right)\right] = z_{i,t}^{\frac{\rho\epsilon}{1+\epsilon\beta_k}}\mathrm{E}_t\left[\exp\left(\frac{e_{i,t+1}\epsilon}{1+\epsilon\beta_k}\right)\right] \tag{54}$$

Since $e_{i,t+1}$ is normally distributed, $\exp\left(\frac{e_{i,t+1}\epsilon}{1+\epsilon\beta_k}\right)$ has a log-normal distribution. Let x denote $\frac{e_{i,t+1}\epsilon}{1+\epsilon\beta_k}$, then $x \sim N\left(0,\frac{\sigma_z^2\epsilon^2}{(1+\epsilon\beta_k)^2}\right)$ and $\exp(x) \sim LogNormal\left(0,\frac{\sigma_z^2\epsilon^2}{(1+\epsilon\beta_k)^2}\right)$. Use the fact that $\mathrm{E}[\exp(x)] = \exp(\mathrm{E}[x]+\frac{1}{2}\mathrm{Var}[x])$, so

$$\mathrm{E}_t\left[\exp\left(\frac{e_{i,t+1}\epsilon}{1+\epsilon\beta_k}\right)\right] = \exp\left(\frac{\sigma_z^2\epsilon^2}{2(1+\epsilon\beta_k)^2}\right) \tag{55}$$

Substitute (55) into (54) and take logs to get:

$$\ln\mathrm{E}_t\left[z_{i,t+1}^{\frac{\epsilon}{1+\epsilon\beta_k}}\right] = \frac{\rho\epsilon}{1+\epsilon\beta_k}\ln z_{i,t} + \frac{\sigma_z^2\epsilon^2}{2(1+\epsilon\beta_k)^2} \tag{56}$$

Finally, substitute (56) into (52) to get (13):

$$\begin{aligned}\ln k_{i,t}^U =&\epsilon\rho\ln z_{i,t}+(1+\epsilon\beta_k)\left\{\ln\left(\beta_k\beta_l^{\frac{\epsilon\beta_l}{1+\epsilon\beta_k}}\beta_m^{\frac{\epsilon\beta_m}{1+\epsilon\beta_k}}\right)-\ln(r_t+\delta)\right.\\ &\left.+\ln\mathrm{E}_t\left[\left(\frac{Z_{t+1}}{w_{t+1}^{\beta_l}p_{m,t+1}^{\beta_m}}\right)^{\frac{\epsilon}{1+\epsilon\beta_k}}\right]+\frac{\epsilon}{1+\epsilon\beta_k}\ln z_i+\frac{\sigma_z^2\epsilon^2}{2(1+\epsilon\beta_k)^2}\right\}\end{aligned} \tag{13}$$

A.4 Dispersion of MRPK within Unconstrained Firms

Using (13) and (45), the marginal revenue product of capital of an unconstrained firm is:

$$\begin{aligned}
\mathrm{lnMRPK}^U_{i,t} =& \frac{\epsilon}{1+\epsilon\beta_k}\ln(Z_t z_i z_{i,t}) - \frac{\epsilon\rho}{1+\epsilon\beta_k}\ln z_{i,t-1} - \ln\left(\beta_k \beta_l^{\frac{\epsilon\beta_l}{1+\epsilon\beta_k}} \beta_m^{\frac{\epsilon\beta_m}{1+\epsilon\beta_k}}\right) \\
& + \ln(r_{t-1}+\delta) - \ln\mathrm{E}_{t-1}\left[\left(\frac{Z_t}{w_t^{\beta_l} p_{m,t}^{\beta_m}}\right)^{\frac{\epsilon}{1+\epsilon\beta_k}}\right] - \frac{\epsilon}{1+\epsilon\beta_k}\ln z_i - \frac{\sigma_z^2\epsilon^2}{2(1+\epsilon\beta_k)^2} \\
& - \frac{\epsilon\beta_l}{1+\epsilon\beta_k}\ln w_t - \frac{\epsilon\beta_m}{1+\epsilon\beta_k}\ln p_{m,t} + \ln\left(\beta_k \beta_l^{\frac{\epsilon\beta_l}{1+\epsilon\beta_k}} \beta_m^{\frac{\epsilon\beta_m}{1+\epsilon\beta_k}}\right) \\
=& \frac{\epsilon}{1+\epsilon\beta_k}(\ln z_{i,t} - \rho\ln z_{i,t-1}) + \ln\left(\frac{Z_t}{w_t^{\beta_l} p_{m,t}^{\beta_m}}\right)^{\frac{\epsilon}{1+\epsilon\beta_k}} - \ln\mathrm{E}_{t-1}\left[\left(\frac{Z_t}{w_t^{\beta_l} p_{m,t}^{\beta_m}}\right)^{\frac{\epsilon}{1+\epsilon\beta_k}}\right] \\
& + \ln(r_{t-1}+\delta) - \frac{\sigma_z^2}{2(1-\beta_l-\beta_m)^2} \\
=& \frac{\epsilon}{1+\epsilon\beta_k}e_{i,t} + \ln\left(\frac{Z_t}{w_t^{\beta_l} p_{m,t}^{\beta_m}}\right)^{\frac{\epsilon}{1+\epsilon\beta_k}} - \ln\mathrm{E}_{t-1}\left[\left(\frac{Z_t}{w_t^{\beta_l} p_{m,t}^{\beta_m}}\right)^{\frac{\epsilon}{1+\epsilon\beta_k}}\right] + \ln(r_{t-1}+\delta) \\
& - \frac{\sigma_z^2\epsilon^2}{2(1+\epsilon\beta_k)^2}
\end{aligned} \tag{57}$$

Assuming σ_z is the same for all firms, the dispersion of MRPK among unconstrained firms is:[41]

$$\mathrm{Var}_i(\mathrm{lnMRPK}^U_{i,t}) = \psi_1 \mathrm{Var}_i(e_{i,t}) \tag{16}$$

where $\psi_1 \equiv \left(\frac{\epsilon}{1+\epsilon\beta_l}\right)^2$ and i denotes an unconstrained firm i. As can be seen, the dispersion of MRPK among unconstrained firms is only driven by the cross-section dispersion of the productivity innovation $e_{i,t}$.

A.5 Capital Demand of Financially Constrained Firms

Using the assumption on the financing sources of capital (9), when the borrowing constraint (10) is binding (i.e., $\lambda_{i,t} > 0$), firm i's capital demand $k^C_{i,t}$ is determined by its net worth:

$$k^C_{i,t} = \frac{n_{i,t}}{1-\phi(1-\delta)} \tag{58}$$

Taking logs yields:

$$\ln k^C_{i,t} = \ln n_{i,t} - \ln[1-\phi(1-\delta)] \tag{12}$$

[41] Alternatively, $\mathrm{Var}_i(\mathrm{lnMRPK}^U_{i,t})$ can be found by substituting $\ln k^U_{i,t}$ in (13) for $\ln k_{i,t}$ in (15).

It can be seen from the first order condition with respect to $k_{i,t}$ (35) that:

$$\lambda_{i,t} = \mathrm{E}_t \left[\beta_k \frac{p_{i,t+1} y_{i,t+1}}{k_{i,t}} - (r_t + \delta) \right] > 0 \tag{59}$$

which implies that the expected MRPK is greater than $(r_t + \delta)$. It can be shown that the investment of a constrained firm is determined by its cash flow $\mathrm{CF}_{i,t}$, which is the revenue net of wage payments, materials costs, and net interest payments on loans, i.e., $\mathrm{CF}_{i,t} \equiv p_{i,t} y_{i,t} - w_t l_{i,t} - p_{m,t} m_{i,t} - r_{t-1} b_{i,t-1}$, assuming that debt is not repaid in each period but rolled over.[42] Using the definitions for net worth (8) and cash flow, and $k_{i,t} = n_{i,t} + b_{i,t}$ (9),

$$\begin{aligned} n_{i,t+1} &\equiv p_{i,t+1} y_{i,t+1} - w_{t+1} l_{i,t+1} - p_{m,t+1} m_{i,t+1} - R_t b_{i,t} + (1-\delta) k_{i,t} \\ &= p_{i,t+1} y_{i,t+1} - w_{t+1} l_{i,t+1} - p_{m,t+1} m_{i,t+1} - r_t b_{i,t} - \delta k_{i,t} + n_{i,t} = \mathrm{CF}_{i,t+1} - \delta k_{i,t} + n_{i,t} \end{aligned} \tag{60}$$

where the firm's net income is equal to $p_{i,t+1} y_{i,t+1} - w_{t+1} l_{i,t+1} - p_{m,t+1} m_{i,t+1} - r_t b_{i,t} - \delta k_{i,t}$, and cash flow is the sum of net income and the depreciation of capital stock.

Using (60) and the binding collateral constraint (58),

$$k^C_{i,t} - k^C_{i,t-1} = \frac{1}{1-\phi(1-\delta)} (n_{i,t} - n_{i,t-1}) = \frac{1}{1-\phi(1-\delta)} (\mathrm{CF}_{i,t} - \delta k_{i,t-1}) \tag{61}$$

A.6 Dispersion of MRPK within Constrained Firms

Using (12) and (45), the marginal revenue product of capital of a constrained firm is:

$$\begin{aligned} \ln \mathrm{MRPK}^C_{i,t} =& \frac{\epsilon}{1+\epsilon\beta_k} \ln Z_{i,t} - \frac{1}{1+\epsilon\beta_k} \left[\ln n_{i,t-1} - \ln(1-\phi(1-\delta)) \right] \\ & - \frac{\epsilon\beta_l}{1+\epsilon\beta_k} \ln w_t - \frac{\epsilon\beta_m}{1+\epsilon\beta_k} \ln p_{m,t} + \ln \left(\beta_k \beta_l^{\frac{\epsilon\beta_l}{1+\epsilon\beta_k}} \beta_m^{\frac{\epsilon\beta_m}{1+\epsilon\beta_k}} \right) \\ =& \frac{\epsilon}{1+\epsilon\beta_k} \ln Z_t + \frac{\epsilon}{1+\epsilon\beta_k} \ln z_{i,t} + \frac{\epsilon}{1+\epsilon\beta_k} \ln z_i - \frac{1}{1+\epsilon\beta_k} \ln n_{i,t-1} \\ & + \frac{1}{1+\epsilon\beta_k} \ln[1-\phi(1-\delta)] - \frac{\epsilon\beta_l}{1+\epsilon\beta_k} \ln w_t - \frac{\epsilon\beta_m}{1+\epsilon\beta_k} \ln p_{m,t} \\ & + \ln \left(\beta_k \beta_l^{\frac{\epsilon\beta_l}{1+\epsilon\beta_k}} \beta_m^{\frac{\epsilon\beta_m}{1+\epsilon\beta_k}} \right) \end{aligned} \tag{62}$$

where the last step uses $Z_{i,t} \equiv Z_t z_i z_{i,t}$. Using (62) and the AR(1) process for the idiosyncratic transitory productivity $z_{i,t}$ (5), the dispersion of MRPK among constrained firms is given

[42] Apart from the terminal period where the gross interests on debt $R_t b_t$ are repaid, assume in all the other periods, debt is rolled over and only net interests on debt $r_t b_t$ are repaid.

by:[43]

$$
\begin{aligned}
\mathrm{Var}_i(\ln\mathrm{MRPK}^C_{i,t}) =& \psi_1\mathrm{Var}_i(\ln z_i) + \psi_1\mathrm{Var}_i(e_{i,t}) + \psi_1\rho^2\mathrm{Var}_i(\ln z_{i,t-1}) + \psi_2\mathrm{Var}_i(\ln n_{i,t-1}) \\
& - \psi_3\mathrm{Cov}_i(\ln z_i + \rho\ln z_{i,t-1}, \ln n_{i,t-1}) \qquad (17) \\
=& \psi_1\mathrm{Var}_i(e_{i,t}) + \mathrm{Var}_i(\psi_1^{\frac{1}{2}}\ln z_i + \psi_1^{\frac{1}{2}}\rho\ln z_{i,t-1} - \psi_2^{\frac{1}{2}}\ln n_{i,t-1})
\end{aligned}
$$

where $\psi_1 \equiv \left(\frac{\epsilon}{1+\epsilon\beta_k}\right)^2$, $\psi_2 \equiv \left(\frac{1}{1+\epsilon\beta_k}\right)^2$, $\psi_3 \equiv 2\frac{\epsilon}{(1+\epsilon\beta_k)^2}$, and i denotes a constrained firm i. Using (16) and (17), it can be seen that:

$$
\mathrm{Var}_i(\ln\mathrm{MRPK}^C_{i,t}) > \mathrm{Var}_i(\ln\mathrm{MRPK}^U_{i,t}) \qquad (63)
$$

since $\mathrm{Var}_i(\psi_1^{\frac{1}{2}}\ln z_i + \psi_1^{\frac{1}{2}}\rho\ln z_{i,t-1} - \psi_2^{\frac{1}{2}}\ln n_{i,t-1}) > 0$.

