Universidade Federal de Pernambuco
Centro de Ciências Sociais Aplicadas
Departamento de Economia

Programa de Pós-Graduação em Economia

Risk, productivity and misallocation in Brazilian listed companies

Paulo Silvio Mourão Veras Filho

Dissertação de Mestrado

Recife
2021

Universidade Federal de Pernambuco
Centro de Ciências Sociais Aplicadas
Departamento de Economia

Paulo Silvio Mourão Veras Filho

Risk, productivity and misallocation in Brazilian listed companies

Dissertação apresentada ao Programa de Pós-Graduação em Economia do Departamento de Economia da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Mestre em Economia.

Orientador: Prof. Dr. Paulo Henrique Pereira de Meneses Vaz

Recife
2021

Catalogação na Fonte
Bibliotecária Ângela de Fátima Correia Simões, CRB4-773

V476r Veras Filho, Paulo Silvio Mourão
Risk, productivity and misallocation in Brazilian listed companies / Paulo Silvio Mourão Veras Filho. - 2021.
42 folhas: il. 30 cm.

Orientador: Prof. Dr. Paulo Henrique Pereira de Meneses Vaz
Dissertação (Mestrado em Economia) – Universidade Federal de Pernambuco, CCSA, 2021.
Inclui referências e apêndices.

1. Produtividade. 2. Países em desenvolvimento. 3. Ativos financeiros. I. Vaz, Paulo Henrique Pereira de Meneses (Orientador). II. Título.

336 CDD (22. ed.) UFPE (CSA 2021 – 056)

PAULO SILVIO MOURÃO VERAS FILHO

Risk, productivity and misallocation in Brazilian listed companies

Dissertação apresentada ao Programa de Pós-Graduação em Economia do Departamento de Economia da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Mestre em Economia.

Aprovada em: 23/03/2021.

BANCA EXAMINADORA

Prof. Dr. Paulo Henrique Pereira de Meneses Vaz
Orientador

Prof. Dr. Rafael da Silva Vasconcelos
Examinador Interno

Prof. Dr. Ricardo Dias de Oliveira Brito
Examinador Externo

Resumo

Este trabalho investiga como a exposição das firmas a riscos sistemáticos de investimento e os retornos de suas ações interagem com características tais como produtividade e a má-alocação de capital no contexto de um país em desenvolvimento. Para tanto, foram adaptadas e avaliadas recentes contribuições teóricas das literaturas de finanças empíricas e "*misallocation*" à luz dos dados brasileiros. Foram coletados dados dos demonstrativos financeiros das companhias listadas na bolsa Brasileira, bem como dos retornos de suas ações e dos fatores de risco de mercado a fim de se estimar medidas de exposição ao risco (betas), produtividade total dos fatores (PTF) e má alocação de capital (representada aqui pela dispersão nos produtos marginais do capital (MPK)). Em seguida, foi executada uma série de exercícios empíricos. Os resultados sugerem que estas características das firmas preveem racionalmente os retornos das ações brasileiras. Firmas de alta produtividade apresentam um prêmio significativo sobre firmas de baixa produtividade. Além disso, firmas de alto MPK tendem a oferecer retornos contemporâneos e futuros mais elevados, o que sugere que diferenças de MPK refletem a exposição a fatores de risco para as quais os investidores demandam compensação na forma de maiores taxas de retorno. Ainda que preliminares, estes resultados indicam novas possíveis causas para o hiato de produtividade entre países ricos e pobres, dada a volatilidade relativamente mais alta do ciclo de negócios em países em desenvolvimento como o Brasil e o impacto disso nas relações estudadas aqui.

JEL Classifications: D25, E32, G12, O47.

Palavras-chave: Produtividade. Má-alocação de capital. Precificação de ativos. Países em desenvolvimento.

Abstract

This work investigates how firm-level risk exposure to aggregate conditions and stock returns interact with characteristics such as productivity and capital misallocation in the context of a developing country. To do so, we adapt recent theoretical contributions from empirical finance and misallocation literatures and we extend and evaluate their analysis by focusing on Brazil. We collect balance sheet data from Brazilian publicly listed companies, stock market returns and risk factors to estimate measures of firm-level risk exposure (betas), total factor productivity (TFP) and capital misallocation (dispersion in marginal products of capital (MPK)), then we proceed with a series of empirical exercises. Our results suggest that these firm characteristics rationally predict returns of Brazilian stocks. High productivity firms earn a significant premium over low productivity firms. Also, high MPK firms tend to offer high stock returns, both in a realized and an expected sense, suggesting that MPK differences reflect exposure to risk factors for which investors demand compensation in the form of a higher rate of return. Although preliminary, these results shed light on potential causes for the productivity gap between rich and poor countries, given the relatively higher volatility faced by developing countries like Brazil and its impact on the relations studied here.

JEL Classifications: D25, E32, G12, O47.

Keywords: Productivity. Misallocation. Cross-section of returns. Developing countries.

List of Figures

List of Tables

Contents

1 Introduction

The empirical finance literature has documented that many firm characteristics, such as size, book-to-market ratio, investment, and hiring behavior of firms, forecast future stock returns. Recent works introduced neoclassical investment models to explain such differences in firm characteristics and returns. For example, Kogan and Papanikolau (2013) and Belo et al. (2014) show how ex ante identical firms are exposed to firm-level total factor productivity (TFP) shocks, which lead to different firm characteristics. Cochrane (1991) and Balvers et al. (2015) show that stock returns and investment returns are closely linked and works like Gomes et al. (2006) and Zhang (2017) interpret common risk factors through firms' investment policies and show that investment-based factors are priced in the cross-section of returns. In this line, İmrohoroğlu and Tüzel (2014) provide evidence about the link between firm-level total factor productivity (TFP) and stock returns. The key implication of these studies is that financial market considerations can have sizable effects on real outcomes by affecting capital allocation decisions.

On the other hand, what drives productivity gaps between rich and poor countries is one of the fundamental questions in economics. The large misallocation literature argues that low income countries are not as effective in allocating their factors of production to their most efficient use (Restuccia and Rodgerson, 2017). In fact, since Hsieh and Klenow (2009) there is a growing body of evidence on how the dispersion in the marginal product of inputs across firms represents adverse effects on productivity and output. Recent studies evaluated a host of candidates for causing these dispersions, but a significant portion of observed misallocation seems to come from other (and possibly unexplored) firm-specific factors.

With these two aspects in mind, David, Schmid and Zeke (2020) propose a new theory linking capital misallocation to systematic investment risks. According to the authors, firms differ on their degree of exposure to systematic investment risks for a number of reasons, such as heterogeneous technologies/markups or heterogeneous demand sensitivities. Thus, firms with higher exposure to the aggregate risk factors require a higher risk premium on investments, which translates into a higher expected marginal product of capital (MPK). And this firm-specific risk premium appears exactly as what would otherwise be labeled a persistent distortion or "wedge" in the firm's investment decision. This channel implies that, by inducing MPK dispersion, cross-sectional variation in factor risk exposures and a higher price of risk (which depends on the degree of aggregate volatility) reduce the long-run (average) level of achieved TFP.

In this paper, we investigate how the standard notions of the risk-return tradeoff relate to productivity and the allocative efficiency of firms in the universe of public listed companies from a developing country. To the best of our knowledge, we are the first to study this connection using Brazilian data. Although preliminary, the results established here open the way for a specific research agenda focusing on applying these recent theoretical developments to the context of developing countries. This is specially important for those countries because of their particular characteristics, such as the higher volatility in their business cycles and the potential higher dispersion in risk premia arising from this, as suggested by Figure 1. Additionally, those countries are going through a process of "financial deepening", facing an increasing importance of stock markets for their economies. Figure 2 illustrates this fact, showing the increase of total value of yearly stocks traded as a percentage of Brazilian GDP.

In order to do this analysis, we reviewed two similar models linking productivity, misallocation, risk and returns and we constructed two panels with balance sheet data of Brazilian listed companies and returns for stocks and risk factors. We then estimated measures of firm-level risk exposure, firm-level productivity and misallocation degree for this sample. Finally, we investigated the hypotheses derived from the models in a set of empirical exercises including a characteristic-sorted portfolio analysis, a widespread method in empirical finance.

As main results, we find that both firm-level TFP and MPK rationally predict stock returns. TFP is positively correlated to contemporaneous stock returns, but negatively correlated to future returns. The difference between the contemporaneous returns of high and low TFP firms is 6.27%, while low productivity firms on average earn a 4% annual premium over high productivity firms in the following year. For MPK, although results are not perfectly monotonic, firms with high MPK tend to offer high stock returns both in a realized and an expected sense, suggesting that MPK differences reflect exposure to risk factors for which investors demand compensation in the form of a higher rate of return. We also find that measures of risk exposures have significant explanatory power for subsequent MPK and that the price of risk in Brazil (proxied by the aggregate market price/dividend ratio) is negatively correlated to MPK dispersion.

1.1 Related literature

This paper relates to several branches of literature. First, there is the large body of work in empirical finance relating firm's characteristics to stock returns since the pioneer work of Fama and French (1992). For example, Kogan and Papanikolau (2013), Novy-Marx (2013) and Belo, Lin and Bazdresch (2014). Cochrane (1991) and Balvers et al. (2015) show that stock returns and investment returns are closely linked. Recent works interpret common risk factors through firms' investment policies and show that investment-based factors are priced in the cross-section of returns, e. g., Gomes et al. (2006) and Zhang (2017).

In addition, by examining the link proposed by David, Schmid and Zeke (2020) between misallocation and firm-level risk exposure to aggregate risk, this work relates to the literature investigating resource misallocation, pioneered by Hsieh and Klenow (2009) and Restuccia and Rogerson (2008). Recent papers explored the role of financial frictions, for example, Moll (2014) and Buera et al. (2011). Particularly, we explore the implications of a different dimension of financial markets for marginal product dispersion, namely, the risk-return tradeoff faced by risk-averse agents.

Finally, by focusing on Brazil, this work also relates to Vasconcelos (2017) in his pioneer study of misallocation of resources in the Brazilian manufacturing sector and to numerous works on empirical finance that studied the relation between firm characteristics and stock returns in Brazil, such as Blank et al. (2014), Varga and Brito (2016) and Garcia and Santos (2018).

1.2 Work organization

This paper is organized as follows: Section 2 describes the data used in the empirical exercises. Section 3 describes the procedure to estimate the measures of aggregate risk exposure (betas) and Section 4 describes the characteristic-sorted portfolio approach used in the next sections. Section 5 describes a simple model relating firm-level productivity, firm characteristics and stock returns, estimates firm-level TFP and contains the first set of empirical exercises involving the productivity estimations. Section 6 describes our misallocation measures, develop the previous model to explain the link between MPK and stock returns and proceeds with the second set of empirical exercises involving the measures of misallocation. Finally, Section 7 concludes.

2 Data

The major difficulty of this work was the lack of access to private sources of financial market data for Brazil. We had to rely on public sources and, although our sources are detailed and accurate, the time span of the panel is relatively short and we could not use some other technical indicators and firms' characteristics typically made available by private sources. For balance sheet data of Brazilian listed companies, we used the standardized financial statements (*Demonstrações Financeiras Padronizadas*) of companies publicly listed in B3 – *Brasil, Bolsa, Balcão*, the Brazilian stock exchange, made publicly available by the *Comissão de Valores Mobiliários* (CVM), the securities market authority in Brazil. The panel was constructed using financial statements from 2010 to 2019 because of a discontinuity on the CVM's update policy in 2009.