B Decomposition of the Dispersion of MRPK

Suppose there are M_t firms in a given industry and N_t of them are unconstrained in a given period t, where $N_t \leqslant M_t$, and the remaining $M_t - N_t$ firms are constrained. The distribution of the observed lnMRPK in the data is a mixture of two distributions of $\ln\mathrm{MRPK}^U$ (for unconstrained firms) and $\ln\mathrm{MRPK}^C$ (for constrained firms). It is shown below that the variance of lnMRPK across all firms in a given industry and a given period t can be written in terms of the variances and means over the subgroups (unconstrained U and constrained C) of firms.

Let $X_{i,t}$ denote $\ln\mathrm{MRPK}^U_{i,t}$ and $Y_{i,t}$ denote $\ln\mathrm{MRPK}^C_{i,t}$ in a given time period. For simplicity, the subscripts i and t are suppressed for $X_{i,t}$ and $Y_{i,t}$ in the following proof. Order the firms in such a way that the first N_t firms according to the index i are unconstrained

[43] Alternatively, $\mathrm{Var}_i(\ln\mathrm{MRPK}^C_{i,t})$ can be found by substituting $\ln k^C_{i,t}$ in (12) for $\ln k_{i,t}$ in (15).

and the rest of firms are constrained (i.e., firms N_t+1 to M_t). In a given industry-year:

$$
\begin{aligned}
&\mathrm{Var}_i(\mathrm{lnMRPK}_{i,t}) = \mathrm{E}_i(\mathrm{lnMRPK}^2_{i,t}) - \mathrm{E}_i(\mathrm{lnMRPK}_{i,t})^2\\
=&\frac{1}{M_t}\left[\sum_{i=1}^{N_t}(\mathrm{lnMRPK}^U_{i,t})^2 + \sum_{i=N_t+1}^{M_t}(\mathrm{lnMRPK}^C_{i,t})^2\right]\\
&-\left[\frac{1}{M_t}\left(\sum_{i=1}^{N_t}\mathrm{lnMRPK}^U_{i,t} + \sum_{i=N_t+1}^{M_t}\mathrm{lnMRPK}^C_{i,t}\right)\right]^2\\
=&\frac{1}{M_t}\left[\sum_{i=1}^{N_t}X^2 + \sum_{i=N_t+1}^{M_t}Y^2\right] - \left[\frac{1}{M_t}\left(\sum_{i=1}^{N_t}X + \sum_{i=N_t+1}^{M_t}Y\right)\right]^2\\
=&\frac{N_t}{M_t}\frac{1}{N_t}\sum_{i=1}^{N_t}X^2 + \frac{M_t-N_t}{M_t}\frac{1}{M_t-N_t}\sum_{i=N_t+1}^{M_t}Y^2 - \left(\frac{N_t}{M_t}\frac{1}{N_t}\sum_{i=1}^{N_t}X + \frac{M_t-N_t}{M_t}\frac{1}{M_t-N_t}\sum_{i=N_t+1}^{M_t}Y\right)^2\\
=&\frac{N_t}{M_t}\mathrm{E}_i(X^2) + \frac{M_t-N_t}{M_t}\mathrm{E}_i(Y^2) - \left(\frac{N_t}{M_t}\mathrm{E}_i(X) + \frac{M_t-N_t}{M_t}\mathrm{E}_i(Y)\right)^2\\
=&\frac{N_t}{M_t}\mathrm{E}_i(X^2) + \frac{M_t-N_t}{M_t}\mathrm{E}_i(Y^2) - \left(\frac{N_t}{M_t}\right)^2\mathrm{E}_i(X)^2 - \frac{2N_t(M_t-N_t)}{M^2}\mathrm{E}_i(X)\mathrm{E}_i(Y)\\
&-\left(\frac{M_t-N_t}{M_t}\right)^2\mathrm{E}_i(Y)^2\\
=&\frac{N_t}{M_t}\left[\mathrm{E}_i(X^2)-\mathrm{E}_i(X)^2\right] + \frac{N_t}{M_t}\left(1-\frac{N_t}{M_t}\right)\mathrm{E}_i(X)^2 + \frac{M_t-N_t}{M_t}\left[\mathrm{E}_i(Y^2)-\mathrm{E}_i(Y)^2\right]\\
&+\frac{M_t-N_t}{M_t}\left(1-\frac{M_t-N_t}{M_t}\right)\mathrm{E}_i(Y)^2 - \frac{2N_t(M_t-N_t)}{M^2}\mathrm{E}_i(X)\mathrm{E}_i(Y)\\
=&\frac{N_t}{M_t}\mathrm{Var}_i(X) + \frac{M_t-N_t}{M_t}\mathrm{Var}_i(Y) + \frac{N_t(M_t-N_t)}{M^2}\left[\mathrm{E}_i(X)^2 + \mathrm{E}_i(Y)^2 - 2\mathrm{E}_i(X)\mathrm{E}_i(Y)\right]\\
=&\frac{N_t}{M_t}\mathrm{Var}_i(X) + \frac{M_t-N_t}{M_t}\mathrm{Var}_i(Y) + \frac{N_t(M_t-N_t)}{M^2}\left[\mathrm{E}_i(X) - \mathrm{E}_i(Y)\right]^2
\end{aligned}
\tag{64}
$$

where $\mathrm{E}_i(X)$ denotes the mean of $\mathrm{lnMRPK}_{i,t}$ across all the unconstrained firms i and $\mathrm{E}_i(Y)$ denotes the mean of $\mathrm{lnMRPK}_{i,t}$ across all the constrained firms i in a given period t. Similarly, $\mathrm{E}_i(X^2)$ and $\mathrm{Var}_i(X)$ are defined over the subgroup of unconstrained firms and $\mathrm{E}_i(Y^2)$ and $\mathrm{Var}_i(Y)$ are defined over the subgroup of constrained firms.

The last term in (64) is the squared difference between the mean values of lnMRPK within the two subgroups of firms, weighted by the product of the two fractions $\frac{N_t}{M_t}$ and $\frac{M_t-N_t}{M_t}$, which are the proportions of unconstrained firms and constrained firms respectively for each industry-year.

C Production Function Estimation

Once the revenue elasticities (β_k, β_l and β_m) for each two-digit NACE Rev.2 industry are estimated, the log of revenue-based productivity (TFPR) for firm i in a given industry at time t is the residual term after subtracting the weighted sum of inputs from $\ln(p_{i,t}y_{i,t})$:

$$\log\text{TFPR}_{i,t} \equiv \ln Z_{i,t} = \ln(p_{i,t}y_{i,t}) - \beta_k \ln k_{i,t} - \beta_l \ln l_{i,t} - \beta_m \ln m_{i,t} \tag{65}$$

where $p_{i,t}y_{i,t}$ is measured by the nominal revenue, $k_{i,t}$, $m_{i,t}$ and $l_{i,t}$ are measured by the book value of fixed tangible assets, material costs, and the wage bill, respectively. Wage bill is used to measure $l_{i,t}$ to control for the quality differences of labor across firms, following Gopinath et al. (2017). Labor and materials are variable inputs, whereas capital $k_{i,t}$ is the state variable, which is equivalent to $k_{i,t-1}$ in the model described in Section 2.

This paper uses the Wooldridge (2009) estimation-based approach. Wooldridge (2009) show that the two-step estimation proposed by Olley and Pakes (1996) (OP) and Levinsohn and Petrin (2003) (LP) can be implemented in one step using GMM, by applying different instruments to each of the two equations. As he pointed out, there are two advantages of using the joint GMM estimation compared to the two-step methods. First, if the variable input (labor) is also determined by unobserved productivity and state variables, then the coefficient on labor is unidentified in the first-stage estimation (Ackerberg, Caves and Frazer, 2006) and hence two-step estimation does not work in this case. Second, it is easy to obtain fully robust standard errors using joint estimation.

The capital, labor, and materials coefficients are estimated for each two-digit NACE Rev.2 industry separately. The use of two-digit industries is to make sure there are enough observations in each industry to carry out the estimation.[44] For each firm i within a two-digit industry in period t:

$$\ln(p_{i,t}y_{i,t}) = \beta_0 + \beta_k \ln k_{i,t} + \beta_l \ln l_{i,t} + \beta_m \ln m_{i,t} + \ln Z_{i,t} + \varsigma_{i,t} \tag{66}$$

where the sequence $\ln Z_{i,t}$ is the unobserved revenue-based productivity and $\varsigma_{i,t}$ is a sequence of shocks that are assumed to be conditional mean independent of current and past inputs. Under OP and LP, the unobserved productivity is proxied by an unknown function of capital

[44] In most countries, the manufacture of tobacco products (industry 12 from NACE Rev.2 code) appears to be quite concentrated and the number of firm-year observations is very small, so the revenue elasticities for this industry are not estimated in those countries. In this paper, at least 200 firm-year observations are required to implement this method.

and investment (under OP) or intermediate inputs (under LP):

$$\ln Z_{i,t} = f(\ln k_{i,t}, \ln m_{i,t}) \tag{67}$$

where the log of material costs $\ln m_{i,t}$ is the proxy variable. To estimate β_k, β_l and β_m jointly, Wooldridge (2009) assumes:

$$\mathrm{E}_i(\varsigma_{i,t}|\ln l_{i,t}, \ln k_{i,t}, \ln m_{i,t}, \ln l_{i,t-1}, \ln k_{i,t-1}, \ln m_{i,t-1}, ..., \ln l_{i,1}, \ln k_{i,1}, \ln m_{i,1}) = 0 \tag{68}$$

where $t = 1, 2, ..., T$. It can be seen that serial dependence in the idiosyncratic shocks $\varsigma_{i,t}$ is allowed in the above assumption, since past values of $\varsigma_{i,t}$ do not appear in the conditioning set. The following sufficient condition is used to restrict the dynamics of the productivity process $\ln Z_{i,t}$:

$$\begin{aligned} &\mathrm{E}_i(\ln Z_{i,t}|\ln k_{i,t}, \ln l_{i,t-1}, \ln k_{i,t-1}, \ln m_{i,t-1}, ..., \ln l_{i,1}, \ln k_{i,1}, \ln m_{i,1}) \\ &\qquad = \mathrm{E}_i(\ln Z_{i,t}|\ln Z_{i,t-1}) \equiv g[f(\ln k_{i,t-1}, \ln m_{i,t-1})] \end{aligned} \tag{69}$$

where $g(.)$ is an unknown function representing the process of productivity $\ln Z_{i,t}$. The last equality follows from $\ln Z_{i,t-1} = f(\ln k_{i,t-1}, \ln m_{i,t-1})$. The above assumption means that $\ln k_{i,t}$, past outcomes on $(\ln l_{i,t}, \ln k_{i,t}, \ln m_{i,t})$, and all functions of these are uncorrelated with the innovations $e_{Z,i,t} = \ln Z_{i,t} - E(\ln Z_{i,t}|\ln Z_{i,t-1})$.

Using $\ln Z_{i,t} = f(\ln k_{i,t}, \ln m_{i,t})$ and $\ln Z_{i,t} = g[f(\ln k_{i,t-1}, \ln m_{i,t-1})] + e_{Z,i,t}$, the two equations used to identify β_k, β_l, and β_m are:

$$\ln(p_{i,t}y_{i,t}) = \beta_0 + \beta_k \ln k_{i,t} + \beta_l \ln l_{i,t} + \beta_m \ln m_{i,t} + f(\ln k_{i,t}, \ln m_{i,t}) + \varsigma_{i,t}, \;\; t = 1, ..., T \tag{70}$$

$$\ln(p_{i,t}y_{i,t}) = \beta_0 + \beta_k \ln k_{i,t} + \beta_l \ln l_{i,t} + \beta_m \ln m_{i,t} + g[f(\ln k_{i,t-1}, \ln m_{i,t-1})] + \xi_{i,t}, \;\; t = 2, ..., T \tag{71}$$

where $\xi_{i,t} \equiv e_{Z,i,t} + \varsigma_{i,t}$. The orthogonality condition on the error term for the first equation is (68), and for the second equation, it is:

$$\mathrm{E}_i(\ln \xi_{i,t}|\ln k_{i,t}, \ln l_{i,t-1}, \ln k_{i,t-1}, \ln m_{i,t-1}, ..., \ln l_{i,1}, \ln k_{i,1}, \ln m_{i,1}) = 0, \;\; t = 2, ..., T \tag{72}$$

These two different orthogonality conditions on the error terms for the two equations imply that different instruments can be used for each equation. For instance, the state variable (capital $\ln k_{i,t}$), any lagged inputs or functions of these variables can be used as instrumental variables for both equations. In addition, the intermediate inputs (investment or intermediate inputs $\ln m_{i,t}$) can also be used as instruments for the first equation.

I use the prodest (Rovigatti and Mollisi, 2016) in Stata to calculate the productivity measure used in this paper. A third-degree polynomial is used to estimate the unknown functions $f(.,.)$ and $g(.)$, as suggested by Petrin, Poi and Levinsohn (2004). $f(\ln k_{i,t}, \ln m_{i,t})$ is approximated by all polynomials of order three or less, (i.e., $(\ln k_{i,t})^{q_1}(\ln m_{i,t})^{q_2}$ where $q_1 + q_2 \leqslant 3$, with $q_1 \geqslant 0$ and $q_2 \geqslant 0$) and can be written as:

$$f(\ln k_{i,t}, \ln m_{i,t}) \approx \vartheta_0 + \mathbf{\Gamma}(\ln k_{i,t}, \ln m_{i,t})\boldsymbol{\vartheta} \equiv \vartheta_0 + \boldsymbol{\Gamma_{i,t}\vartheta} \tag{73}$$

where $\mathbf{\Gamma}_{i,t} \equiv \mathbf{\Gamma}(\ln k_{i,t}, \ln m_{i,t})$ is a vector of $1 \times Q$ vector of functions (polynomials) and $\boldsymbol{\vartheta}$ is a vector of $Q \times 1$ parameters.[45] In addition, $g(.)$ is assumed to be approximated by a G-degree polynomial in $\ln Z_{i,t}$:

$$g(\ln Z_{i,t}) = \rho_0 + \rho_1 \ln Z_{i,t} + ...\rho_G(\ln Z_{i,t})^G \tag{74}$$

where $G = 1$ is used in the prodest package. Substitute the polynomial approximations for the unknown functions into (70) and (71) and rearrange to write the two equations as a vector of residuals $\boldsymbol{\Lambda_{i,t}(\theta)}$:

$$\begin{aligned}\boldsymbol{\Lambda_{i,t}(\theta)} &= \begin{pmatrix} \varsigma_{i,t}(\boldsymbol{\theta}) \\ \xi_{i,t}(\boldsymbol{\theta}) \end{pmatrix} \\ &= \begin{pmatrix} \ln(p_{i,t}y_{i,t}) - \alpha_0 - \beta_k \ln k_{i,t} - \beta_l \ln l_{i,t} - \beta_m \ln m_{i,t} - \boldsymbol{\Gamma_{i,t}\vartheta} \\ \ln(p_{i,t}y_{i,t}) - \zeta_0 - \beta_k \ln k_{i,t} - \beta_l \ln l_{i,t} - \beta_m \ln m_{i,t} - \rho_1 \ln(\boldsymbol{\Gamma_{i,t-1}\vartheta})... - \rho_G(\boldsymbol{\Gamma_{i,t-1}\vartheta})^G \end{pmatrix}\end{aligned} \tag{75}$$

where $\alpha_0 = \beta_0 + \vartheta_0$ and ζ_0 are the new intercepts and $\boldsymbol{\theta}$ is a vector of coefficients to be estimated. The assumption of exogenous instruments $\boldsymbol{\tau_{i,t}}$ gives rise to the following moment conditions:

$$\mathrm{E}_i[\boldsymbol{\tau_{i,t}}'\boldsymbol{\Lambda_{i,t}(\theta)}] = 0 \quad t = 2, ..., T \tag{76}$$

GMM estimation can then be applied to find the vector of coefficients $\hat{\boldsymbol{\theta}}$. All the instruments for the second equation are also valid for the first equation, while the first equation has two additional instruments, the contemporaneous values of $\ln l_{i,t}$ and $\ln m_{i,t}$. The instruments used are:

$$\boldsymbol{\tau_{i,t}} \equiv \begin{pmatrix} \boldsymbol{\tau_{i,t,1}} & 0 \\ 0 & \boldsymbol{\tau_{i,t,2}} \end{pmatrix}, \quad t = 2, ..., T \tag{77}$$

where

$$\boldsymbol{\tau_{i,t,1}} = (\ln l_{i,t}, \boldsymbol{\Gamma_{i,t}}) \tag{78}$$

[45] Wooldridge (2009) assumes that $\boldsymbol{\Gamma_{i,t}}$ includes at least $\ln k_{i,t}$ and $\ln m_{i,t}$ separately to nest the linear version of $f(\ln k_{i,t}, \ln m_{i,t})$ as a special case.

$$\boldsymbol{\tau_{i,t,2}} = (\text{ln}k_{i,t}, \text{ln}l_{i,t-1}, \boldsymbol{\Gamma_{i,t-1}}) \tag{79}$$

where $\text{ln}m_{i,t}$ is included in $\boldsymbol{\Gamma_{i,t}}$. A key difference in the sets of instruments is that $\boldsymbol{\tau_{i,t,2}}$ does not include the contemporaneous values of $\text{ln}l_{i,t}$ and $\text{ln}m_{i,t}$.

Table 7: Revenue Elasticities of Inputs and Correlations

	Revenue Elasticities			Correlations between $\Delta\text{logTFPR}_{i,t-1}$ and				
Country	β_k	β_l	β_m	$\Delta\text{lnSales}_{i,t-1}$	$\Delta\text{lnVA}_{i,t-1}$	$\Delta\text{lnFTA}_{i,t}$	$\frac{\text{CF}_{i,t-1}}{\text{FTA}_{i,t-2}}$	$\Delta\text{ln}n_{i,t-1}$
Bulgaria	0.09	0.31	0.44	0.595***	0.706***	0.039***	0.054***	0.170***
Croatia	0.03	0.28	0.67	0.441***	0.562***	0.030***	0.080***	0.171***
Czech Republic	0.02	0.34	0.52	0.537***	0.647***	0.031***	0.089***	0.213***
Finland	0.06	0.40	0.34	0.533***	0.570***	0.054***	0.127***	0.285***
France	0.04	0.49	0.29	0.530***	0.547***	0.054***	0.110***	0.285***
Germany	0.04	0.42	0.33	0.689***	0.410***	0.058***	0.066***	0.186***
Italy	0.06	0.32	0.40	0.525***	0.569***	0.046***	0.102***	0.158***
Korea	0.02	0.09	0.84	0.403***	0.634***	0.046***	0.086***	0.198***
Norway	0.02	0.39	0.34	0.577***	0.601***	0.047***	0.092***	0.233***
Poland	0.04	0.30	0.52	0.552***	0.651***	0.045***	0.079***	0.243***
Portugal	0.07	0.42	0.39	0.545***	0.416***	0.041***	0.152***	0.226***
Romania	0.14	0.29	0.45	0.560***	0.629***	0.047***	0.078***	0.209***
Serbia	0.11	0.26	0.56	0.504***	0.616***	0.031***	0.060***	0.092***
Slovakia	0.08	0.27	0.55	0.549***	0.633***	0.037***	0.070***	0.180***
Slovenia	0.06	0.36	0.42	0.476***	0.542***	0.031***	0.034***	0.112***
Spain	0.04	0.41	0.42	0.502***	0.358***	0.045***	0.115***	0.193***
Sweden	0.05	0.35	0.33	0.390***	0.545***	0.028***	0.090***	0.196***
Ukraine	0.11	0.46	0.38	0.579***	0.628***	0.025***	0.076***	0.115***
United Kingdom	0.04	0.22	0.60	0.491***	0.663***	0.060***	0.106***	0.226***

Note: The tables shows the mean revenue elasticities of capital, labor and materials (i.e., β_k, β_l, β_m) calculated using the Wooldridge (2009) estimation-based method, and the correlations between the estimated lagged productivity $\Delta\text{lnTFPR}_{i,t-1}$ and different variables, including the lagged sales growth $\Delta\text{lnSales}_{i,t-1}$, lagged value added growth $\Delta\text{lnVA}_{i,t-1}$, net capital investment or capital growth $\Delta\text{lnFTA}_{i,t}$, lagged cash flow $\frac{\text{CF}_{i,t-1}}{\text{FTA}_{i,t-2}}$, and lagged net worth growth $\Delta\text{ln}n_{i,t-1}$. The time period covered is early 1990s to 2015. The exact sample period differs across countries, as can be found in Table 1. The stars indicate the significance of the correlation coefficients. Note that Japan is excluded because material costs are not available to estimate TFPR.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

D Data Cleaning and Summary Statistics

The following cleaning steps are applied to datasets extracted from the Orbis Historical Financial database for each country:

- Drop if industry code is missing.
- Consolidation code: only keep C1, U1, U2
- Only keep the entire calendar year: drop if the number of months is not equal to 12
- Accounting year: following Kalemli-Ozcan et al. (2015), if the closing date is before June 1st, then it should be counted as the previous year.
- Basic reporting mistakes: 1) Drop if both operating revenue and sales are missing. 2) Drop negative number of employees, negative fixed tangible assets and negative sales. Note that operating revenue in Orbis equals the sum of sales, other operating revenues and stock variations, so operating revenue can be negative.
  3) Drop if interest paid, depreciation, long-term debt, short-term debt, employees cost and material costs are negative.
- Following Appendix A3 of Gopinath et al. (2017), drop if age is negative, where age is computed as the difference between year and incorporation year plus one.
- Keep only one filing type for each firm throughout the years. Each firm can have a mixture of two filing types throughout time, i.e., annual report and local registry filing (majority). I find that annual report is often associated with consolidated account (C1), whereas local registry filing is often associated with unconsolidated accounts (U1 or U2). Since empirical analysis looks at within-firm over time variation, it is important to make sure that each firm only has one filing type or consolidation code over time. Whenever a firm has a mixture of filing types across years, the filing type that has more observations is kept. If the two filing types occur with the same frequency for a given firm, then one filing type is chosen if it has greater availability of other variables.
- Keep either consolidated or unconsolidated account for each firm throughout the years. After the previous step, the consolidation code for a firm should be consistent over time.
- Drop duplicates: each firm can have multiple entries for the same year. Duplicates are dropped according to several criteria.
  1) Accounting years can differ across countries. The month of the closing date that has the largest observations is the preferred month. Suppose it is 12 (December), then

when dropping duplicates based on month, then month 12 is kept if this also occurs most frequently within the firm over time, also conditional on firm id, year, Ticker and industry code.
2) After the previous step, if there are still duplicates, drop the duplicate entry with missing Ticker, conditional on firm id, year, industry code, month of the account closing date, and total assets being the same.
3) After the first two steps, if there are still duplicates, drop the duplicate entry with missing ISIN number, conditional on firm id, year, industry code, month of the account closing date, Ticker and total assets being the same.

- The original dataset is in US dollars. Convert the variables (with monetary value) into domestic currency using the exchange rate variable in the dataset.

- This paper focuses on the manufacturing industry so that the capital stock can be well measured by the fixed tangible assets. For each country, only the manufacturing industry (two-digit NACE Rev.2 Code in the range of 10-33) is kept. The description for each two-digit industry can be found in Table 11.