The literature on productivity and misallocation typically uses value added or revenue from sales to measure output and fixed assets, total book assets or estimates using the perpetual inventory method to obtain a measure of capital stock; labor is usually measured by the number of employees, wage bill or labor compensation, investment, when available, is measured by Capex and the intermediate inputs (used to apply some of the control function approaches of productivity estimation) are such as intermediate materials and energy consumption in production. In our panel, we used fixed assets to measure capital as the depreciated value of plant, property, and equipment, although results were similar when using total book assets, which is a broader measure of capital stock. Value added was used to measure output, but the results remain similar when we use revenue from sales[1]. Wage bill was used to measure the labor factor, Capex was used as a measure of investment and materials and energy were used to measure intermediate inputs. The choices used in results reported here also took into account the consistency and accuracy of the data, such as the number of missings and outliers. Finally, we winsorized all variables at 1% and 99% levels to mitigate the impact of extreme observations.

We used Python to scrap the data from CVM's website[2] and to format the panel. Another difficulty was that part of the balance sheet accounts' codes and names was not perfectly standardized between companies. For those cases we used very similar terms (which we call "synonyms"). Table 10 shows the exact portuguese synonyms we used for each account in case of replications. Once we constructed the panel, all values were adjusted for inflation using the *Índice Nacional de Preço ao Consumidor Amplo* (IPCA), the Brazilian government official price index. Excluding financial companies yields a sample with around 300 firms by year for the administrative data panel. These firms are classified by B3 into 11 sectors and 45 subsectors, described in Table 11. We used these 45 subsectors as industries to measure misallocation in Section 6. Table 1 presents some descriptive statistics of the sample. In general, we can see that, on average, the cost of capital is greater than the labor expenses. However, labor expenses increased more than 50% during the time covered by our panel, much more than capital cost, which increased approximately 20%.

The second panel we constructed contains the stocks and Fama-French risk factors returns. Stock market data was obtained through the Python library `yfinance`, which provides data from Yahoo Finance®. Besides excluding financial companies, we followed the eligibility criteria typically used in national empirical finance works[3]: for companies with different classes of shares (ON, PN or UNIT) we used only the most liquid (highest traded volume) class; we excluded stocks with less than R$ 500,000.00 in daily volume of trade and that were not traded in at least 80% of the days in year $t-1$. This process reduced this second panel considerably, leaving an average of 132 stocks that met these criteria. Figure 3 shows the dynamics of the three Fama-French risk factors and the PD ratio. The three factors behaved somewhat similar, with a sharp decline near the year of 2016 and, as expected, PD ratio presents cyclical behavior.

For risk factors and and the price of risk, we used data from the Brazilian Center for Research in Financial Economics (NEFIN)[4]. We collected data on the three risk factors included in the Fama-French three factors model described in Section 3, which are the aggregate market return (MKT), defined as the value-weighted

[1] Although the majority of works studying misallocation use value added to measure output, recent works like David, Schmid and Zeke (2020) and Gorodnichenko et al. (2020) used revenue from sales instead.
[2] http://dados.cvm.gov.br/dados/CIA_ABERTA/DOC/DFP/DADOS/
[3] For example, NEFIN uses these criteria: http://nefin.com.br/Metodologia/Methodology.pdf.
[4] http://nefin.com.br/

daily return of the market portfolio (using all the eligible stocks that met the criteria mentioned above), the Small Minus Big Factor (SMB), which is the return of a portfolio long on stocks with low market capitalization (small) and short on stocks with high market capitalization (big), and the High Minus Low Factor (HML), that is the return of a portfolio long on stocks with high book-to-market ratio and short on stocks with low book-to-market ratio. We used the aggregate market price/dividend (PD) Ratio as a proxy for the price of risk. It consists of the dividend yield for the Brazilian stock market as the ratio of total dividend payments in the last 12 months and the market value of equity. Finally, as we work with excess returns, we also used the risk-free rate from the 30-day DI Swap to subtract returns.

3 Betas and expected returns

Following the analysis of David, Schmid and Zeke (2020), we estimate stock market betas by performing time-series regressions of firm-level excess returns (realized returns in excess of the risk-free rate), r^e_{it}, on aggregate factors, denoted by the $N \times 1$ vector $\mathbf{F}_t$. For each firm, the specification takes the form

$$r^e_{it} = \alpha_{i\tau} + \beta_{i\tau}\mathbf{F}_t + \epsilon_{it}. \tag{1}$$

We estimate these regressions at the quarterly frequency using backwards-looking five-year rolling windows, i.e., for $\tau \in \{\tau - N_t + 1, \tau - \tau_T + 2, \ldots, \tau\}$, where $\beta_{i\tau}$ denotes the $1 \times N$ vector of factor loadings and N_τ the length of the window. Under the CAPM, the single risk factor is the aggregate market return. Under the Fama-French 3 factor model, the risk factors are the market return (MKT), the return on a portfolio that is long in small firms and short in large ones (SMB) and the return on a portfolio that is long in high book-to-market firms and short on low ones (HML).

To obtain a single measure of risk exposure from the multi-factor Fama-French model, we combine the betas into a single value using estimated prices of risk from Fama and MacBeth (1973) regressions. Specifically, we estimate the following cross-sectional regression in each period:

$$r^e_{it} = \alpha_t + \lambda_t\beta_{it} + \epsilon_{it} \tag{2}$$

where λ_t denotes the $1 \times N$ vector of period t factor risk prices and β_{it} the $N \times 1$ vector of exposures, estimated as just described. We then calculate a single index of exposure to these factors as

$$\beta_{it,FF} = \lambda\beta_{it} = \sum_x \lambda_x\beta_{it,x}, \;\; x \in \{MKT, HML, SMB\},$$

where $\lambda_{\mathbf{x}} = \frac{1}{T}\sum_{t=1}^{T}\lambda_{\mathbf{x}t}$. Finally, expected stock returns are given by $\alpha_i + \lambda\beta_{it}$, where $\alpha_i = \frac{1}{T}\sum_{t=1}^{T}(\alpha_{it} + \varepsilon_{it})$.

4 Characteristic sorted portfolios

Portfolio sorting is an widespread and important tool of modern empirical finance, especially after Fama and French (1992) work on the cross-section of expected returns. It has been used to test fundamental theories in asset pricing, to establish a number of different pricing anomalies, and to identify profitable investment strategies. Cattaneo et al. (2020) provide a formalization of this methodology, which they cast as a nonparametric estimator. The premise behind portfolio sorting is to discover whether expected returns of an asset are related to a certain characteristic. A natural, and popular, way to investigate this is to sort observed returns by the characteristic value, divide the assets into portfolios according to the characteristic, and then compare differences in average returns across the portfolios.

To begin, suppose we observe both the return, R, and value of a single continuous characteristic, z, for n assets over T time periods, that are related through a regression-type model of the form

$$R_{it} = \mu(z_{it}) + \varepsilon_{it}, \qquad i = 1, \ldots, n \text{ and } t = 1, \ldots, T. \tag{3}$$

Here $\mu(.)$ is the unknown object of interest that dictates how expected returns vary with the characteristic, and is assumed to be twice continuously differentiable. We first form portfolios by partitioning the support of z into quantile-spaced bins. For each periot t, it is common practice to form J disjoint portfolios, denoted by P_{jt}, as follows $P_{jt} = [z_{(\lfloor n(j-1)/J \rfloor)t}, z_{(\lfloor nj/J \rfloor)t})$ if $J = 1, \ldots, J-1$, and $P_{jt} = [z_{(\lfloor n(J-1)/J \rfloor)t}, z_{(n)t}]$, where $z_{(\ell)t}$ denotes the ℓ-th order statistic of the sample of characteristics $\{z_{it} : 1 \leq i \leq n\}$ at each time period $t = 1, 2, \ldots, T$, and $\lfloor . \rfloor$ is the floor operator. In other words, each portfolio is a random interval containing roughly $(100/J)$-percent of the observations at each moment in time. With the portfolios thus formed, we estimate $\mu(z_*)$ at some fixed point z_* with the average returns within the portfolio containing z_*. If we let P^*_{jt} represent the appropriate portfolio at each time t, then the basic portfolio-sorted estimate is

$$\hat{\mu}(z_*) = \frac{1}{T}\sum_{t=1}^{T}\hat{\mu}_t(z_*), \quad \text{with} \quad \hat{\mu}_t(z_*) = \frac{1}{N^*_{jt}}\sum_{i:z_{it}\in P^*_{jt}} R_{it}, \tag{4}$$

where N^*_{jt} is the number of assets in P^*_{jt} at time t.

In this work, we use this approach to analyze the two specific firm characteristics we hypothesize that are related to stock returns according to the recent works of David, Schmid and Zeke (2020) and İmrohoroğlu and Tüzel (2014), that is, firm-level TFP and marginal product of capital (MPK).

5 Productivity analysis

In this section we investigate a channel relating firm-level productivity to expected returns. To do so, we describe the theoretical framework, then we estimate firm-level productivity and proceed with some empirical exercises. This section is mainly based on the work of İmrohoroğlu and Tüzel (2014).

5.1 Theoretical framework

According to Kogan and Papanikolau (2013), the literature on expected stock returns and firm characteristics considers several sources of firm heterogeneity. Many of these models assume that all firms have identical long-run properties but differ from each other at each time point because of the firm-specific productivity shocks. Other models focus on the structural differences between firms, emphasizing, for instance, persistent cross-sectional differences in the firms' technologies. Here we develop a standard production-based asset pricing model where firms are subject to both aggregate and idiosyncratic productivity shocks to account for the cross-sectional relationship between TFP, firm-level characteristics, and stock returns.

We begin with many firms that produce a homogeneous good using capital and labor. They are subject to different productivity shocks and produce according to

$$Y_{it} = F(X_t, Z_{it}, K_{it}, L_{it}) = X_t Z_{it} K_{it}^{\alpha} L_{it}^{\alpha-1}, \text{ with } \theta_1 + \theta_2 < 1, \tag{5}$$

where K_{it} denotes the beginning of period t capital stock of firm i, and L_{it} denotes the labor used in production by firm i during period t. θ_1 and θ_2 are capital and labor shares, respectively. Aggregate productivity is denoted by $x_t = \log(X_t)$ and has a stationary and monotone Markov transition function, given by $p_x(x_{t+1}|x_t)$ as follows:

$$x_{t+1} = \rho_x x_t + \varepsilon_{t+1}^x, \qquad \varepsilon_{t+1}^x \sim \text{ i.i.d } N(0, \sigma_x^2). \tag{6}$$

The firm productivity, $z_{it} = \log(Z_{it})$, has a stationary and monotone Markov transition function, denoted by $p_{z_i}(z_{i,t+1}|z_{it})$, as follows:

$$z_{i,t+1} = \rho_z z_{it} + \varepsilon_{i,t+1}^z, \qquad \varepsilon_{i,t+1}^z \sim \text{ i.i.d } N(0, \sigma_z^2). \tag{7}$$

And $\varepsilon_{i,t+1}^z$ and $\varepsilon_{j,t+1}^z$ are uncorrelated for any pair of firms (i, j) with $i \neq j$.