- Further cleaning: Missing operating revenue (used to calculate the sales growth) and missing or zero fixed tangible assets (used to measure capital stock $k_{i,t}$) are dropped.[46] Firm-year observations with fewer than 3 consecutive years are dropped, since in the empirical regressions, lagged growth rates are used. Years with fewer than 50 firms are dropped, which happens in the earlier sample period in some countries.

- Winsorization: before running regressions for each industry or country, variables are winsorized at the 1st and 99th percentiles in the relevant sample. Variables that need winsorization include: capital growth or firm investment, sales growth, value added growth, productivity growth, cash flow over lagged capital stock, net worth growth, net worth-to-assets ratio, cash-to-assets ratio. Variables such as log of MRPK and log of total assets do not have high kurtosis and winsorization is not necessary.

[46] In Orbis data, the variable 'operating revenue' represents the turnover or sales, while the 'sales' variable represents the net sales.

Table 8: Summary Statistics of Selected Variables for Each Country in the Baseline Sample

	$\Delta\ln\text{FTA}_{i,t}$			$\Delta\ln\text{Sales}_{i,t}$			$\Delta\ln\text{VA}_{i,t}$			$\Delta\ln\text{TFPR}_{i,t}$		
Country	Mean	Median	sd	Mean	Median	sd	Mean	Median	sd	Mean	Median	sd
Bulgaria	0.059	-0.026	0.616	0.046	0.047	0.654	0.053	0.048	0.687	-0.009	-0.002	0.403
Croatia	0.010	-0.038	0.717	0.003	0.020	0.652	0.033	0.031	0.692	-0.009	-0.007	0.298
Czech Republic	0.049	-0.026	0.568	0.036	0.034	0.433	0.042	0.035	0.493	-0.002	-0.001	0.221
Finland	0.002	-0.066	0.421	0.017	0.026	0.398	0.027	0.033	0.366	0.005	0.007	0.208
France	-0.021	-0.102	0.527	0.028	0.024	0.216	0.029	0.026	0.208	0.003	0.004	0.138
Germany	0.041	-0.033	0.449	0.041	0.012	0.219	0.040	0.038	0.348	0.010	0.009	0.131
Italy	0.023	-0.033	0.505	-0.000	0.020	0.444	0.020	0.026	0.389	-0.007	0.001	0.210
Japan	0.008	-0.024	0.321	0.004	0.009	0.279	0.083	0.018	0.558			
Korea	0.154	-0.000	0.670	0.131	0.084	0.483	0.138	0.101	0.557	0.006	0.005	0.130
Norway	-0.002	-0.059	0.606	0.051	0.039	0.393	0.051	0.043	0.358	0.014	0.011	0.190
Poland	0.055	-0.019	0.462	0.045	0.048	0.389	0.052	0.052	0.402	0.003	0.003	0.199
Portugal	-0.030	-0.074	0.521	0.015	0.012	0.394	0.022	0.019	0.495	-0.004	0.001	0.225
Romania	0.086	-0.003	0.824	0.083	0.108	0.852	0.117	0.137	0.926	0.003	0.003	0.455
Serbia	0.135	-0.008	0.653	0.081	0.095	0.780	0.150	0.130	0.799	-0.002	-0.008	0.425
Slovakia	-0.009	-0.053	0.666	0.005	0.027	0.610	0.023	0.033	0.608	-0.007	-0.001	0.306
Slovenia	-0.057	-0.060	0.821	-0.018	0.034	0.706	0.045	0.034	0.440	-0.003	0.001	0.220
Spain	0.032	-0.038	0.496	0.011	0.027	0.333	0.022	0.034	0.401	-0.007	-0.000	0.187
Sweden	-0.034	-0.077	0.550	0.023	0.031	0.378	0.024	0.031	0.364	-0.006	0.007	0.303
Ukraine	0.023	-0.040	0.619	0.031	0.076	0.899	0.088	0.109	0.853	-0.015	-0.009	0.518
United Kingdom	-0.007	-0.039	0.442	0.028	0.030	0.316	0.052	0.050	0.372	0.005	0.005	0.125

Note: The table shows the mean, median and standard deviation for each of the four variables: capital investment or capital growth $\Delta\ln\text{FTA}_{i,t}$, sales growth $\Delta\ln\text{Sales}_{i,t}$, value added growth $\Delta\ln\text{VA}_{i,t}$, and the productivity growth $\Delta\ln\text{TFPR}_{i,t}$, where the productivity is estimated using the Wooldridge (2009) approach. The time period covered is early 1990s to 2015. The exact sample period differs across countries, as can be found in Table 1. Note that TFPR cannot be estimated for Japan due to the lack of data on material costs.

Table 9: Summary Statistics of Selected Variables for Each Country in the Baseline Sample

	$\frac{\text{CF}_{i,t-1}}{\text{FTA}_{i,t-2}}$			$\Delta\ln n_{i,t-1}$			Net worth/Assets			Cash/Assets		
Country	Mean	Median	sd	Mean	Median	sd	Mean	Median	sd	Mean	Median	sd
Bulgaria	1.403	0.326	4.326	0.176	0.093	0.513	0.388	0.446	0.524	0.196	0.082	0.245
Croatia	1.266	0.313	4.540	0.093	0.046	0.508	0.260	0.269	0.495	0.087	0.032	0.132
Czech Republic	1.148	0.291	3.711	0.109	0.064	0.435	0.356	0.430	0.522	0.157	0.087	0.180
Finland	1.180	0.424	3.350	0.075	0.059	0.459	0.375	0.422	0.469	0.182	0.109	0.196
France	1.621	0.652	4.197	0.083	0.067	0.368	0.328	0.349	0.316	0.177	0.115	0.183
Germany	1.632	0.384	6.184	0.097	0.056	0.452	0.340	0.304	0.257	0.138	0.063	0.171
Italy	0.813	0.284	2.846	0.084	0.041	0.442	0.214	0.182	0.292	0.082	0.028	0.120
Japan	0.269	0.122	0.904	0.051	0.035	0.307	0.185	0.220	0.494	0.203	0.161	0.162
Korea	1.353	0.237	3.984	0.182	0.131	0.480	0.395	0.365	0.270	0.060	0.021	0.094
Norway	1.760	0.414	6.157	0.094	0.068	0.488	0.276	0.285	0.362	0.185	0.111	0.197
Poland	1.250	0.316	4.376	0.103	0.070	0.406	0.434	0.486	0.406	0.103	0.045	0.138
Portugal	0.434	0.254	2.685	0.089	0.053	0.438	0.216	0.268	0.504	0.130	0.058	0.170
Romania	1.289	0.231	5.066	0.193	0.085	0.944	0.047	0.214	0.980	0.112	0.041	0.165
Serbia	0.797	0.250	3.365	0.209	0.116	0.562	0.391	0.368	0.308	0.055	0.017	0.093
Slovakia	0.781	0.267	2.709	0.066	0.045	0.684	0.168	0.277	0.740	0.152	0.073	0.191
Slovenia	1.221	0.298	3.699	-0.011	0.044	0.783	0.478	0.488	0.337	0.104	0.039	0.152
Spain	0.617	0.258	2.210	0.101	0.067	0.366	0.284	0.283	0.363	0.116	0.060	0.142
Sweden	1.384	0.395	4.814	0.077	0.057	0.396	0.422	0.419	0.278	0.180	0.105	0.201
Ukraine	0.717	0.100	5.769	0.090	0.019	0.618	0.338	0.519	0.801	0.073	0.015	0.137
United Kingdom	2.018	0.432	7.571	0.087	0.071	0.529	0.246	0.348	0.657	0.149	0.069	0.194

Note: The table shows the mean, median and standard deviation for each of the four variables: lagged cash flow over twice lagged fixed tangible assets $\frac{\text{CF}_{i,t-1}}{\text{FTA}_{i,t-2}}$, lagged net worth growth $\Delta\ln n_{i,t-1}$, net-worth-to-assets ratio, and cash-to-assets ratio. The time period covered is early 1990s to 2015. The exact sample period differs across countries, as can be found in Table 1.

Table 10: Correlations between Lagged Sales Growth and Different Variables for Each Country

Country	$\Delta \ln \text{FTA}_{i,t}$	$\frac{\text{CF}_{i,t-1}}{\text{FTA}_{i,t-2}}$	$\Delta \ln n_{i,t-1}$	$\Delta \ln \text{VA}_{i,t-1}$	$\Delta \ln \text{TFPR}_{i,t-1}$
Bulgaria	0.133***	0.187***	0.397***	0.815***	0.595***
Croatia	0.122***	0.166***	0.333***	0.727***	0.441***
Czech Republic	0.091***	0.156***	0.329***	0.759***	0.537***
Finland	0.102***	0.167***	0.322***	0.872***	0.533***
France	0.107***	0.169***	0.341***	0.892***	0.530***
Germany	0.091***	0.106***	0.198***	0.593***	0.689***
Italy	0.114***	0.181***	0.252***	0.846***	0.525***
Japan	0.075***	0.219***	0.302***	0.301***	
Korea	0.103***	0.175***	0.297***	0.729***	0.403***
Norway	0.080***	0.100***	0.247***	0.832***	0.577***
Poland	0.133***	0.155***	0.355***	0.842***	0.552***
Portugal	0.121***	0.182***	0.314***	0.659***	0.545***
Romania	0.069***	0.225***	0.501***	0.842***	0.560***
Serbia	0.166***	0.197***	0.346***	0.736***	0.504***
Slovakia	0.078***	0.154***	0.403***	0.807***	0.549***
Slovenia	0.023***	0.121***	0.732***	0.860***	0.476***
Spain	0.120***	0.168***	0.291***	0.607***	0.502***
Sweden	0.102***	0.160***	0.331***	0.803***	0.390***
Ukraine	0.114***	0.158***	0.272***	0.802***	0.579***
United Kingdom	0.105***	0.122***	0.262***	0.517***	0.491***

Note: The table shows the correlations between lagged sales growth $\Delta \ln \text{Sales}_{i,t-1}$ and different variables, including net capital investment or capital growth $\Delta \ln \text{FTA}_{i,t}$, lagged cash flow over twice lagged capital stock $\frac{\text{CF}_{i,t-1}}{\text{FTA}_{i,t-2}}$, lagged net worth growth $\Delta \ln n_{i,t-1}$, lagged value added growth $\Delta \ln \text{VA}_{i,t-1}$, and lagged productivity growth $\Delta \ln \text{TFPR}_{i,t-1}$, where the productivity is estimated using the Wooldridge (2009) approach. The stars indicate the significance of the correlation coefficients. Note that TFPR cannot be estimated for Japan due to the lack of data on material costs.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Industry Classification: NACE Rev.2 Code for Manufacturing

Nace Code	Descriptions
10	Manufacture of food products
11	Manufacture of beverages
12	Manufacture of tobacco products
13	Manufacture of textiles
14	Manufacture of wearing apparel
15	Manufacture of leather and related products
16	Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials
17	Manufacture of paper and paper products
18	Printing and reproduction of recorded media
19	Manufacture of coke and refined petroleum products
20	Manufacture of chemicals and chemical products
21	Manufacture of basic pharmaceutical products and pharmaceutical preparations
22	Manufacture of rubber and plastic products
23	Manufacture of other non-metallic mineral products
24	Manufacture of basic metals
25	Manufacture of fabricated metal products, except machinery and equipment
26	Manufacture of computer, electronic and optical products
27	Manufacture of electrical equipment
28	Manufacture of machinery and equipment
29	Manufacture of motor vehicles, trailers and semi-trailers
30	Manufacture of other transport equipment
31	Manufacture of furniture
32	Other manufacturing
33	Repair and installation of machinery and equipment

Note: The table shows the NACE Rev. 2 Code for the two-digit industries in the manufacturing sector and their corresponding descriptions. More detailed industry classification can be found: https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF.