The capital accumulation rule is $K_{i,t+1} = (1 - \delta)K_{it} + I_{it}$, where I_{it} denotes investment and δ denotes the depreciation rate of installed capital. Investment is subject to quadratic adjustment costs given by g_{it}:

$$g(I_{it}, K_{it}) = \frac{1}{2}\eta\Big(\frac{I_{it}}{K_{it}} - \delta\Big)^2 K_{it}, \text{ with } \eta > 0. \tag{8}$$

Firms are equity financed and face a perfectly elastic supply of labor at a given stochastic equilibrium real wage rate W_t, as in Belo et al. (2014). The equilibrium wage rate, given by

$$W_t = \exp(\omega x_t), \text{ with } \omega \in [0, 1], \tag{9}$$

is assumed to be increasing with aggregate productivity, with ω determining the sensitivity of wages to aggregate conditions. Hiring decisions are made after firms observe the productivity shocks and labor is adjusted freely; hence, for each firm, marginal product of labor equals the wage rate

$$F_{L_{it}} = F_L(X_t, Z_{it}, K_{it}, L_{it}) = W_t.$$

Dividends to shareholders are equal to shareholders are equal to

$$D_{it} = Y_{it} - [I_{it} + g_{it}] - W_t L_{it}. \tag{10}$$

At each date t, firms choose $\{I_{it}, L_{it}\}$ to maximize the net present value of their expected dividend stream, V_{it}, which is the firm value

$$V_{it} = \max_{\{I_{i,t+\tau}, L_{i,t+\tau}\}} E_t\left[\sum_{k=0}^{\infty} M_{t,t+\tau} D_{i,t+\tau}\right], \tag{11}$$

subject to equations 6 to 9, where $M_{t,t+\tau}$ is the stochastic discount factor between time t and $t + \tau$.

The first order conditions for the firms's optimization problem leads to the pricing equation

$$1 = \int\int M_{t,t+1} R^I_{i,t+1} p_{z_i}(z_{i,t+1}|z_{it}) p_a(a_{t+1}|a_t) d_{z_i} d_a, \tag{12}$$

where the returns to investment are given by

$$R^I_{i,t+1} = \frac{F_{K_{i,t+1}} + (1-\delta)q_{i,t+1} + \frac{1}{2}\eta((I_{i,t+1}/K_{i,t+1})^2 - \delta^2}{q_{it}}, \tag{13}$$

and where $F_{K_{it}} = F_K(A_t, Z_{it}, K_{it}, L_{it})$ and $q_{it} = 1 + \eta\Big((I_{it}/K_{it}) - \delta\Big)$, which is the Tobin's (marginal) q.

Finally, the returns to the firm are defined as

$$R^S_{i,t+1} = \frac{V_{i,t+1}}{V_{it} - D_{it}}. \tag{13}$$

The model assumes an exogenous pricing kernel given by

$$\log M_{t+1} = \log \rho - \gamma_t \epsilon^x_{t+1} - \frac{1}{2}\gamma_t^2 \sigma_x^2, \text{ with } \log \gamma_t = \gamma_0 + \gamma_1 x_t, \tag{14}$$

where ρ, $\gamma_0 > 0$ and $\gamma_1 < 0$ are constant parameters.

In this model, the key mechanism relating firm-level productivity to expected returns involves the interaction of convex adjustment costs and the countercyclical Sharpe ratios assumed in our pricing kernel. Firm risk derives from its inability to freely adjust its capital following shocks to aggregate and firm-level productivity. In this economy, aggregate productivity shocks drive the business cycles. In bad times (low aggregate productivity level), net present value of investments go down because of lower expected cash flows and higher discount rates. Hence, all firms would like to invest less and hire less. Even though firms can freely adjust their labor, they incur adjustment costs when they change their capital stock. In states of low aggregate productivity, a bad aggregate shock tends to have a larger negative effect on the low TFP firms. These firms would have a lower investment rate (i.e., reduce their capital stock relatively more) than the high TFP firms.

5.2 Productivity estimation

Total factor productivity is a measure of the overall effectiveness with which capital and labor are used in a production process. It provides a broader gauge of firm-level performance than some of the more conventional measures, such as labor productivity or firm profitability. In this exercise, we consider the following specification:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} + \eta_{it}, \tag{15}$$

where y_{it} is the log of value added for firm i in period t; l_{it} and k_{it} are log values of labor and capital of the firm, respectively; ω_{it} is the productivity and η_{it} is an error term not known by the firm. Following İmrohoroğlu and Tüzel (2014), we employ the semiparametric procedure suggested by Olley and Pakes (1996) to estimate the parameters of this production function. The major advantage of this approach over more traditional estimation techniques such as the ordinary least squares (OLS) is its ability to control for selection and simultaneity biases and deal with the within firm serial correlation in productivity that plagues many production function estimates. Once we estimate the production function parameters, we obtain the firm-level (log) TFPs by

$$\omega_{it} = y_{it} - \beta_0 - \beta_k k_{it} - \beta_l l_{it}. \tag{16}$$

We followed Rovigatti and Mollisi (2018) to run this exercise[5]. Given that our panel also contains data on energy and materials, we also estimate productivity using the Levinsohn and Petrin (2003) methodology as a robustness test. The main advantage of the latter is to overcome the presence of many zero observations in the investment variable used as a proxy in the Olley-Pakes approach. Results are similar, as reported in Table 2. The estimations presented here are pooled by all industries, given the reduced sample size. However, results obtained by estimations using the sectors described in Table 11 are similar. The Olley-Pakes results for the entire sample period and all firms yield a labor share (and standard errors) of 0.87 (0.06) and a capital share of 0.14 (0.06). Using the Levinsohn-Petrin method we obtained a labor share (and standard errors) of 0.70 (0.07) and a capital share of 0.20 (0.08). All estimations are statistically significant. Besides that, as shown in Figure 5, although the estimated productivities using the Levinsohn-Petrin approach are smaller, their dynamic behavior is very similar to that of Olley-Pakes estimates, which is the main concern for our analysis here. Finally, Figure 4 shows the distribution of TFP estimates by three arbitrary years, results suggest that the productivity dispersion decreased during the time period of our panel.

5.3 Empirical results from estimated productivity analysis

TFP and other firm characteristics. Table 3 presents the summary statistics and various firm characteristics for firms sorted into 5 portfolios based on the level of their TFP over the 2010-2019 period. Results in Table 3 indicate a strongly monotonic relationship between TFP and many firm characteristics, including firm size and book-to-market ratios of firms. Market capitalization of firms monotonically increase with TFP. The average size of the firms in the lowest TFP quantile is 12% of the average size of all firms in that year, whereas the average size of firms in the highest TFP decile is 89% of the average size. The book-to-market ratios of the firms monotonically decline with TFP, indicating that high TFP firms are typically growth firms and low TFP firms are value firms.

The hiring rate, fixed investment to capital ratio, asset growth, investment to capital for organizational capital are all monotonically increasing in firm-level TFP. The firms in the lowest productivity quantile almost do not increase their workforce and their assets grow on average 2%, whereas firms in the highest productivity quantile increase their workforce by 10% and experience 12% asset growth. We also investigate the relationship between firm TFP and measures of profitability. Productivity and profitability have often been used interchangeably in finance literature (e.g., Novy-Marx, 2013)) where unobserved productivity is

[5] Also, we used the Stata® package `prodest`, created by the authors.

frequently proxied by measures of profitability. Our first measure is the return on equity (ROE), calculated as net income available to common stockholders/book equity. The second measure is the return on assets (ROA), defined as net income/total assets. The last profitability measure (GPR) is gross profits/book assets from Novy-Marx (2013). All three measures of profitability are monotonically and positively related to TFP.

Portfolio sorts. With the estimated firm-level productivities in hand, we are in position to investigate the relation between this firm characteristic and the expected stock returns by applying the methdology described in Section 4. To form the portfolios, we rank firms into five quantiles based on their estimated TFP. The portfolios are equally-weighted and rearranged every year t. Table 4 shows the average monthly returns of the five portfolios and a high-minus-low portfolio, which is short in low-TFP companies and long in high-TFP companies. All returns are statistically significant and our results show that TFP is positively and monotonically related to contemporaneous stock returns. The difference between the contemporaneous returns of high and low TFP firms is 6.27%. The relationship between the level of TFP and future excess returns is equally striking for equal-weighted portfolios: low productivity firms on average earn a 4.46% annual premium over high productivity firms in the following year and the return spread is statistically significant. These results go in line with İmrohoroğlu and Tüzel (2014) findings for the American stock market.

Alphas and betas of Portfolios sorted on TFP. In Table 5, we investigate whether widely used asset pricing models such as the capital asset pricing model (CAPM) and Fama–French (FF) three-factor model capture the variation in excess returns of TFP-sorted portfolios. As we demonstrate in Table 3, TFP is significantly related to size and B/M at the firm level. Hence, we explore whether the returns of TFP-sorted portfolios are systematically related to SMB and HML.

Table 5 presents the alphas and betas of TFP-sorted portfolios for the CAPM and FF models. Betas are estimated by regressing the portfolio excess returns on the factors. Alphas are estimated as intercepts from the regressions of excess portfolio returns. Monthly alphas are annualized by multiplying by 12. We find that low TFP portfolios load heavily on SMB, while high TFP portfolios loadings are low. HML loadings are not monotonic, but higher TFP portfolios have lower loadings than lower TFP portfolios. The MKT relation is not monotonic for CAPM, but show higher values for low TFP loadings in the case of the Fama-French model.

Finally, as stated by İmrohoroğlu and Tüzel (2014), these results do not necessarily suggest that TFP is a separate risk factor that is not captured by these factors but rather that TFP is systematically related to SMB, and to some extent to HML.

6 Misallocation analysis

In this section, we will investigate the existence of misallocation of resources in the Brazilian listed companies and its relation to expected returns. First, we describe the theoretical framework used to compute the misallocation measures, then we report and analyze the obtained results. This section is mainly based on the works of Hsieh and Klenow (2009), Oberfield (2013) and Vasconcelos (2017).

6.1 Theoretical framework

Vasconcelos (2017) already found evidence of misallocation of resources in the Brazilian manufacturing sector. The author used the National Industrial Survey (PIA), an annual survey conducted by the Brazilian Institute of Geography and Statistics. It is a database with more than 40,000 firms on average each year, making it a much more appropriate sample to measure misallocation at the national level. The sample used here is rather small and potentially biased, given that listed companies are expected to face less financial constraints. That said, however, they offer the great advantage to allow a direct risk measure through the analysis of stock returns, something difficult to do for non-listed firms in general[6]. Therefore, misallocation measures here are specifically useful for the analysis intended in this paper in the context of the universe of Brazilian publicly listed companies.