E Exogenous Switching Regression Model

The likelihood function $L_{i,t}$ (31) of an observation in the exogenous switching regression model is derived below:

$$\begin{aligned}
L_{i,t} =& f(\varepsilon_{C,i,t}|\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^S})P(\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^S}) \\
&+ f(\varepsilon_{U,i,t}|\varepsilon_{S,i,t} \leqslant -\boldsymbol{x_{S,i,t}\gamma^S})P(\varepsilon_{S,i,t} \leqslant -\boldsymbol{x_{S,i,t}\gamma^S}) \\
=& \frac{\int_{-\boldsymbol{x_{S,i,t}\gamma^S}}^{\infty} f(\varepsilon_{C,i,t}, \varepsilon_{S,i,t})d\varepsilon_{S,i,t}}{P(\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^S})} P(\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^S}) \\
&+ \frac{\int_{-\infty}^{-\boldsymbol{x_{S,i,t}\gamma^S}} f(\varepsilon_{U,i,t}, \varepsilon_{S,i,t})d\varepsilon_{S,i,t}}{P(\varepsilon_{S,i,t} \leqslant -\boldsymbol{x_{S,i,t}\gamma^S})} P(\varepsilon_{S,i,t} \leqslant -\boldsymbol{x_{S,i,t}\gamma^S}) \\
=& \int_{-\boldsymbol{x_{S,i,t}\gamma^S}}^{\infty} f(\varepsilon_{C,i,t}, \varepsilon_{S,i,t})d\varepsilon_{S,i,t} + \int_{-\infty}^{-\boldsymbol{x_{S,i,t}\gamma^S}} f(\varepsilon_{U,i,t}, \varepsilon_{S,i,t})d\varepsilon_{S,i,t} \\
=& f(\varepsilon_{C,i,t}) \int_{-\boldsymbol{x_{S,i,t}\gamma^S}}^{\infty} f(\varepsilon_{S,i,t})d\varepsilon_{S,i,t} + f(\varepsilon_{U,i,t}) \int_{-\infty}^{-\boldsymbol{x_{S,i,t}\gamma^S}} f(\varepsilon_{S,i,t})d\varepsilon_{S,i,t} \\
=& f(\varepsilon_{C,i,t})P(\varepsilon_{S,i,t} > -\boldsymbol{x_{S,i,t}\gamma^S}) + f(\varepsilon_{U,i,t})P(\varepsilon_{S,i,t} \leqslant -\boldsymbol{x_{S,i,t}\gamma^S}) \qquad (31)
\end{aligned}$$

where $f(\varepsilon_{C,i,t}|.)$ and $f(\varepsilon_{U,i,t}|.)$ denote general conditional probability densities and $f(.)$ is the marginal density. The fourth step uses the assumption that $\varepsilon_{C,i,t}$ and $\varepsilon_{U,i,t}$ are each independent from the error term $\varepsilon_{S,i,t}$ in the selection equation.

F Robustness Checks

Table 12: Capital Investment-Cash Flow Sensitivity and Marginal Revenue Product of Capital (MRPK)

Country	ΔlnTFPR	ΔlnTFPR $* d$	$\frac{\text{CF}}{\text{FTA}}$	$\frac{\text{CF}}{\text{FTA}} * d$	$d(\text{MRPK} > p70)$	Within R^2	Observations
Bulgaria	-0.000	0.037**	-0.001	0.024***	0.380***	0.0679	62,361
	(0.0059)	(0.0186)	(0.0022)	(0.0026)	(0.0117)		
Croatia	0.003	-0.020	0.005***	0.021***	0.420***	0.0682	76,801
	(0.0094)	(0.0333)	(0.0017)	(0.0021)	(0.0106)		
Czech Republic	0.003	0.005	0.006***	0.017***	0.409***	0.0686	106,834
	(0.0086)	(0.0277)	(0.0016)	(0.0020)	(0.0086)		
Finland	0.024***	0.009	0.010***	0.005***	0.287***	0.0559	107,782
	(0.0073)	(0.0200)	(0.0015)	(0.0018)	(0.0062)		
France	0.040***	0.007	0.021***	0.005***	0.302***	0.0712	972,611
	(0.0043)	(0.0118)	(0.0004)	(0.0005)	(0.0022)		
Germany	0.057***	-0.001	0.002	0.011***	0.285***	0.0552	60,566
	(0.0151)	(0.0370)	(0.0022)	(0.0025)	(0.0098)		
Italy	0.031***	0.007	0.018***	0.009***	0.293***	0.0512	1,198,195
	(0.0028)	(0.0072)	(0.0007)	(0.0008)	(0.0021)		
Korea	0.099***	-0.065**	0.004***	0.028***	0.574***	0.1074	341,295
	(0.0088)	(0.0318)	(0.0009)	(0.0012)	(0.0058)		
Norway	0.050***	-0.080**	0.010***	0.008***	0.374***	0.0669	73,268
	(0.0131)	(0.0405)	(0.0012)	(0.0014)	(0.0096)		
Poland	0.030***	0.000	0.008***	0.008***	0.280***	0.0505	83,927
	(0.0110)	(0.0293)	(0.0021)	(0.0023)	(0.0082)		
Portugal	0.033***	-0.001	0.007***	0.012***	0.304***	0.0430	252,792
	(0.0047)	(0.0157)	(0.0013)	(0.0016)	(0.0045)		
Romania	0.016***	0.066***	0.005***	0.013***	0.381***	0.0550	336,141
	(0.0032)	(0.0096)	(0.0006)	(0.0008)	(0.0047)		
Serbia	0.020***	0.025	0.007***	0.016***	0.384***	0.0642	99,047
	(0.0054)	(0.0154)	(0.0017)	(0.0021)	(0.0084)		
Slovakia	0.010	0.067*	0.005	0.026***	0.444***	0.0655	42,936
	(0.0126)	(0.0363)	(0.0035)	(0.0044)	(0.0147)		
Slovenia	0.026*	0.069	-0.000	0.045***	0.375***	0.0646	43,656
	(0.0148)	(0.0440)	(0.0047)	(0.0061)	(0.0144)		
Spain	0.007***	0.016*	0.013***	0.016***	0.280***	0.0520	960,187
	(0.0026)	(0.0097)	(0.0007)	(0.0009)	(0.0023)		
Sweden	0.016***	0.008	0.014***	0.004***	0.308***	0.0514	183,344
	(0.0052)	(0.0115)	(0.0009)	(0.0011)	(0.0058)		
Ukraine	0.012***	0.001	-0.001	0.011***	0.359***	0.0413	185,898
	(0.0031)	(0.0082)	(0.0013)	(0.0015)	(0.0064)		
United Kingdom	0.046***	0.002	0.021***	-0.003	0.194***	0.0350	94,157
	(0.0134)	(0.0330)	(0.0034)	(0.0036)	(0.0067)		

Note: The table shows the coefficients from regressing $\Delta\text{lnFTA}_{i,t}$ on lagged productivity growth $\Delta\text{lnTFPR}_{i,t-1}$ and lagged cash flow over twice lagged fixed tangible assets $\frac{\text{CF}_{i,t-1}}{\text{FTA}_{i,t-2}}$, and each of which interacted with a dummy that equals one if lagged log MRPK is in the top 30% and zero if otherwise. The last column shows the number of firm-year observations used in each regression. Firm and four-digit industry*year fixed effects are included in all regressions. Firm-level clustered standard errors are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Proportion of Constrained Firms for Each Two-digit Industry in Each Country

Country	10	13	14	16	18	20	22	23	25	26	27	28	31	32
Bulgaria	0.34	0.42	0.47	0.46	0.52	0.32	0.40	0.42	0.39	0.43	0.40	0.36	0.48	0.53
Croatia	0.40	0.44	0.41	0.41	0.40	0.47	0.41	0.43	0.41	0.47	0.48	0.42	0.43	0.39
Czech Republic	0.32	0.30	0.45	0.41	0.48	0.31	0.34	0.29	0.41	0.48	0.43	0.35	0.36	0.45
Finland	0.26	0.27	0.22	0.33	0.25	0.18	0.22	0.25	0.23	0.24	0.26	0.24	0.22	0.22
France	0.28	0.29	0.29	0.28	0.32	0.30	0.28	0.28	0.24	0.30	0.31	0.31	0.30	0.30
Germany	0.19	0.29	0.39	0.39	0.25	0.21	0.23	0.28	0.26	0.29	0.33	0.27	0.28	0.30
Italy	0.33	0.35	0.32	0.38	0.32	0.33	0.32	0.35	0.34	0.32	0.34	0.35	0.41	0.46
Japan	0.17	0.22	*	0.40	0.20	0.16	0.14	0.22	0.25	0.24	0.24	0.22	0.26	0.26
Korea	0.30	0.37	0.63	0.42	0.47	0.30	0.30	0.29	0.36	0.26	0.10	0.48	0.47	0.19
Norway	0.32	0.32	0.49	0.34	0.37	0.32	0.35	0.31	0.34	0.39	0.41	0.37	0.36	0.33
Poland	0.25	0.25	0.39	0.26	0.37	0.30	0.29	0.26	0.27	0.34	0.26	0.34	0.30	0.31
Portugal	0.28	0.32	0.30	0.32	0.33	0.26	0.32	0.26	0.37	0.32	0.39	0.40	0.33	0.30
Romania	0.36	0.37	0.19	0.19	0.37	0.36	0.35	0.12	0.44	0.34	0.18	0.35	0.36	0.15
Serbia	0.29	0.34	0.32	0.30	0.31	0.34	0.27	0.31	0.31	0.33	0.31	0.15	0.33	0.32
Slovakia	0.30	0.37	0.44		0.41	0.30	0.37	0.37	0.40	0.51	0.45	0.35	0.45	0.40
Slovenia	0.38	0.35	0.49	0.33	0.35	0.34	0.32	0.32	0.32	0.38	0.29	0.33	0.37	0.34
Spain	0.27	0.27	0.29	0.31	0.31	0.28	0.28	0.29	0.30	0.29	0.29	0.31	0.30	0.29
Sweden	0.32	0.41	0.39	0.29	0.32	0.35	0.33	0.29	0.32	0.29	0.36	0.36	0.39	0.30
Ukraine	0.30	0.37	0.30	0.32	0.29	0.31	0.31	0.31	0.33	0.34	0.34	0.35	0.30	0.33
United Kingdom	0.27	0.22	0.24	0.27	0.27	0.28	0.24	0.33	0.24	0.46	0.29	0.28	0.28	0.30
Unlisted - Listed	0.15	0.14	0.19	0.17	0.14	0.16	0.12	0.17	0.18	0.12	0.14	0.14	0.22	0.08
Number of Countries	11	9	9	3	4	10	7	11	11	12	11	12	7	7

Note: The table summarizes the results for the proportion of constrained firms from applying the switching regression model in the baseline analysis (where MRPK is computed using nominal revenue over fixed tangible assets) to 14 different two-digit industries selected based on the number of observations. Each column summarizes the average proportion of constrained firms across all firms and years in a given industry for different countries. The last two rows show the mean difference between the proportion of constrained firms in a sample of unlisted firms and that in a sample of listed firms, and the number of countries used to calculate this mean difference. The blank cell is due to the failure of convergence in log likelihood. * indicates none of MRPK, age or size is significant in the selection equation.