To obtain preliminary evidence on misallocation between listed Brazilian companies, we followed the same framework constructed by Hsieh and Klenow (2009) and adapted by authors like Oberfield (2013) and Vasconcelos (2017). It begins with a monopolistic competition model with heterogeneous firms. Suppose an infinite-horizon closed economy. There is a single final good Y_t produced by a representative firm in a perfectly competitive market. Y_t is produced through a Cobb-Douglas technology using the combination of intermediate goods X_{st} of S manufacturing industries:

$$Y_t = \prod_{s=1}^{S} X_{st}^{\theta_s}, \text{ where } \sum_{s=1}^{S} \theta_s = 1 \text{ and } \theta_s \in (0, 1). \tag{17}$$

Cost minimization implies:

$$\theta_s = \frac{P_{st} Y_{st}}{P_t Y_t}, \tag{18}$$

where P_{st} is the price of intermediate good X_{st} and $P_t = \prod_{s=1}^{S} (P_{st}/\theta_s)^{\theta_s}$ is the price of the final good. Industry output X_{st} is itself a CES aggregate of I_s differentiated products:

[6]However, David, Schmid and Zeke (2020 offer a promising approach in this direction through the estimation of betas relating changes in firm-level productivity and aggregate productivity shocks.

$$X_{st} = \left(\sum_{i \in I_s} X_{ist}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \text{ where } \sigma \in (0, \infty). \tag{19}$$

Each differentiated product is produced according to

$$X_{ist} = M_{ist} K_{ist}^{\alpha_s} L_{ist}^{1-\alpha_s}, \text{ where } \alpha_s \in (0, 1). \tag{20}$$

Here, α_s is the capital share, M_{ist} is TFP, L_{ist} is a labor factor and K_{ist} is physical capital. Note that capital and labor shares are allowed to differ across industries, but not across firms within an industry.

6.2 Measuring distortions

Because there are two factors of production, we can separately identify distortions that affect both capital and labor from distortions that change the marginal product of one of the factors relative to the other factor of production. We denote distortions that increase the marginal products of capital *and* labor by the same proportion as an output distortion τ_Y. In turn, we denote distortions that raise the marginal product of capital relative to labor as the capital distortion τ_K. Therefore, profits are given by

$$\pi_{is} = (1 - \tau_{Yis}) P_{ist} Y_{ist} - w_{ist} L_{ist} - (1 + \tau_{Kis}) R K_{ist}, \tag{21}$$

where w_{ist} is the wage and R_t is the cost of capital, both of which are time-variant. Hsieh and Klenow (2009) show that the allocation of resources across firms depends not only on firm TFP levels, but also on the output and capital distortions they face. To the extent resource allocation is driven by distortions rather than firm TFP, this will result in differences in the marginal revenue products of labor and capital across firms. The marginal revenue product of labor is proportional to revenue per worker and the marginal revenue product of capital is proportional to the revenue-capital ratio. They are, respectively:

$$MRPL_{ist} = \frac{1}{1 - \tau_{Yis}} w_{ist} \tag{22}$$

and

$$MRPK_{ist} = \frac{1 + \tau_{Kis}}{1 - \tau_{Yis}} R_t. \tag{23}$$

Foster, Haltiwanger, and Syverson (2008) stress that, when industry deflators are used, differences in plant-specific prices show up in the customary measure of plant TFP. They stress the distinction between "physical productivity," which they denote TFPQ, and "revenue productivity," which they call TFPR. The use of a plant-specific deflator yields TFPQ, whereas using an industry deflator gives TFPR. We define these objects as follows:

$$TFPQ_{ist} \equiv M_{ist} = \frac{X_{ist}}{K_{ist}^{\alpha_s} L_{ist}^{1-\alpha_s}}, \tag{24}$$

and

$$TFPQ_{ist} \equiv P_{ist} M_{ist} = \frac{P_{ist} X_{ist}}{K_{ist}^{\alpha_s} L_{ist}^{1-\alpha_s}}. \tag{25}$$

Using equations 22 and 23, we get

$$TFPR_{ist} \propto (MRPK_{ist})^{\alpha_s} (MRPL_{ist})^{1-\alpha_s} \propto \frac{(1+\tau_{Kis})^{\alpha_s}}{1-\tau_{Yis}}. \tag{26}$$

With the expression for TFPR in hand, we can express industry TFP as

$$TFPR_{st} = \left[\sum_{i=1}^{M_s} \left(M_{ist} \frac{\overline{TFPR}_{st}}{TFPR_{\ ist}} \right) \right], \tag{27}$$

where $\overline{TFPR}_{st}$ is a geometric average of the average marginal revenue product of capital and labor in the sector. Additionally, assume that $\{\log M_{ist},\ \log(1-\tau_{Yis}),\ \log(1+\tau_{Kis})\}$ has a multivariate lognormal distribution. Let ϕ_Y and ϕ_K be the standard deviations of $\log(1-\tau_{Yis})$ and $\log(1+\tau_{Kis})$ respectively, and ϕ_{YK} be their covariance. Then,

$$\log TFP_{st} = \frac{1}{1-\sigma} \log E_t[TFPQ_{ist}^{\sigma-1}] - \frac{\sigma}{2}\phi_Y^2 - \left(\frac{\alpha_s + \alpha_s^2(\sigma-1)}{2} \right) \phi_K^2 + \sigma \alpha_s \phi_{YK}. \tag{28}$$

In this special case, the negative effect of distortions on aggregate TFP can be summarized by the variance of log TFPR. Intuitively, the extent of misallocation is worse when there is greater dispersion of marginal products.

6.3 MPK and risk premia

Now, following David, Schmid and Zeke (2020), we lay out a version of the standard, frictionless neoclassical theory of investment to illustrate the link between firm-level MPK and risk premia. Consider a slightly modified version of the model developed in Section 5. There is a continuum of firms with measure one producing a homogeneous good according to

$$Y_{it} = X_t^{\hat{\beta}_i} \hat{Z}_{it} K_{it}^{\theta_1} L_{it}^{\theta_2}, \text{ where } \theta_1 + \theta_2 < 1. \tag{29}$$

Firm productivity in logs is equal to $\hat{\beta}_i x_t + \hat{z}_{it}$, where $\hat{z}_{it}$ denotes a firm-specific idiosyncratic component of productivity, x_t is an aggregate component that is common across firms and $\hat{\beta}_i$ captures the exposure of the productivity of firm i to aggregate conditions, with $\hat{\beta}_i \sim N\left(\overline{\hat{\beta}}, \sim \sigma^2_{\hat{\beta}}\right)$ across firms. The two productivity components follow AR(1) (in logs) processes as before. We define the stochastic discount factor, $m_{t+1} \equiv \log M_{t+1}$ as in Section 5.

We keep the labor market simple and also identical to Section 5. Thus, maximizing over the static labor decision gives operating profits - revenue less labor costs - as

$$\Pi_{it} = G X_t^{\beta_i} Z_{it} K_{it}^{\theta}, \tag{30}$$

where $G \equiv (1-\theta_2)\theta_2^{\frac{\theta_2}{1-\theta_2}}$, $\beta_i \equiv \frac{1}{1-\theta_2}(\hat{\beta}_i - \omega\theta_2)$, $Z_{it} \equiv \hat{Z}_{it}^{\frac{1}{1-\theta_2}}$ and $\theta \equiv \frac{\theta_1}{1-\theta_2}$. The exposure of firm profits to aggregate conditions is captured by β_i, which is a simple transformation of the underlying exposure of firm productivity to the aggregate component, $\hat{\beta}_i$, and the sensitivity of wages, ω. The idiosyncratic component of productivity is similarly scaled, by $\frac{1}{1-\theta_2}$. The curvature of the profit function is equal to θ, which depends on the relative elasticities of capital and labor in production. These scalings reflect the leverage effects of labor liabilities on profits. From here on, we will primarily work with z_{it}, which has the same persistence as $\hat{z}_{it}$, i.e., ρ_z, and innovations $\varepsilon_{it+1} = \frac{1}{1-\theta_2}$ with variance $\sigma^2_{\tilde{\varepsilon}} = \left(\frac{1}{1-\theta_2}\right)^2 \sigma^2_{\varepsilon}$. We will also use the fact that $\sigma^2_{\beta} = \left(\frac{1}{1-\theta_2}\right)^2 \sigma^2_{\hat{\beta}}$.

In addition, we show in Appendix 7 that the firm's optimal investment policy is given by:

$$k_{it+1} = \frac{1}{1-\theta}(\tilde{\alpha} + \beta_i \rho_x x_t + \rho_z z_{it} - \beta_i \gamma_t \sigma^2_{\varepsilon}), \tag{31}$$

where $\tilde{\alpha} \equiv \log\theta + \log G - \alpha$, $\alpha \equiv \log(r_f + \delta)$. The firm's choice of capital is increasing in x_t and z_{it} due to their direct effect on expected future productivity (i.e., $\beta_i \rho_x x_t + \rho_z z_{it} = \mathbb{E}_t[\beta_i x_{t+1} + z_{it+1}]$), but, *ceteris paribus*, firms with higher betas choose a lower level of capital.

Finally, by definition the realized *mpk* is given by $mpk_{it+1} = \log\theta + \pi_{it+1} - k_{it+1}$. Substituting for k_{it+1},

$$mpk_{it+1} = \alpha + \varepsilon_{it+1} + \beta_i \varepsilon_{t+1} + \beta_i \gamma_t \sigma^2_{\varepsilon}, \tag{32}$$

and taking conditional expectations,

$$\mathbb{E}_t[mpk_{it+1}] = \alpha + \beta_i \gamma_t \sigma^2_{\varepsilon}. \tag{33}$$

Expression 32 shows that dispersion in the realized *mpk* can stem from uncertainty over the realization of shocks, as well as the risk premium term, which is persistent at the firm level and depends on (i) the firm's exposure to the aggregate shock, β_i (and is increasing in β_i), and (ii) the time t price of risk, which is reflected in the term $\gamma_t \sigma_\varepsilon^2$. Intuitively, firm-level mpk deviations are composed of both a transitory component due to uncertainty and a persistent component due to the risk premium. The transitory components are i.i.d. over time and lead to purely temporary deviations in mpk (even though the underlying productivity processes are autocorrelated); the risk premium, on the other hand, leads to persistent deviations -– firms that are more exposed to aggregate shocks, and so are riskier, will have persistently high mpk.

And the cross-sectional variance of $\mathbb{E}_t[mpk_{it+1}]$ is given by

$$\sigma^2_{\mathbb{E}_t[mpk_{it+1}]} = \sigma^2_\beta (\gamma_t \sigma^2_\varepsilon)^2. \tag{34}$$

Cross-sectional variation in $\mathbb{E}_t[mpk_{it+1}]$ depends on the dispersion in beta and the price of risk. Dispersion will be greater when risk prices, reflected by $\gamma_t \sigma_\varepsilon^2$, are high and so will be countercyclical.

6.4 Basic results

With the measures described above, we are able to use the balance sheet data panel to measure the misallocation degree in our sample. Figure 6 shows the distribution of the logarithm of TFPR in the Brazilian listed companies arbitrarily for 2010, 2015 and 2019. Fundamentally, we are interested in the dispersion of TFP. The more disperse TFP, the greater can be the misallocation. A high dispersion implies that some firms are more able to produce output with the same amount of inputs, given the technology process in each sector. According to this figure, despite the particularities of the sample, results are in line with previous studies. The dispersion is relatively elevate in each time period. Furthermore, the median is greater than the average for the time periods. Therefore, the distribution of TFP is asymmetric. These two points suggest that there exists misallocation of resources at the firm-level in the Brazilian listed manufacturing companies.

Although more dispersed, results resemble those from Olley-Pakes estimation in Figure 4 in terms of means and behavior along the three years showed. Additionally, Figure 7 shows the (log) distribution of input wedges for Brazilian listed companies in the year of 2019[7]. Capital wedge distribution shows an elevated dispersion and a positive mean, suggesting the presence of over-investment by the firms in our sample. These results must be viewed with caution, given that the reduced size of our sample makes it inappropriate for a rigorous misallocation analysis similar to the one performed by Vasconcelos (2017) using PIA, but the central point of this exercise was to measure the firms' MPKs to be used in the empirical exercises below.

[7]The previous years of the sample showed a similar picture.

6.5 Empirical results from measured misallocation analysis

Portfolio sorts. With the estimated MPKs in hand, we are in position to investigate the relation between this firm characteristic and the expected stock returns by applying the methdology described in Section 4. To form the portfolios, we rank firms into five quantiles based on their estimated MPK. The portfolios are equal-weighted and rearranged every year t. Table 6 shows the average monthly returns of the five portfolios and a high-minus-low portfolio, which is short in low-MPK companies and long in high-MPK companies. All results are statistically significant. Even though the relation seems to be not perfectly monotonic, the first row shows that the difference in contemporaneous returns between high and low MPK firms, i.e., the excess return on the MPK-HML portfolio, is over 6%.

The second row confirms that this finding does not simply result from the simultaneous response of stock returns and MPK to the realization of unexpected shocks – one-period ahead excess returns are in fact predictable by MPK. The predictable spread on the MPK-HML portfolio is over 5% annually. Both the contemporaneous and future MPK-HML spreads are statistically different from zero at the 99% or 95% levels. Thus, high MPK tend to offer high stock returns, both in a realized and an expected sense, suggesting that MPK differences reflect exposure to risk factors for which investors demand compensation in the form of a higher rate of return.

Measures of risk exposures and expected MPK. In this exercise we directly relate firm MPK to measures of risk exposures estimating regressions of the form

$$mpk_{it+1} = \psi_0 + \psi_\beta \beta_{it} + \zeta_{it+1}, \tag{35}$$

where β_{it} is a measure of firm i exposure to aggregate risk at time t. The specification tests whether observable measures of firm-level risk exposures are indeed correlated with higher MPK. We estimate (5) at an annual frequency and lag the right-hand side variable to control for the simultaneous effect of unexpected shocks on contemporaneous measures of beta and MPK. Following David, Schmid and Zeke (2020), we construct two different measures of these exposures. First, we compute standard CAPM and Fama-French stock market betas, β_{CAPM} and β_{FF} respectively, by estimating firm-level regressions of stock returns on the risk factors from each of these models based on the procedure described in Section 3. With these measures in hand, we are in a position to estimate Equation 35. Table 7 reports the results of this exercise. Both measures have significant explanatory power for subsequent MPK. For example, the estimate in column (1) implies that each unit increase in the CAPM beta is associated with a 26% increase in expected MPK. In columns (3) and (4) of Table 7, we estimate analogous regressions with the addition of industry-year fixed-effects. All of the beta coefficients remain positive and statistically significant.

Price of risk and mean expected MPK. Expression 32 implies that the price of risk should positively predict the level of expected MPK. To test this implication, we estimate a time-series regression of the form:

$$E[mpk_{it+1}] = \psi_0 + \psi_1 \lambda_t + \zeta_{t+1}, \tag{36}$$

where $E[mpk_{it+1}]$ denotes the average mpk in period $t+1$ and λ_t denotes the price/dividend (PD) ratio on the aggregate stock market (likely negatively correlated with the price of risk). The estimated coefficient is approximately -0.38, in line with the theory indicating a negative correlation between PD ratio and future MPK, but the estimation is not statistically significant, probably because of the reduced time span of the panel. To illustrate this relation a little more, we report in Table 8 some contemporaneous correlations between MPK dispersion and indicators of the price of risk and the business cycle[8]. Results indicate negative relations between MPK and PD ratio, GDP and TFP. Besideds that, these correlations suggest that our measure of the price of risk (PD ratio) is countercyclical, implying that variation in risk premia induce a countercyclical component in MPK.

MPK dispersion and beta dispersion. Expression 34 implies that across groups of firms or segments of the economy, dispersion in expected MPK should be positively related to dispersion in risk exposures. Again, following David, Schmid and Zeke (2020), we investigate this prediction using variation in the dispersion of firm-level betas and expected stock market returns across industries. Specifically, for each industry in each year, we compute the standard deviation of MPK, $\sigma(mpk)$, expected stock returns, $\sigma(\mathbb{E}[r])$, and beta, $\sigma(\beta)$. We then estimate regressions of industry-level MPK dispersion on the dispersion in expected returns and betas, i.e.,

$$\sigma(mpk_{jt+1}) = \phi_0 + \phi_1 \sigma(x_{jt}) + \zeta_{jt+1}$$

where $x_{jt} \in \{\mathbb{E}[r_{jt}], \beta_{jt}\}$, j denotes industry. Again, to avoid potential simultaneity biases from the realization of shocks, we lag the independent variables (dispersion in expected returns and betas). Table 9 reports the results of these regressions and demonstrates that indeed, industries with higher dispersion in expected stock returns and beta exhibit greater dispersion in MPK. Column (1) reveals this fact using expected returns calculated from the Fama-French model. Variation in expected return dispersion predicted by the Fama-French model explains approximately 30% of the variation in MPK dispersion across industry-

[8] In all estimations we extract the cyclical components of GDP, TFP and the PD ratio using a one-sided Hodrick-Prescott filter

years. Column (2) regresses MPK dispersion on dispersion in each of the three individual factors -– variation in the beta on each factor is significantly related to MPK dispersion. Columns (3) and (4) add year fixed-effects. Across these specifications, measures of within-industry heterogeneity in expected returns and aggregate risk exposures remain positive and significant predictors of within-industry dispersion in MPK, except for the SMB beta. Results here are in line with those from David, Schmid and Zeke (2020) but, interestingly, are relatively larger than the estimates for the American stock market obtained by the authors.

7 Concluding remarks

In this paper, we studied the link between productivity and misallocation of resources, risk and stock returns focusing on a developing country reality plagued by high risk and severe market frictions. From a panel data set of Brazilian publicly listed companies we conducted a series of empirical exercises to test the predictions of two similar models that investigate these issues.

The results obtained here seem to support the hypotheses that firm-level total factor productivity and marginal product of capital rationally predict stock returns and are related to firm-level risk exposure to aggregate conditions. The first analysis found that TFP is positively related to contemporaneous stock returns, but negatively related to future stock returns. The second analysis suggests that expected firm-level marginal products should reflect exposure to factor risks, and their pricing. To the extent that firms are heterogeneously exposed to these risks, as the literature on cross-sectional asset pricing suggests, the implication is that cross-sectional dispersion in MPK may not only reflect true misallocation, but also risk-adjusted capital allocation.

These results, however, should be interpreted with caution since we did not provide causal evidence on the hypotheses derived from the models reviewed. Thus, the main goal of this paper through its empirical exercises is to initiate a research agenda by applying recent theoretical developments relating empirical finance and firm dynamics to the context of developing countries. There is still a lot to do be done in this direction. A natural next step for future research is to include more developing countries in our sample in order to avoid the caveats arising from the relatively reduced size and particularities of the Brazilian stock market. In addition, a more detailed and reliable analysis could be done by the quantitative analysis of calibrated versions of the models presented here. Nevertheless, results here remain important for reinforcing the notion that financial market considerations can have sizable effects on real outcomes by affecting capital allocation decisions and therefore explain part of the productivity gaps between rich and poor countries.

References

Balvers, J., Gu, L., Huang, D. (2017). Profitability, value and stock returns in production-based asset pricing without frictions. *Journal of Money, Credit, and Banking*, 49(7):1621-1651.

Blank, F., Samanez, C., Baidya, T., Aiube, F. (2014). Conditional CAPM: Time-varying Betas in the Brazilian Market. *Brazilian Review of Finance*, 12(2):163-199.

Belo, F., Lin, X. and Bazdresch, S. (2014). Labor hiring, investment, and stock return predictability in the cross section, *Journal of Political Economy*, 122: 129–177.

Buera, F., Kaboski, J. and Shin, Y. (2011). Finance and Development: A Tale of Two Sectors, *American Economic Review*, 101: 1964–2002.

Cattaneo, M.; Crump, R.; Farrell, M. and Schaumburg, E. (2020). Characteristic-Sorted Portfolios: Estimation and Inference *The Review of Economics and Statistics*, 102(3):531-551.

Cochrane, J. (1991). Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations. *Journal of Finance*, 46, 207–234.

David, J., Schmid, L. and Zeke, D. (2020). Risk-Adjusted Capital Allocation and Misallocation. *Working Paper.*

Fama, E. and French, K. (1992). The cross section of expected stock returns. *Journal of Finance*, 47:427–465.

Fama, E. and MacBeth, J. (1973). Risk, Return, and Equilibrium: Empirical Tests, *Journal of Political Economy*, 81: 607–636.

Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1): 394-425.

Garcia, A. and Santos, A. (2018). Dissecting Anomalies with the Five-factor Model for the Brazilian Stock Market. *Brazilian Review of Finance*, 16(1): 81-122.

Gomes, J., A. Yaron, and L. Zhang (2006). Asset pricing implications of firms financing constraints, *Review of Financial Studies*, 19: 1321–1356.

Gorodnichenko et al. (2020). Resource Misallocation in European Firms: The Role of Constraints, Firm Characteristics and Managerial Decisions, *Working Paper.*

Hsieh, C. and Klenow, P. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.

İmrohoroğlu, A and Tüzel, Ş. (2014). Firm-level productivity, risk and return. *Management Science*, 60(8):2073-2090.

Kogan, L. and D. Papanikolaou (2013). Firm characteristics and stock returns: The role of investment-specific shocks, *The Review of Financial Studies*, 26, 2718–2759.

Kohn, D., Leibovici, F. and Tretvoll, H. (2021). Trade in Commodities and Business Cycle Volatility. *American Economic Jounral: Macroeconomics*, forthcoming.

Levinsohn, J. and Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies*, 70(2): 317-341.

Moll, B. (2014). Productivity losses from financial frictions: can self-financing undo capital misallocation? *The American Economic Review*, 104: 3186–3221.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108: 1–28.

Oberfield, E. (2013). Productivity and misallocation during a crisis: evidence from the Chilean crisis of 1982. *Review of Economic Dynamics*, 16:100–119.

Olley, G. and Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64(6): 1263-1297.

Restuccia, D., and Rogerson, R. (2017). The Causes and Costs of Misallocation. *Journal of Economic Perspectives*, 31 (3), 151-74.