Table 14: Credit Distortion for Each Two-digit Industry in Each Country

Country	10	13	14	16	18	20	22	23	25	26	27	28	31	32
Bulgaria	0.53	0.75	0.69	0.64	0.79	0.55	0.65	0.71	0.64	0.76	0.62	0.63	0.70	0.71
Croatia	0.61	0.72	0.60	0.62	0.67	0.73	0.61	0.65	0.64	0.67	0.69	0.61	0.67	0.58
Czech Republic	0.59	0.62	0.69	0.65	0.74	0.58	0.66	0.63	0.69	0.70	0.71	0.71	0.66	0.75
Finland	0.45	0.28	0.30	0.43	0.35	0.27	0.37	0.40	0.31	0.31	0.35	0.34	0.30	0.26
France	0.54	0.49	0.46	0.52	0.55	0.55	0.55	0.52	0.46	0.51	0.55	0.53	0.53	0.55
Germany	0.70	0.62	0.79	0.72	0.51	0.57	0.63	0.70	0.56	0.54	0.57	0.58	0.54	0.54
Italy	0.51	0.55	0.40	0.54	0.48	0.51	0.49	0.53	0.51	0.43	0.46	0.47	0.55	0.58
Japan	0.43	0.52	*	0.69	0.64	0.47	0.45	0.48	0.59	0.65	0.66	0.59	0.66	0.59
Korea	0.60	0.76	0.80	0.71	0.80	0.62	0.62	0.60	0.68	0.53	0.16	0.74	0.76	0.36
Norway	0.45	0.41	0.56	0.51	0.49	0.57	0.51	0.47	0.43	0.50	0.49	0.49	0.54	0.40
Poland	0.50	0.56	0.71	0.52	0.70	0.59	0.62	0.51	0.57	0.67	0.62	0.70	0.65	0.58
Portugal	0.53	0.61	0.54	0.61	0.68	0.57	0.57	0.54	0.68	0.51	0.70	0.68	0.64	0.51
Romania	0.47	0.52	0.30	0.31	0.53	0.48	0.50	0.23	0.57	0.51	0.29	0.50	0.49	0.25
Serbia	0.38	0.43	0.38	0.40	0.41	0.45	0.38	0.43	0.44	0.39	0.45	0.23	0.45	0.40
Slovakia	0.49	0.68	0.64		0.65	0.44	0.63	0.62	0.64	0.72	0.66	0.62	0.73	0.55
Slovenia	0.66	0.74	0.76	0.60	0.69	0.65	0.61	0.63	0.66	0.71	0.54	0.68	0.68	0.64
Spain	0.48	0.48	0.47	0.53	0.55	0.54	0.53	0.50	0.52	0.47	0.47	0.52	0.53	0.50
Sweden	0.53	0.62	0.56	0.50	0.50	0.53	0.59	0.54	0.56	0.40	0.55	0.56	0.67	0.43
Ukraine	0.40	0.47	0.42	0.43	0.47	0.39	0.44	0.44	0.47	0.41	0.45	0.46	0.42	0.42
United Kingdom	0.58	0.52	0.35	0.53	0.53	0.66	0.56	0.65	0.48	0.72	0.48	0.53	0.47	0.54
Unlisted - Listed	0.24	0.22	0.18	0.16	0.15	0.24	0.19	0.20	0.27	0.09	0.16	0.21	0.37	0.06
Number of Countries	11	9	9	3	4	10	7	11	11	12	11	12	7	7

Note: The table summarizes the results for credit distortion from applying the switching regression model in the baseline analysis (where MRPK is computed using nominal revenue over fixed tangible assets) to 14 different two-digit industries selected based on the number of observations. Each column summarizes the average credit distortion across all firms and years in a given industry for different countries. The last two rows show the mean difference between the credit distortion in a sample of unlisted firms and that in a sample of listed firms, and the number of countries used to calculate this mean difference. The blank cell is due to the failure of convergence in log likelihood. * indicates none of MRPK, age or size is significant in the selection equation.

Table 15: Proportion of Constrained Firms for Each Two-digit Industry in Each Country

Country	10	13	14	16	18	20	22	23	25	26	27	28	31	32
Bulgaria	0.34	0.42	0.46	0.46	0.53	0.28	0.37	0.43	0.39	0.44	0.39	0.38	0.47	0.55
Croatia	0.40	0.45	0.40	0.40	0.40	0.48	0.41	0.42	0.41	0.48	0.46	0.42	0.43	0.40
Czech Republic	0.33	0.32	0.47	0.42	0.48	0.31	0.34	0.30	0.42	0.49	0.44	0.36	0.36	0.47
Finland	0.26	0.29	0.22	0.15	0.26	0.19	0.23	0.25	0.23	0.25	0.27	0.24	0.22	0.24
France	0.28	0.30	0.29	0.28	0.32	0.72	0.29	0.28	0.25	0.30	0.31	0.31	0.30	0.30
Germany	0.20		0.39	0.41	0.24	0.22	0.23	0.29	0.27	0.31	0.25	0.26	0.27	0.29
Italy	0.33	0.35	0.31	0.38	0.31	0.33	0.32	0.35	0.34	0.32	0.34	0.35	0.39	0.43
Japan	0.24	0.23	0.18	0.39	*	0.25	0.25	0.27	0.25	0.31	0.26	0.30	0.28	0.31
Korea	0.30	0.36	0.19	0.14	0.45	0.30	0.30	0.29	0.27	0.35	0.46	0.16	0.46	0.52
Norway	0.32	0.28	0.46	0.34	0.37	0.38	0.36	0.31	0.33	0.33	0.41	0.37	0.37	0.33
Poland	0.24	0.25	0.31	0.26	0.37	0.31	0.29	0.27	0.27	0.34	0.25	0.34	0.30	0.31
Portugal	0.29	0.32	0.29	0.32	0.32	0.31	0.32	0.35	0.36	0.32	0.38	0.39	0.34	0.30
Romania	0.36	0.36	0.36	0.35	0.24	0.51	0.19	0.19	0.41	0.35	0.36	0.36	0.41	0.23
Serbia	0.29	0.35	0.32	0.31	0.31	0.35	0.28	0.30	0.31	0.33	0.31	0.28	0.33	0.32
Slovakia	0.31	0.36	0.45	0.42	0.42	0.33	0.37	0.38	0.41	0.51	0.45	0.36	0.44	0.40
Slovenia	0.38	0.35	0.49	0.33	0.36	0.32	0.32	0.33	0.32	0.38	0.30	0.34	0.37	0.46
Spain	0.27	0.27	0.29	0.30	0.31	0.28	0.28	0.29	0.30	0.29	0.28	0.31	0.30	0.30
Sweden	0.32	0.38	0.39	0.30	0.32	0.41	0.33	0.30	0.32	0.30	0.39	0.37	0.40	0.30
Ukraine	0.28	0.39	0.30	0.30	0.28	0.33	0.30	0.30	0.32	0.33	0.32	0.36	0.29	0.35
United Kingdom	0.25	0.27	0.28	0.33	0.32	0.25	0.19	0.30	0.24	0.31	0.30	0.29	0.36	0.28
Unlisted - Listed	0.14	0.17	0.19	0.16	0.13	0.18	0.08	0.16	0.18	0.09	0.16	0.16	0.24	0.11
Number of Countries	10	6	7	2	3	9	5	9	9	10	9	11	4	5

Note: The table summarizes the results for the proportion of constrained firms from applying the switching regression model (where MRPK is computed using nominal value added over fixed tangible assets) to 14 different two-digit industries selected based on the number of observations. Each column summarizes the average proportion of constrained firms across all firms and years in a given industry for different countries. The last two rows show the mean difference between the proportion of constrained firms in a sample of unlisted firms and that in a sample of listed firms, and the number of countries used to calculate this mean difference. The blank cell is due to the failure of convergence in log likelihood. * indicates none of MRPK, age or size is significant in the selection equation.

Table 16: Credit Distortion for Each Two-digit Industry in Each Country

Country	10	13	14	16	18	20	22	23	25	26	27	28	31	32
Bulgaria	0.58	0.74	0.69	0.67	0.80	0.47	0.64	0.73	0.64	0.77	0.67	0.68	0.71	0.76
Croatia	0.66	0.74	0.61	0.64	0.67	0.78	0.65	0.69	0.65	0.69	0.65	0.64	0.66	0.60
Czech Republic	0.66	0.67	0.70	0.69	0.72	0.62	0.67	0.64	0.70	0.72	0.74	0.70	0.68	0.78
Finland	0.45	0.42	0.31	0.23	0.35	0.29	0.39	0.41	0.33	0.31	0.36	0.36	0.33	0.27
France	0.56	0.53	0.49	0.54	0.56	0.89	0.57	0.56	0.48	0.53	0.58	0.55	0.55	0.55
Germany	0.72		0.82	0.73	0.48	0.59	0.64	0.68	0.56	0.52	0.55	0.56	0.51	0.54
Italy	0.58	0.57	0.41	0.58	0.49	0.55	0.52	0.57	0.53	0.45	0.49	0.50	0.55	0.57
Japan	0.47	0.53	0.43	0.69	*	0.42	0.50	0.40	0.52	0.62	0.71	0.63	0.57	0.60
Korea	0.62	0.74	0.23	0.27	0.78	0.61	0.61	0.59	0.60	0.63	0.76	0.28	0.77	0.79
Norway	0.57	0.42	0.57	0.56	0.56	0.72	0.56	0.55	0.47	0.53	0.55	0.56	0.61	0.50
Poland	0.59	0.57	0.62	0.55	0.72	0.61	0.62	0.54	0.59	0.71	0.61	0.71	0.69	0.57
Portugal	0.57	0.65	0.50	0.63	0.66	0.61	0.59	0.65	0.68	0.56	0.73	0.72	0.66	0.54
Romania	0.51	0.52	0.47	0.49	0.40	0.64	0.37	0.34	0.55	0.52	0.50	0.52	0.55	0.37
Serbia	0.41	0.45	0.38	0.42	0.45	0.51	0.40	0.39	0.44	0.42	0.46	0.42	0.49	0.42
Slovakia	0.55	0.63	0.68	0.64	0.71	0.58	0.66	0.68	0.66	0.76	0.68	0.64	0.75	0.58
Slovenia	0.71	0.77	0.76	0.64	0.71	0.67	0.68	0.66	0.70	0.72	0.57	0.71	0.71	0.78
Spain	0.53	0.52	0.47	0.56	0.54	0.56	0.55	0.52	0.54	0.49	0.48	0.55	0.53	0.52
Sweden	0.57	0.62	0.58	0.58	0.53	0.64	0.62	0.58	0.59	0.43	0.59	0.59	0.69	0.45
Ukraine	0.43	0.52	0.45	0.39	0.48	0.47	0.45	0.44	0.44	0.41	0.43	0.48	0.42	0.49
United Kingdom	0.56	0.54	0.51	0.61	0.63	0.56	0.41	0.58	0.50	0.53	0.50	0.58	0.62	0.50
Unlisted - Listed	0.21	0.29	0.26	0.22	0.18	0.21	0.09	0.26	0.29	0.12	0.24	0.24	0.45	0.19
Number of Countries	10	6	7	2	3	9	5	9	9	10	9	11	4	5

Note: The table summarizes the results for credit distortion from applying the switching regression model (where MRPK is computed using nominal value added over fixed tangible assets) to 14 different two-digit industries selected based on the number of observations. Each column summarizes the average credit distortion across all firms and years in a given industry for different countries. The last two rows show the mean difference between the credit distortion in a sample of unlisted firms and that in a sample of listed firms, and the number of countries used to calculate this mean difference. The blank cell is due to the failure of convergence in log likelihood. * indicates none of MRPK, age or size is significant in the selection equation.

Table 17: Switching Regression Model of Firm Investment in Fabricated Metal Products Industry (Without Firm Fixed Effects)

	Unconstrained Regime		Constrained Regime				
Country	ΔlnSales$_{i,t-1}$	$\frac{\text{CF}_{i,t-1}}{k_{i,t-2}}$	ΔlnSales$_{i,t-1}$	$\frac{\text{CF}_{i,t-1}}{k_{i,t-2}}$	Observations	Prob > Chi2	df
Bulgaria	0.021***	0.007	0.109***	0.017***	4,243	0.0000	69
	(0.0061)	(0.0041)	(0.0324)	(0.0041)			
Croatia	0.031***	-0.004***	0.222***	0.016***	12,652	0.0000	69
	(0.0053)	(0.0016)	(0.0241)	(0.0026)			
Czech Republic	0.035***	-0.003**	0.150***	0.021***	25,421	0.0000	89
	(0.0040)	(0.0015)	(0.0207)	(0.0022)			
Finland	0.048***	-0.012***	0.120***	0.020***	27,429	0.0000	73
	(0.0037)	(0.0007)	(0.0232)	(0.0034)			
France	0.169***	-0.002***	0.358***	0.021***	170,850	0.0000	75
	(0.0045)	(0.0005)	(0.0177)	(0.0010)			
Germany	0.084***	-0.000	0.252***	0.008***	12,100	0.0000	91
	(0.0094)	(0.0008)	(0.0404)	(0.0023)			
Italy	0.032***	-0.012***	0.233***	0.028***	246,989	0.0000	87
	(0.0014)	(0.0005)	(0.0069)	(0.0011)			
Japan	0.026***	0.043**	0.138**	0.030	6,830	0.0000	79
	(0.0078)	(0.0172)	(0.0679)	(0.0243)			
Korea	0.010***	0.000	0.178***	0.024***	55,900	0.0000	53
	(0.0017)	(0.0007)	(0.0138)	(0.0022)			
Norway	0.044***	-0.003***	0.189***	0.018***	12,676	0.0000	71
	(0.0072)	(0.0006)	(0.0373)	(0.0023)			
Poland	0.068***	0.000	0.208***	0.011***	13,237	0.0000	75
	(0.0059)	(0.0014)	(0.0344)	(0.0025)			
Portugal	0.037***	-0.009***	0.272***	0.017***	47,373	0.0000	69
	(0.0035)	(0.0012)	(0.0162)	(0.0018)			
Romania	0.031***	-0.003***	0.102***	0.016***	44,863	0.0000	75
	(0.0023)	(0.0004)	(0.0088)	(0.0012)			
Serbia	0.028***	-0.000	0.149***	0.037***	12,866	0.0000	67
	(0.0033)	(0.0012)	(0.0207)	(0.0044)			
Slovakia	0.051***	-0.013***	0.124***	0.031***	10,806	0.0000	75
	(0.0063)	(0.0018)	(0.0267)	(0.0050)			
Slovenia	0.056***	-0.010***	0.315***	0.016***	12,476	0.0000	59
	(0.0065)	(0.0014)	(0.0335)	(0.0035)			
Spain	0.049***	-0.004***	0.270***	0.027***	193,141	0.0000	77
	(0.0019)	(0.0006)	(0.0128)	(0.0014)			
Sweden	0.082***	-0.009***	0.239***	0.016***	56,662	0.0000	71
	(0.0041)	(0.0005)	(0.0176)	(0.0015)			
Ukraine	0.015***	0.000	0.144***	0.009***	20,782	0.0000	63
	(0.0022)	(0.0007)	(0.0134)	(0.0018)			
United Kingdom	0.081***	-0.006***	0.237***	0.006***	26,117	0.0000	75
	(0.0057)	(0.0003)	(0.0289)	(0.0012)			