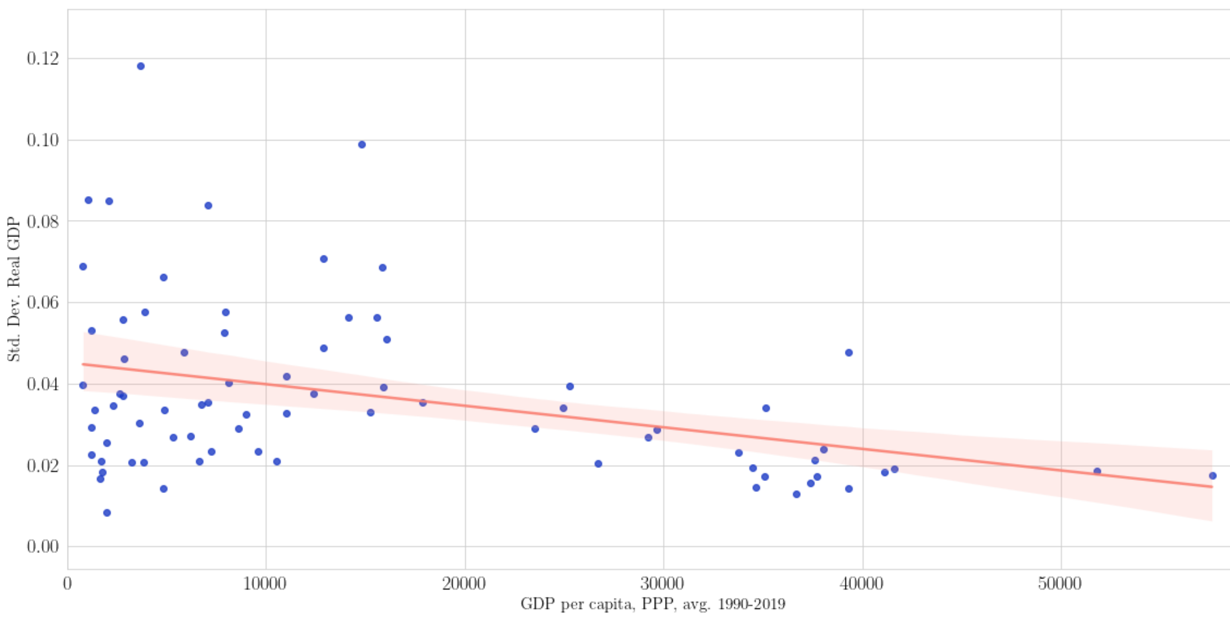
Rovigatti, G. and Mollisi, V. (2018). Theory and practice of total-factor productivity estimation: The control function approach using Stata. *Stata Journal*, 18(3): 618–662.

Varga, G. and Brito, R. (2016). The Cross-Section of Expected Stock Returns in Brazil. *Brazilian Review of Finance*, 14(2): 151-187.

Vasconcelos, R. (2017). Misallocation in the Brazilian manufacturing sector. *Brazilian Review of Econometrics*, 37(2): 191-232.

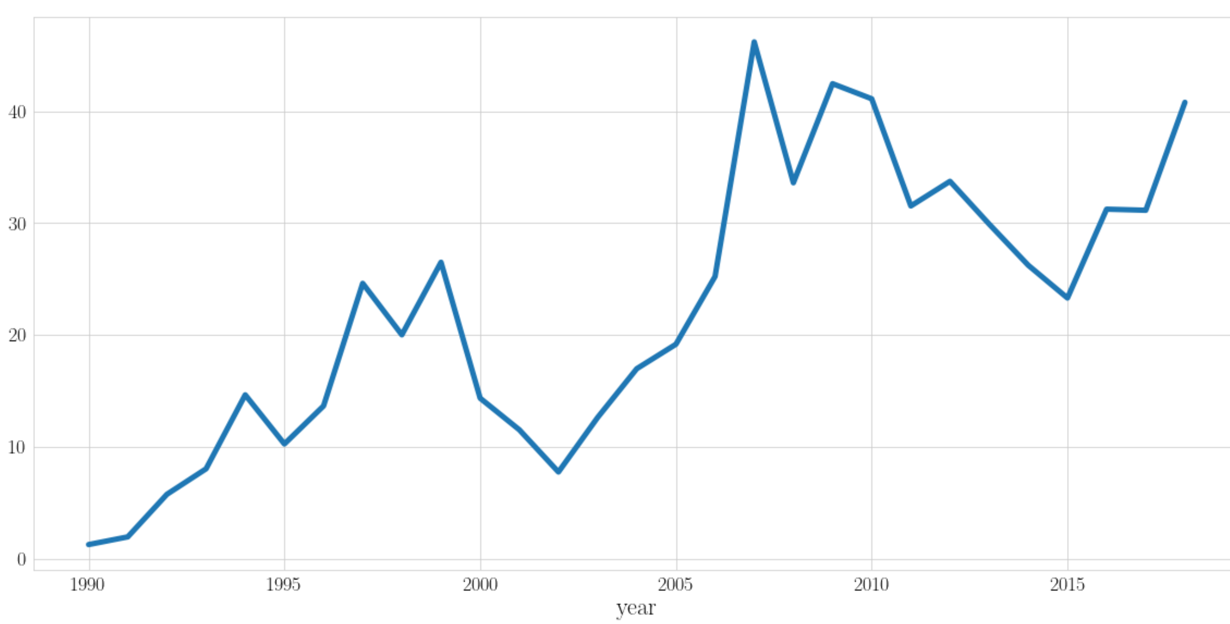
Zhang, L. (2017). The Investment CAPM. *European Financial Management*, 23(4): 545-603.

Figure 1: Economic development and business cycle volatility



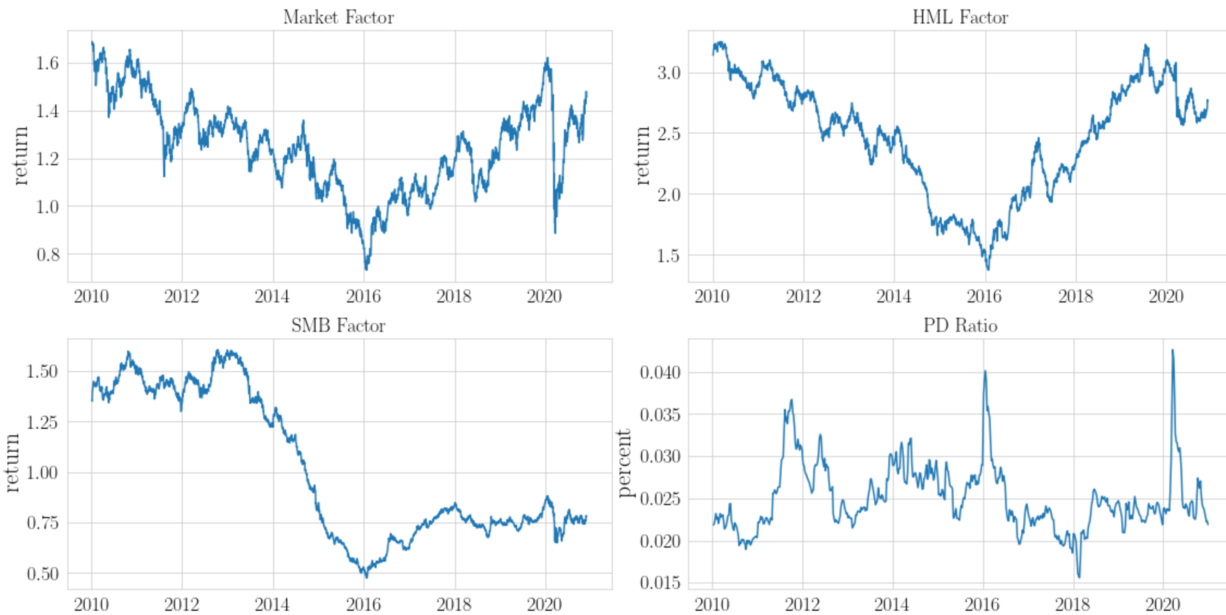
Note: This figure is an updated version of the exercise conducted by Kohn, Leibovici and Tretvoll (2021).

Figure 2: Stocks traded, total value (% of GDP) - Brazil



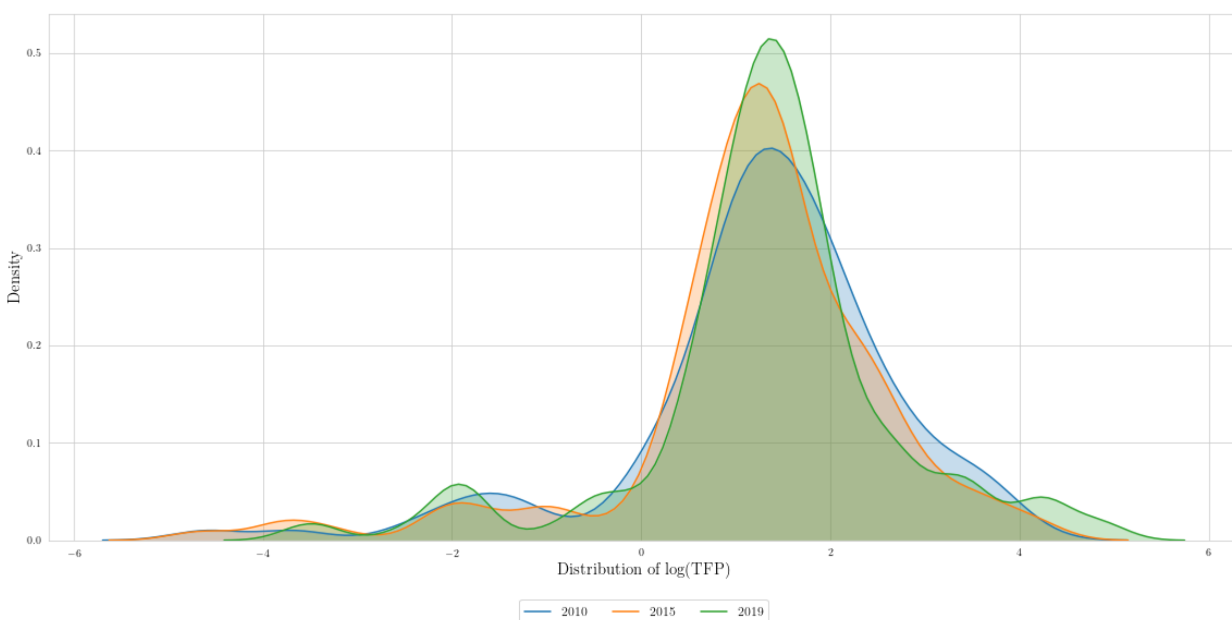
Note: This figure was constructed using data from the World Bank.

Figure 3: Risk factors and PD Ratio (2010-2020)



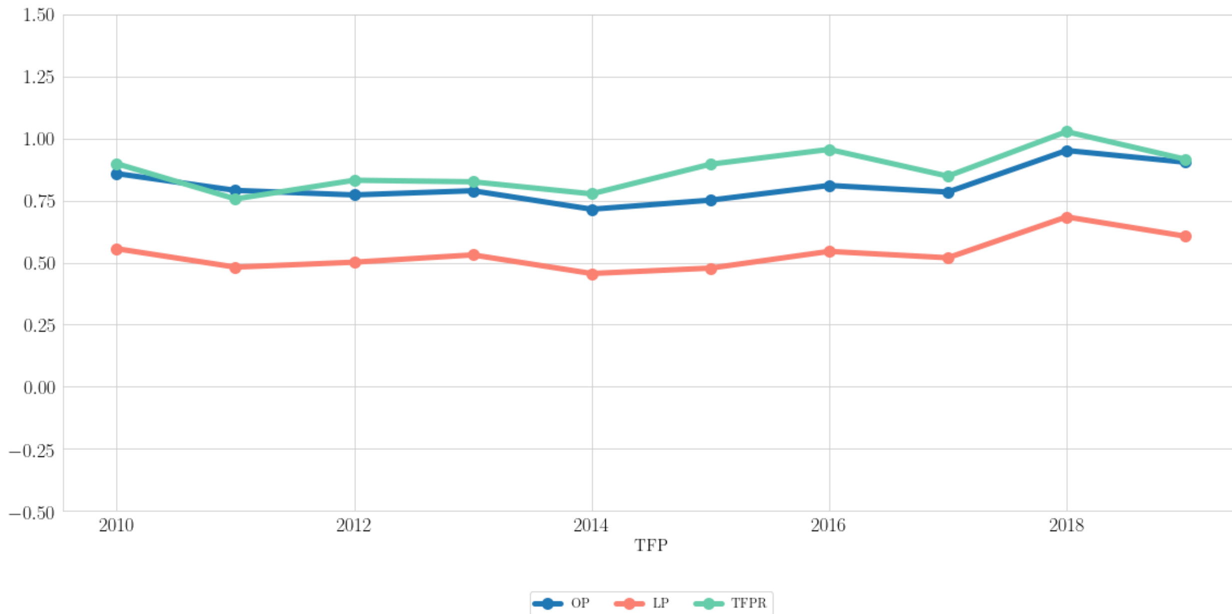
Note: This figure was constructed using data from Brazilian Center for Research in Financial Economics (NEFIN). It shows the dynamics of the three risk factors present in the Fama-French asset pricing model and the price/dividend (PD) ratio for the aggregate Brazilian stock market.