Note: The dependent variable is firm investment ΔlnFTA$_{i,t}$. The coefficients for lagged sales growth and lagged cash flow in two different investment regimes are reported. Four-digit industry and year fixed effects are included in the switching regression. Firm fixed effects are partially controlled by adding the means of the firm-specific variables in each equation, whose coefficients are not reported here. The last two columns show the p-value for the likelihood ratio test and the degrees of freedom for the χ^2 distribution respectively. A small p-value suggests that the switching regression (less restrictive model) fits the data significantly better than an OLS regression. Robust standard errors are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

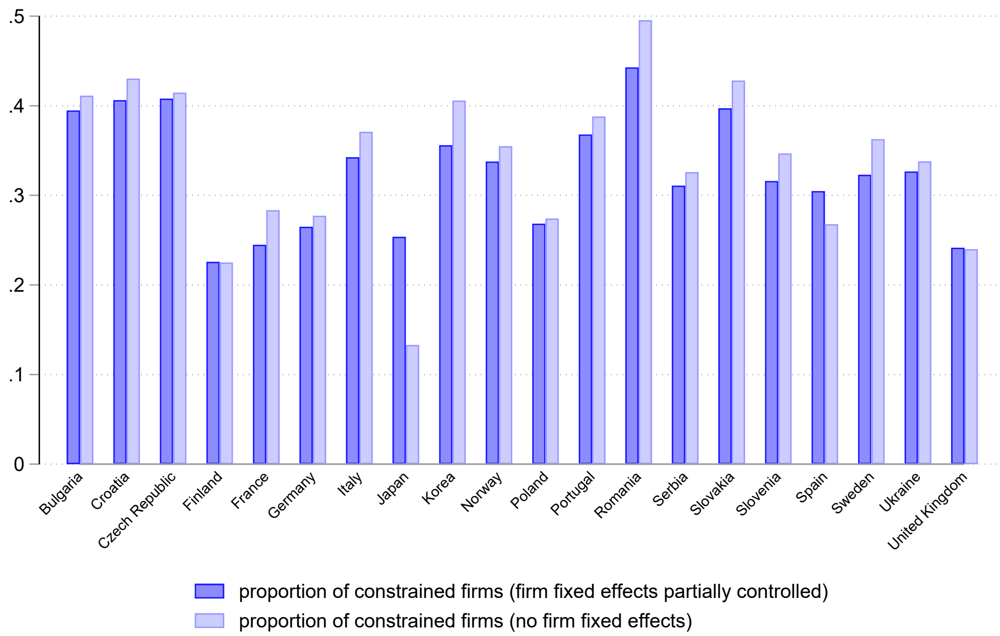
Table 18: The Selection Equation of the Switching Regression in Fabricated Metal Products Industry (Without Firm Fixed Effects)

Country	Age	ln(Assets)	ln(MRPK)	$\frac{\text{Net worth}}{\text{Assets}}$	$\frac{\text{Cash}}{\text{Assets}}$	Fraction constrained
Bulgaria	-0.027***	-0.127***	0.790***	-0.449**	1.531***	0.41
	(0.0063)	(0.0417)	(0.0602)	(0.1784)	(0.3521)	
Croatia	-0.023***	-0.214***	0.907***	-0.504***	1.588***	0.43
	(0.0054)	(0.0229)	(0.0360)	(0.1102)	(0.3344)	
Czech Republic	-0.083***	-0.287***	0.933***	-0.308***	2.109***	0.41
	(0.0046)	(0.0190)	(0.0246)	(0.0710)	(0.1555)	
Finland	-0.020***	-0.142***	0.618***	0.041	0.398***	0.23
	(0.0022)	(0.0157)	(0.0225)	(0.0720)	(0.1246)	
France	-0.014***	-0.331***	1.032***	0.196***	1.186***	0.28
	(0.0008)	(0.0094)	(0.0122)	(0.0468)	(0.0615)	
Germany	-0.006***	-0.345***	0.789***	0.222	0.957***	0.28
	(0.0013)	(0.0257)	(0.0369)	(0.1709)	(0.2922)	
Italy	-0.012***	-0.236***	0.794***	-0.002	1.050***	0.37
	(0.0007)	(0.0064)	(0.0078)	(0.0371)	(0.0663)	
Japan	-0.024***	-0.142***	0.918***	-0.025	0.879*	0.13
	(0.0042)	(0.0432)	(0.0632)	(0.2414)	(0.4683)	
Korea	-0.029***	-0.267***	0.857***	-0.074	0.148	0.41
	(0.0021)	(0.0110)	(0.0138)	(0.0616)	(0.1640)	
Norway	-0.019***	-0.147***	0.671***	-0.129	0.402**	0.35
	(0.0049)	(0.0231)	(0.0273)	(0.1263)	(0.1662)	
Poland	-0.037***	-0.190***	0.712***	-0.345***	1.469***	0.27
	(0.0067)	(0.0244)	(0.0320)	(0.0962)	(0.2373)	
Portugal	-0.034***	-0.251***	1.108***	-0.106*	0.951***	0.39
	(0.0019)	(0.0156)	(0.0224)	(0.0616)	(0.1379)	
Romania	-0.039***	-0.149***	0.650***	-0.228***	1.038***	0.50
	(0.0034)	(0.0120)	(0.0181)	(0.0336)	(0.1250)	
Serbia	-0.010**	-0.059***	0.540***	-0.579***	1.758***	0.33
	(0.0042)	(0.0192)	(0.0276)	(0.1070)	(0.3700)	
Slovakia	-0.073***	-0.453***	0.858***	-0.381***	1.534***	0.43
	(0.0075)	(0.0294)	(0.0393)	(0.0943)	(0.2681)	
Slovenia	-0.043***	-0.333***	0.984***	-0.449***	1.930***	0.35
	(0.0054)	(0.0267)	(0.0393)	(0.1285)	(0.3262)	
Spain	-0.023***	-0.091***	0.887***	-0.044	0.740***	0.27
	(0.0011)	(0.0072)	(0.0089)	(0.0298)	(0.0634)	
Sweden	-0.012***	-0.320***	0.890***	0.310***	0.889***	0.36
	(0.0013)	(0.0131)	(0.0166)	(0.0743)	(0.0980)	
Ukraine	-0.016***	0.039***	0.494***	-0.246***	0.289	0.34
	(0.0034)	(0.0115)	(0.0165)	(0.0398)	(0.1909)	
United Kingdom	-0.004***	-0.113***	0.696***	-0.345***	-0.160	0.24
	(0.0012)	(0.0130)	(0.0241)	(0.0697)	(0.1396)	

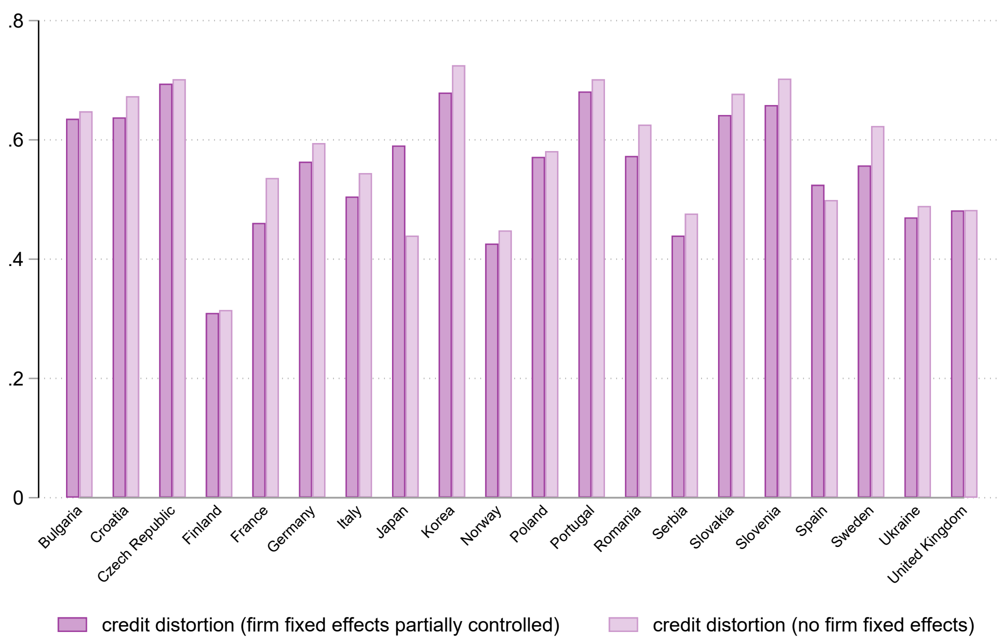
Note: The table shows the coefficients for the key variables in the selection equation that determines the probability of a firm being constrained, including age, log of assets, log of MRPK, net worth-to assets ratio, and cash-to-assets ratio, and the average proportion of constrained firms over the sample period. All variables apart from age are lagged. Four-digit industry and year fixed effects are included. The last column shows the average proportion of constrained firms over the sample period, where firms are classified as constrained based on the estimated posterior probabilities. Robust standard errors are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure 4: Proportion of Constrained Firms and Credit Distortion in Fabricated Metal Products Industry

(a) Proportion of Constrained Firms Using Different Fixed Effects



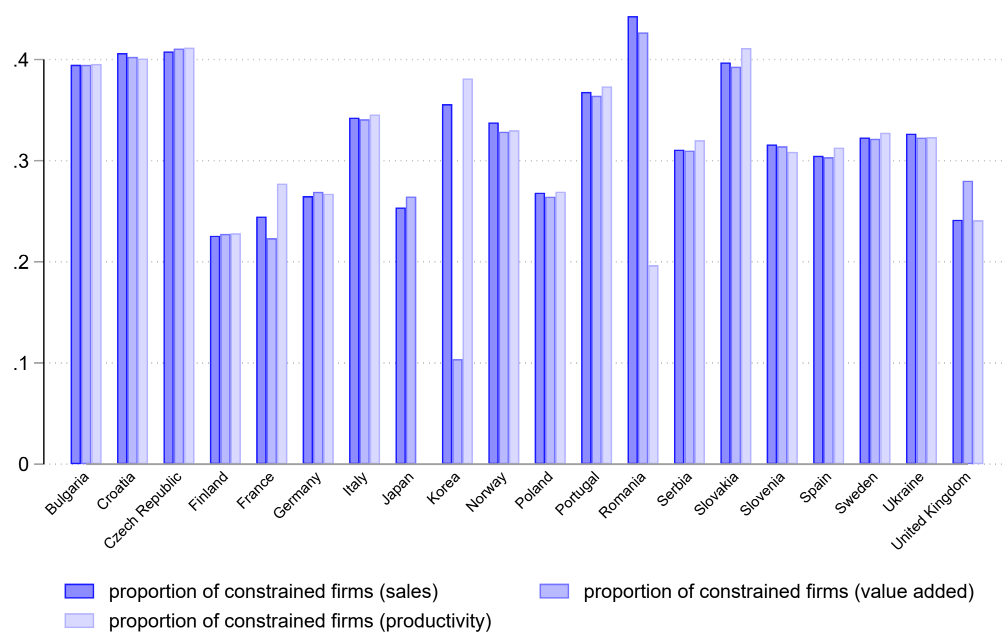
(b) Credit Distortion Using Different Fixed Effects