Figure 4: Olley-Pakes estimated TFP distribution



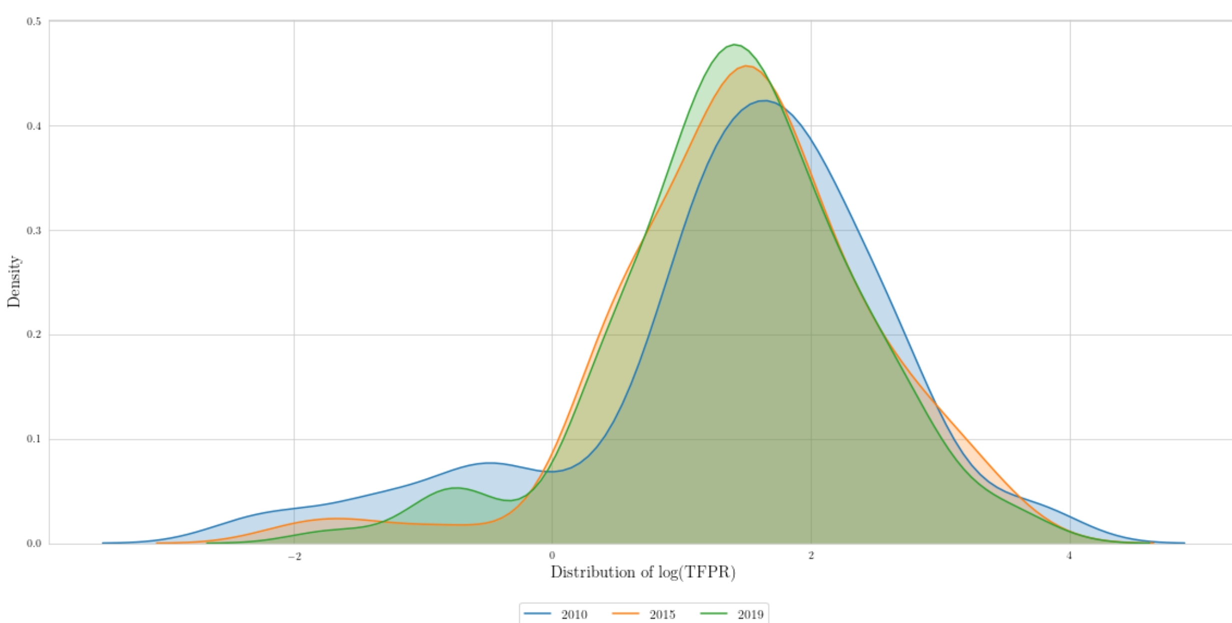
Note: This figure shows the dispersion of estimated firm-level total factor productivity using the Olley and Pakes (1996) approach for three arbitrary years.

Figure 5: TFP mean by year



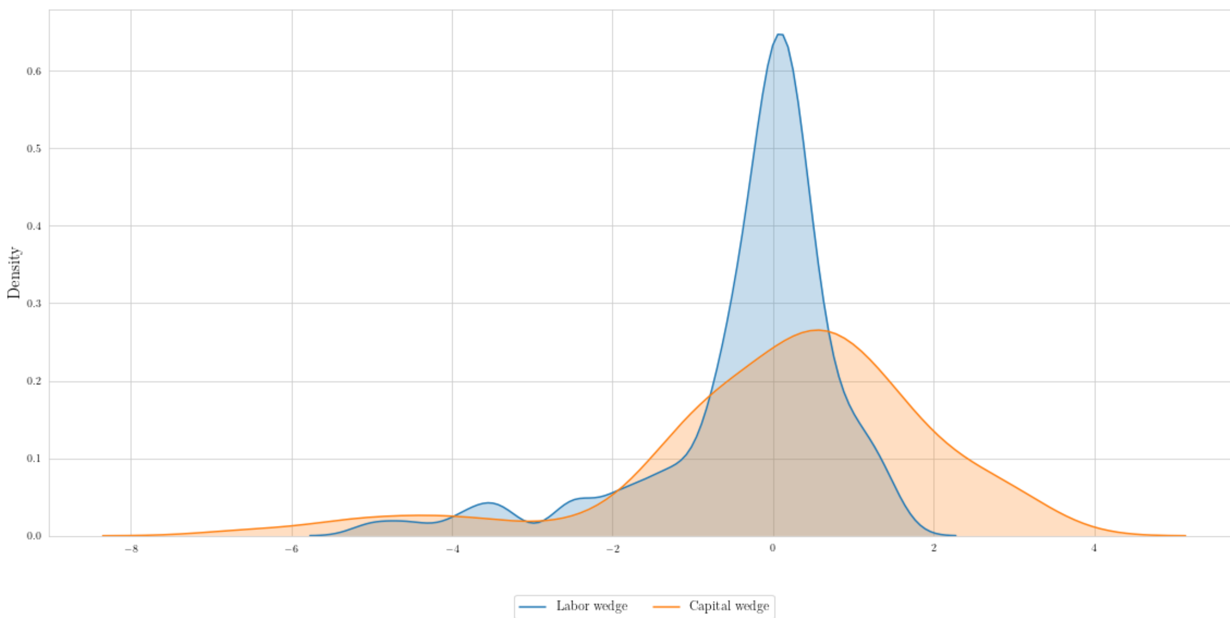
Note: This figure shows the dynamics of the mean estimated total factor productivity using three different methods.

Figure 6: TFPR distribution



Note: This figure shows the dispersion of estimated firm-level total factor productivity using the Hsieh and Klenow (2009) approach for three arbitrary years.

Figure 7: Wedges distribution (2019)



Note: This figure shows the dispersion of estimated input wedges for the year of 2019.

Table 1: Descriptive statistics

Year	Revenue from sales		Fixed assets		Wage bill		Energy and materials		Capex		Value Added		N
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.	
2010	13,334.06	53,826.10	7,467.19	31,372.11	1,875.04	9,292.35	2,045.29	9,997.26	824.23	3,389.20	8,057.72	35,868.86	261
2011	12,684.27	52,228.06	6,851.37	30,081.62	1,944.45	9,300.25	2,009.36	9,305.18	795.70	3,194.42	7,383.02	33,639.47	271
2012	16,049.08	62,247.56	8,266.21	33,065.19	2,729.17	11,855.88	2,809.10	12,429.57	824.32	3,217.93	8,596.29	36,161.20	278
2013	16,205.62	61,564.90	8,171.77	32,407.22	2,806.19	12,036.88	2,868.37	13,683.25	866.76	3,228.00	8,932.90	36,511.53	281
2014	15,822.49	60,481.35	7,960.89	31,194.06	2,571.83	11,559.08	3,067.50	14,319.58	863.00	3,178.10	8,005.87	32,571.56	287
2015	15,213.72	55,548.86	7,639.81	30,006.92	2,690.94	11,425.44	2,672.15	12,177.45	1,050.64	3,706.03	8,234.18	33,315.64	289
2016	13,943.68	48,876.39	6,924.89	27,321.22	2,330.79	10,660.35	1,974.27	7,870.36	796.01	2,698.57	7,633.89	32,347.48	301
2017	14,088.01	49,477.85	6,804.47	26,851.82	2,314.14	10,523.32	1,909.37	7,970.31	733.58	2,481.32	7,840.81	3,2817.77	306
2018	14,115.86	50,087.81	6,401.87	25,922.57	2,424.63	10,532.08	1,929.73	7,638.65	764.99	2,346.00	7,898.53	32,195.70	314
2019	16,034.05	56,267.30	8,511.21	33,321.22	2,813.07	11,769.93	2,974.35	14,990.81	1,045.57	3,401.47	9,142.24	37,142.14	332

Note: This table reports descriptive statistics of the main variables used in the empirical exercises. N is the number of unique firms by year after all the cleaning of the administrative data panel. All values are divided by R$1,000.00.

Table 2: Productivity estimation coefficients

Parameter	TFP-method	
	Olley-Pakes	LP
Labor	0.8668***	0.7019***
	(0.0638)	(0.0697)
Capital	0.1492***	0.2036***
	(0.0564)	(0.0866)
N	1321	1321

Note: This table reports the production function parameters estimated using the two methods described in Olley and Pakes (1996) and Levinsohn and Petrin (2003). Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Descriptive Statistics for TFP-Sorted Portfolios, 2009–2019

	Low	2	3	4	High	HML
TFP	0.18	0.49	0.72	1.07	2.01	1.83
SIZE	0.12	0.36	0.61	0.75	0.89	0.77
B/M	1.09	1.02	1.07	0.91	0.90	-0.19
I/K	0.7	0.9	0.12	0.13	0.17	0.10
AG	0.02	0.05	0.9	0.11	0.12	0.15
HR	0.01	0.02	0.04	0.07	0.10	0.09
ROE	-0.17	-0.90	0.03	0.07	0.09	0.26
ROA	-0.11	-0.42	0.01	0.05	0.08	0.19
GPR	0.25	0.27	0.39	0.42	0.45	0.47

Note: For each variable, averages are first taken over all firms in that portfolio, then over years. Average TFP each year is normalized to be 1. SIZE is the market capitalization of firms in June of year $t-1$. Average size each year is normalized to 1. B/M is the ratio of book equity for the last fiscal year-end in year t divided by market equity in December of year t. I/K is the fixed investment to capital ratio. AG is the change in the natural log of assets, HR is the change in the natural log of wage bill, and ROE is the net income in year t divided by book equity for year t. ROA is the net income in year t divided by total assets for year t. GPR is the gross profits in year t divided by book assets for year t.

Table 4: Excess Returns on TFP-Sorted Portfolios

Portfolio	Low	2	3	4	High	TFP-HML
r^e_t	4.83***	7.33***	7.99***	9.99***	11.10***	6.27***
	(2.89)	(6.27)	(5.51)	(6.19)	(11.96)	(13.64)
r^e_{t+1}	10.17***	7.88***	8.30***	7.43***	5.71***	-4.46***
	(8.19)	(4.51)	(5.12)	(4.74)	(3.09)	(4.43)

Note: This table reports stock market returns for portfolios sorted by TFP. r^e_t denotes equal-weighted contemporaneous annualized monthly excess stock returns (over the risk-free rate) measured in the year of the portfolio formation from January to December of year t. r^e_{t+1} denotes the analogous future returns, measured from July of year $t+1$ to June of year $t+2$. t-statistics in parentheses, computed using Newey-West standard errors. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Alphas and Betas of Portfolios Sorted on TFP (%, Annualized)

	Low	2	3	4	High	High-Low
			CAPM			
Alpha	6.70	5.25	4.13	4.40	3.95	-2.8
	(3.20)	(3.42)	(3.15)	(2.90)	(2.77)	(-2.81)
MKT	1.11	1.06	1.18	1.11	1.23	0.12
	(15.11)	(12.50)	(13.25)	(14.00)	(18.12)	(19.15)
			Fama-French			
Alpha	2.87	1.76	1.25	0.90	0.86	-2.01
	(2.45)	(1.42)	(1.17)	(0.35)	(0.78)	(0.46)
MKT	1.43	1.40	1.16	1.16	1.02	-0.41
	(12.11)	(13.27)	(10.02)	(14.11)	(15.16)	(15.01)
HML	0.41	0.49	0.39	0.32	0.28	-0.13
	(6.38)	(8.57)	(6.33)	(5.99)	(4.25)	(-3.98)
SMB	1.03	0.79	0.74	0.66	0.55	-0.48
	(12.23)	(11.31)	(10.22)	(9.31)	(8.44)	(-8.67)

Note: This table presents the regressions of equal-weighted and value-weighted excess portfolio returns on various factor returns. MKT, SMB, and HML factors are taken from NEFIN's website. The portfolios are sorted on TFP. Alphas are annualized (%). Returns are measured from July 2010 to June 2019. The t-statistics, computed using the Newey–West estimator allowing for one lag of serial correlation in returns, are in parentheses.

Table 6: Excess Returns on MPK-Sorted Portfolios

Portfolio	Low	2	3	4	High	HML
r^e_t	3.20***	5.73***	11.23***	11.46***	9.24***	6.04***
	(2.38)	(4.10)	(7.58)	(7.44)	(8.41)	(6.43)
r^e_{t+1}	4.02***	5.06***	9.54***	11.73***	9.19**	5.17***
	(2.50)	(3.14)	(5.83)	(7.09)	(2.50)	(7.88)

Note: This table reports stock market returns for portfolios sorted by mpk. r^e_t denotes equal-weighted contemporaneous annualized monthly excess stock returns (over the risk-free rate) measured in the year of the portfolio formation from January to December of year t. r^e_{t+1} denotes the analogous future returns, measured from July of year $t+1$ to June of year $t+2$. t-statistics in parentheses, computed using Newey-West standard errors. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Predictive regressions of MPK on Aggregate Risk Exposures

	(1)	(2)	(3)	(4)
β_{CAPM}	0.261***		0.198***	
	(1.03)		(1.03)	
β_{FF}		0.05**		0.03**
		(1.70)		(1.68)
Observations	1112	1091	1112	1091
R^2	0.009	0.006	0.059	0.036
F.E.	No	No	Yes	Yes

Note: This table reports the results of a panel regression of year-ahead *mpk* regressed on measures of firm exposure to aggregate risk. Each observation is a firm-year. F.E. denotes the presence of industry-year fixed effects. Standard errors are two-way clustered by firm and industry-year. t-statistics in parentheses. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Correlations of MPK Dispersion, the Price of Risk and the Business Cycle

	MPK	PD ratio	TFP	GDP
MPK	1.00			
PD ratio	-0.23	1.00		
TFP	-0.21	0.42	1.00	
GDP	-0.53	0.53	0.28	1.00

Note: This table reports time-series correlations of MPK dispersion, measures of the price of risk and the business cycle. MPK dispersion is measured as the within-industry standard deviation in *mpk*. The PD ratio is the aggregate stock market price/dividend ratio. GDP is log GDP and TFP is log TFP. We extract the cyclical components of GDP, TFP and the PD ratio using a one-sided Hodrick-Prescott filter.

Table 9: Industry-Level Dispersion in MPK, Expected Stock Returns and Beta

	(1)	(2)	(3)	(4)
$\sigma(\mathbb{E}[r])$	3.01***		1.49***	
	(10.12)		(3.44)	
$\sigma(\beta_{MKT})$		0.08**		0.05**
		(6.35)		(4.32)
$\sigma(\beta_{HML})$		0.11**		0.07**
		(5.38)		(2.79)
$\sigma(\beta_{SMB})$		0.12***		0.05
		(11.32)		(6.27)
N	298	301	295	300
R^2	0.23	0.25	0.27	0.29
Industries	33	35	32	32
F. E.	No	No	Yes	Yes

Note: This table reports a panel regression of the dispersion in mpk within industries on lagged measures of dispersion in risk exposure within those industries. An observation is an industry-year. $\mathbb{E}[r]$ is the expected return computed from the Fama-French model. β denotes the stock return beta on the Fama-French factors. t-statistics are in parentheses. Significance levels are denoted by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Portuguese synonyms used for scraping variables from the raw data

Variable	Portuguese synonym used
Value Added	*Valor Adicionado Bruto*
Capital	*Imobilizado*
	Ativo Total
Revenue from sales	*Receita de Venda de Bens e/ou Serviços*
Wage bill	*Pessoal*
	Pessoal e encargos
	Pessoal e Encargos
Energy and materials	*Materiais, Energia, Servs. de Terceiros e Outros*
	Materiais, Energia e Outros
	Matérias Primas Consumidas
	Matérias-primas consumidas
	Materias primas consumidas
	Materiais e OutrosMaterial
Capex	*capex*
	Capex
	CAPEX

Note: This table reports the exactly synonyms used to scrap the accounts from CVM's balance sheet data, given the lack of a perfect standardization.

Table 11: B3 sectors and subsectors

Sector	Subsector
Oil, Gas and Biofuels	Oil, Gas and Biofuels
Basic Materials	Mining
	Steel and Metallurgy
	Chemicals
	Wood and Paper
	Packaging
	Diversified Materials
Capital Goods and Services	Construction and Engineering
	Transportation Equipment and Components
	Machines and Equipments
	Transportation
	Diversified Services
	Retail
Consumer Non Cyclical	Farming
	Food Processors
	Beverage
	Personal Care and Cleaning Products
	Retail and Distribution
Consumer Cyclical	Heavy Construction
	Textiles, Apparel and Footwear
	Household Products
	Automobiles and Motorcycles
	Hotels and Restaurants
	Travel, Entertainment and Leisure
	Diversified
	Retail
Health	Pharmaceutical and Others Products
	Medical and Hospitalar Services
	Equipments
	Retail and Distribution
Information Technology	Hardware and Equipments
	Software and Services
Communications	Telecommunications
	Media
Utilities	Eletric Utilities
	Water Utilities
	Gas Utilities
Financial	Financial Intermediaries
	Asset-backed Securitization
	Diversified Financial Services
	Insurance, Life and Multi-line
	Real Estate
	Holdings
	Other Bonds
Others	Others

Appendix

A Solution for the baseline model without adjustment costs

This appendix presents the solution for the model developed in Section 6 for when there are no adjustment costs. First, we solve:

$$\max_{N_{it}} e^{\hat{z}_{it}+\hat{\beta}_i x_t} K_{it}^{\theta_1} N_{it}^{\theta_2} - W_t N_{it},$$

which gives the following F.O.C.:

$$N_{it} = \left(\frac{\theta_2 e^{\hat{z}_{it}+\hat{\beta}_i x_t} K_{it}^{\theta_1}}{W_t} \right)^{\frac{1}{1-\theta_2}}.$$

And, substituting for the wage with $W_t = X_t^{\omega}$ and rearranging, we get

$$\Pi_{it} = G e^{\beta_i x_t + z_{it}} K_{it}^{\theta},$$

where $G \equiv (1-\theta_2)\theta_2^{\frac{\theta_2}{1-\theta_2}}$, $\beta_i \equiv \frac{1}{1-\theta_2}(\hat{\beta}_i - \omega\theta_2)$, $Z_{it} \equiv \hat{Z}_{it}^{\frac{1}{1-\theta_2}}$ and $\theta \equiv \frac{\theta_1}{1-\theta_2}$, which is the Equation 29 in the text, i.e., firm's operating profits.

Now, the first order and envelope conditions associated with 11 give the Euler equation:

$$\begin{aligned}
1 &= \mathbb{E}_t[M_{t+1}(\theta e^{\beta_i x_{t+1}+z_{it+1}} G K_{it+1}^{\theta-1} + 1 - \delta)] \\
&= (1-\delta)\,\mathbb{E}[M_{t+1}] + \theta G K_{it+1}^{\theta-1}\,\mathbb{E}_t[e^{m_{t+1}+z_{it+1}+\beta_i x_{t+1}}].
\end{aligned}$$

Substituting for m_{t+1} and rearranging,

$$\begin{aligned}
\mathbb{E}_t\left[e^{m_{t+1}+z_{it+1}+\beta_i x_{t+1}}\right] &= \mathbb{E}_t\left[e^{\log\rho - \gamma_t\varepsilon_{t+1} - \frac{1}{2}\gamma_t^2\sigma_\varepsilon^2 + z_{it+1} + \beta_i x_{t+1}}\right] \\
&= \mathbb{E}_t\left[e^{\log\rho + \rho_z z_{it} + \varepsilon_{it+1} + \beta_i\rho_x x_t + (\beta_i\gamma_t)\varepsilon_{t+1} - \frac{1}{2}\gamma_t^2\sigma_\varepsilon^2}\right] \\
&= e^{\log\rho + \rho_z z_{it} + \beta_i\rho_x x_t + \frac{1}{2}\sigma_{\tilde{\varepsilon}}^2 + \frac{1}{2}\beta_i^2\sigma_\varepsilon^2 - \beta_i\gamma_t\sigma_\varepsilon^2}
\end{aligned}$$

and

$$\mathbb{E}_t[M_{t+1}] = \mathbb{E}_t\left[e^{\log\rho - \gamma_t\varepsilon_{t+1}} - \frac{1}{2}\gamma_t^2\sigma_\varepsilon^2\right] = e^{\log\rho - \frac{1}{2}\gamma_t^2\sigma_\varepsilon^2 - \frac{1}{2}\gamma_t^2\sigma_\varepsilon^2} = \rho$$

so that

$$\theta G K_{it+1}^{\theta-1} = \frac{1-(1-\delta)\rho}{e^{\log\rho + \rho_z z_{it} + \beta_i\rho_x x_t + \frac{1}{2}\sigma_{\tilde{\varepsilon}}^2 + \frac{1}{2}\beta_i^2\sigma_\varepsilon^2 - \beta_i\gamma_t\sigma_\varepsilon^2}}$$

and, rearranging and taking logs,

$$k_{it+1} = \frac{1}{1-\theta}\Big(\tilde{\alpha} + \frac{1}{2}\sigma_{\tilde{\varepsilon}^2} + \frac{1}{2}\beta_i^2\sigma_\varepsilon^2 + \rho_z z_{it} + \beta_i\rho_x x_t - \beta_i\gamma_t\sigma_\varepsilon^2\Big),$$

where

$$\tilde{\alpha} = \log\theta + \log G - \alpha$$

$$\alpha = -\log\rho + \log(1-(1-\delta)\rho) = r_f + \log(1-(1-\delta)\rho).$$

Finally, ignoring the variance terms gives Equation 31.