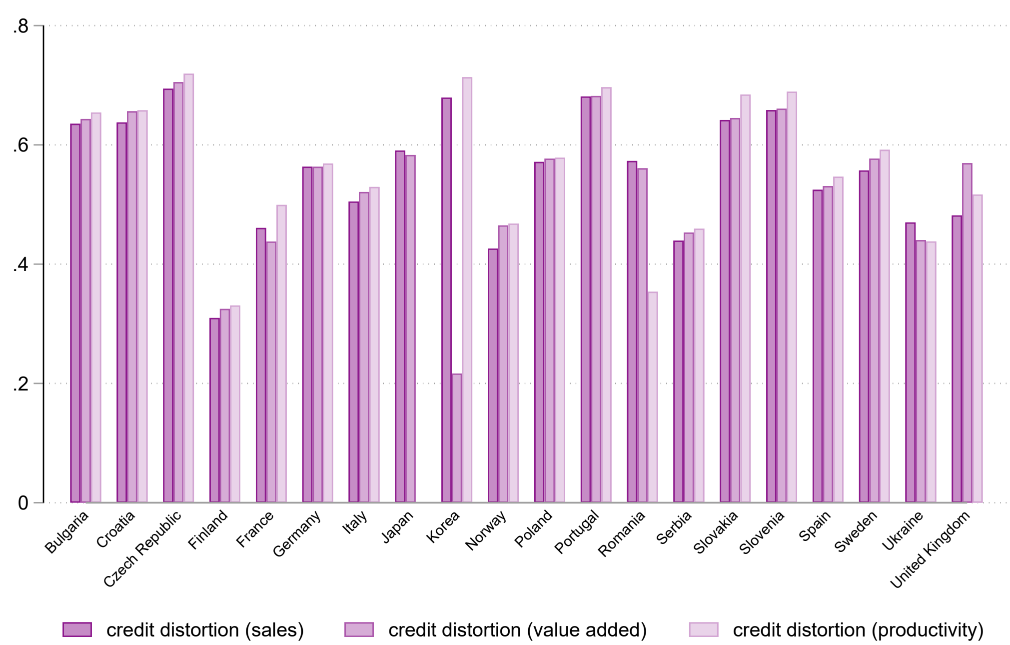
Note: In each graph, the corresponding measure is computed across all firms and years using the results from controlling for different fixed effects in the switching regression model: firm fixed effects partially controlled using the Hu and Schiantarelli (1998) approach and neglecting firm fixed effects. In both cases, four digit industry and year fixed effects are controlled. Graph (a) plots the fraction of constrained firms in industry 25 (manufacture of fabricated metal products) by NACE Rev.2 Code across 20 countries. Graph (b) plots credit distortion in percent points (i.e., the fraction of the observed dispersion (cross-section variance) of MRPK that is caused by the presence of constrained firms) in industry 25, which is computed based on (19). MRPK is computed as the nominal revenue divided by fixed tangible assets.
Data source: Orbis

Figure 5: Proportion of Constrained Firms and Credit Distortion in Fabricated Metal Products Industry

(a) Proportion of Constrained Firms Using Different Proxies



(b) Credit Distortion Using Different Proxies



Note: In each graph, the corresponding measure is computed across all firms and years using the results from applying three different proxies for investment opportunity in the switching regression model: lagged sales growth, lagged value added growth and lagged productivity growth. Graph (a) plots the fraction of constrained firms in industry 25 (manufacture of fabricated metal products) by NACE Rev.2 Code across 20 countries. Graph (b) plots credit distortion in percent points (i.e., the fraction of the observed dispersion (cross-section variance) of MRPK that is caused by the presence of constrained firms) in industry 25, which is computed based on (19). MRPK is computed as the nominal revenue divided by fixed tangible assets. Data source: Orbis

References

Ackerberg, Daniel, Kevin Caves, and Garth Frazer. 2006. "Structural Identification of Production Functions." Mimeo, UCLA Department of Economics.

Ağca, Şenay, and Abon Mozumdar. 2008. "The Impact of Capital Market Imperfections on Investment-cash Flow Sensitivity." *Journal of Banking & Finance*, 32(2): 207–216.

Almeida, Heitor, and Murillo Campello. 2007. "Financial Constraints, Asset Tangibility, and Corporate Investment." *The Review of Financial Studies*, 20(5): 1429–1460.

Alti, Aydoğan. 2003. "How Sensitive Is Investment to Cash Flow When Financing Is Frictionless?" *The Journal of Finance*, 58(2): 707–722.

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2014. "Dynamic Inputs and Resource (mis) Allocation." *Journal of Political Economy*, 122(5): 1013–1063.

Bai, Yan, Dan Lu, and Xu Tian. 2018. "Do Financial Frictions Explain Chinese Firms' Saving and Misallocation?" NBER Working Paper, No. 24436.

Banerjee, Abhijit V., and Esther Duflo. 2005. "Growth Theory through the Lens of Development Economics." *Handbook of Economic Growth*, 1: 473–552.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. 2013. "Cross-country Differences in Productivity: The Role of Allocation and Selection." *The American Economic Review*, 103(1): 305–334.

Beck, Thorsten, Asli Demirgüç-Kunt, and Vojislav Maksimovic. 2005. "Financial and Legal Constraints to Growth: Does Firm Size Matter?" *The Journal of Finance*, 60(1): 137–177.

Busso, Matías, Lucía Madrigal, and Carmen Pagés. 2013. "Productivity and Resource Misallocation in Latin America." *The Be Journal of Macroeconomics*, 13(1): 903–932.

Carpenter, Robert E, and Alessandra Guariglia. 2008. "Cash Flow, Investment, and Investment Opportunities: New Tests Using Uk Panel Data." *Journal of Banking & Finance*, 32(9): 1894–1906.

Chamberlain, Gary. 1980. "Analysis of Covariance with Qualitative Data." *The Review of Economic Studies*, 47(1): 225–238.

David, Joel M, and Venky Venkateswaran. 2017. "The Sources of Capital Misallocation." NBER Working Paper, No 23129.

Fazzari, Steven M., R. Glenn Hubbard, Bruce C. Petersen, Alan S. Blinder, and James M. Poterba. 1988. "Financing Constraints and Corporate Investment." *Brookings Papers on Economic Activity*, 1988(1): 141–206.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98(1): 394–425.

Garcia, Rene, Annamaria Lusardi, and Serena Ng. 1997. "Excess Sensitivity and Asymmetries in Consumption: An Empirical Investigation." *Journal of Money, Credit, and Banking*, 29(2): 154–176.

Gilchrist, Simon, Jae W. Sim, and Egon Zakrajšek. 2013. "Misallocation and Financial Market Frictions: Some Direct Evidence from the Dispersion in Borrowing Cost." *Review of Economic Dynamics*, 16(1): 159–176.

Gopinath, Gita, Şebnem Kalemli-Özcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez. 2017. "Capital Allocation and Productivity in South Europe." *The Quarterly Journal of Economics*, 132(4): 1915–1967.

Greene, William. 2004. "The Behaviour of the Maximum Likelihood Estimator of Limited Dependent Variable Models in the Presence of Fixed Effects." *The Econometrics Journal*, 7(1): 98–119.

Hadlock, Charles J., and Joshua R. Pierce. 2010. "New Evidence on Measuring Financial Constraints: Moving beyond the KZ Index." *The Review of Financial Studies*, 23(5): 1909–1940.

Hovakimian, Gayane, and Sheridan Titman. 2006. "Corporate Investment with Financial Constraints: Sensitivity of Investment to Funds from Voluntary Asset Sales." *Journal of Money, Credit, and Banking*, 38(2): 357–374.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *The Quarterly Journal of Economics*, 124(4): 1403–1448.

Hubbard, R Glenn, Anil K. Kashyap, and Toni M. Whited. 1995. "Internal Finance and Firm Investment." *Journal of Money, Credit and Banking*, 27(3): 683–701.

Hu, Xiaoqiang, and Fabio Schiantarelli. 1998. "Investment and Capital Market Imperfections: A Switching Regression Approach Using Us Firm Panel Data." *Review of Economics and Statistics*, 80(3): 466–479.

Kalemli-Ozcan, Sebnem, Bent Sorensen, Carolina Villegas-Sanchez, Vadym Volosovych, and Sevcan Yesiltas. 2015. "How to Construct Nationally Representative Firm Level Data from the Orbis Global Database." NBER Working Paper No. 21558.

Kaplan, Steven N., and Luigi Zingales. 1997. "Do Investment-cash Flow Sensitivities Provide Useful Measures of Financing Constraints?" *The Quarterly Journal of Economics*, 112(1): 169–215.

Kiyotaki, Nobuhiro, and John Moore. 1997. "Credit Cycles." *Journal of Political Economy*, 105(2): 211–248.

Lamont, Owen, Christopher Polk, and Jesús Saaá-Requejo. 2001. "Financial Constraints and Stock Returns." *The Review of Financial Studies*, 14(2): 529–554.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *The Review of Economic Studies*, 70(2): 317–341.

Maddala, Gangadharrao S. 1986. "Disequilibrium, Self-selection, and Switching Models." *Handbook of Econometrics*, 3: 1633–1688.

Midrigan, Virgiliu, and Daniel Yi Xu. 2014. "Finance and Misallocation: Evidence from Plant-level Data." *The American Economic Review*, 104(2): 422–458.

Moshiriana, Fariborz, Vikram Nandab, Alexander Vadilyevc, and Bohui Zhanga. 2017. "What Drives Investment–cash Flow Sensitivity around the World? an Asset Tangibility Perspective." *Journal of Banking and Finance*, 77: 1–17.

Mulier, Klaas, Koen Schoors, and Bruno Merlevede. 2016. "Investment-cash Flow Sensitivity and Financial Constraints: Evidence from Unquoted European Smes." *Journal of Banking and Finance*, 73: 182–197.

Neyman, Jerzy, and Elizabeth L Scott. 1948. "Consistent Estimates Based on Partially Consistent Observations." *Econometrica*, 16(1): 1–32.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64(6): 1263–97.

Petrin, Amil, Brian Poi, and James Levinsohn. 2004. "Production Function Estimation in Stata Using Inputs to Control for Unobservables." *Stata Journal*, , (4): 113–123.

Restuccia, Diego, and Richard Rogerson. 2008. "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments." *Review of Economic Dynamics*, 11(4): 707–720.

Restuccia, Diego, and Richard Rogerson. 2013. "Misallocation and Productivity." *Review of Economic Dynamics*, 16(1): 1–10.

Restuccia, Diego, and Richard Rogerson. 2017. "The Causes and Costs of Misallocation." *Journal of Economic Perspectives*, 31(3): 151–74.

Rovigatti, Gabriele, and Vincenzo Mollisi. 2016. "PRODEST: Stata Module for Production Function Estimation Based on the Control Function Approach." Statistical Software Components S458239, Boston College Department of Economics.

Schiantarelli, Fabio. 1995. "Financial Constraints and Investment: A Critical Review of Methodological Issues and International Evidence." Boston College Working Paper, No. 293.

Whited, Toni M., and Guojun Wu. 2006. "Financial Constraints Risk." *The Review of Financial Studies*, 19(2): 531–559.

Wooldridge, Jeffrey M. 2009. "On Estimating Firm-level Production Functions Using Proxy Variables to Control for Unobservables." *Economics Letters*, 104(3): 112–114.

Wu, Guiying Laura. 2018. "Capital Misallocation in China: Financial Frictions or Policy Distortions?" *Journal of Development Economics*, 130: 203–223.