Cyclical Fluctuations, Financial Shocks, and the Entry of Fast-Growing Entrepreneurial Startups*

Christoph Albert[†1] and Andrea Caggese[‡2]

[1] *CEMFI*

[2] *UPF, CREI and Barcelona GSE*

June 18, 2019

Abstract

We analyze a multi-year, multi-country entrepreneurship survey with more than one million observations to identify startups with low- and high growth potential, and we confirm the validity of these ex-ante measures with ex post firm-level information on employment growth. We find that negative aggregate financial shocks reduce all startup types, but their effect is significantly stronger for startups with high growth potential, especially when GDP growth is low. Our results uncover a new *composition of entry* channel that significantly reduces employment growth and is potentially important for explaining slow recoveries after financial crises.

*A previous version of this paper was entitled "Financial Frictions, Cyclical Fluctuations, and the Growth Potential of New Firms". For their useful comments and suggestions, we thank Davide Furceri, Peter Karadi, Geert Mesters, Kirill Shakhnov (discussant) and the participants at the 2018 Annual Conference of the Society for Economic Dynamics in Mexico City, the 6th ECB-CBRT Conference in Izmir, Turkey, the 2019 Workshop in Macroeconomics in Marrakech, the 2019 Catalan Economic Society Conference in Barcelona and the Seminars at Manchester University, Danmarks Nationalbank, and CREI. We thank Cristiano Mantovani for his excellent work as a research assistant. All errors are our own. The financial support from La Caixa Foundation, grant n.LCF/PR/RC14/10080001, is gratefully acknowledged. Andrea Caggese acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2015-0563)

†Correspondence: christoph.albert@cemfi.es

‡Correspondence: andrea.caggese@upf.edu

1 Introduction

A well established literature has documented the importance of financial frictions for entry into entrepreneurship and for the survival and growth of new firms (among others, see Holtz-Eakin et al, 1994; Blanchflower and Oswald, 1998; Corradin and Popov, 2015; Schmalz et al., 2015; Adelino et al. 2018). However, less is known about the relation between financial factors, the decision of what type of business to start, and the ex post performance of new firms. Haltiwanger et al. (2016) show that while most new firms grow slowly, a small fraction grow very rapidly, driving a higher mean net employment growth for younger firms than for older firms. Pugsley, Sedlaceck and Sterk (2018) argue that such heterogeneity is primarily driven by the ex ante characteristics of these startups rather than by the ex post shocks they face during their lifetime.

Are these ex ante decisions of the entrepreneurs important for the ex post ability of their businesses to create jobs? Additionally, do financial factors affect these ex ante decisions? This paper provides new evidence and an answer to these questions by combining multiple data sources. Our main dataset is drawn from the Global Entrepreneurship Monitor (GEM), a multi-country survey of entrepreneurial decisions that allows us to identify heterogeneous startup types. We use a sample of this survey that ranges from 2002 to 2013 and includes a total of approximately one million individual-level observations from 21 OECD countries. We merge this dataset with firm-level data, which allows us to measure the ex post performance of these different startup types, and we employ a country-specific business cycle indicator (GDP growth) and several macroeconomic indicators of financial conditions, which have been shown to strongly affect the availability of credit to households and businesses.

Three features make the GEM dataset particularly suited for our purpose. First, it includes individual characteristics such as age, gender, education, income bracket and entrepreneurial experience. Thus, we can study the dynamics of startups while controlling for the quality of the pool of potential entrepreneurs. Second, it is designed to be representative of a country's population and contains harmonized data across countries. Poschke (2018) shows that the firm size distribution obtained from GEM by using survey responses from entrepreneurs matches remarkably well with that obtained from administrative data sources. Third, it includes survey questions such as items ascertaining the expected employment growth of new startups and the innovative nature of the products and services that will be offered, which are helpful in identifying startups with growth potential.

In order to formalize the intuition of the relation between financial frictions and startup selection, we develop a stylized partial equilibrium model in which new entrepreneurs start a business by paying an initial sunk cost that is financed partly with their own wealth and partly with debt for which they pay a premium over the market interest rate. This premium reflects the excess cost of external finance caused by financial frictions. The entrepreneurs can choose between two different types of businesses: Type 1 represents a business model that is reliable and immediately profitable but with limited growth potential, for example, a business model in which the entrepreneur decides to provide well-established services and/or products in well-known markets. Type 2 represents the decision to provide a newer product or service and/or one in less well-known markets. The Type 2 business is initially not as productive as the Type 1 business but has a much larger growth potential in the medium-long term. The entrepreneurs are heterogeneous in their ability to manage these different businesses: in equilibrium, for the marginal entrepreneur who is indifferent between the two types, Type 2 has lower profitability in the short term and higher profitability in the long term. It follows that at the margin, it takes longer to repay the initial debt to finance a Type 2 startup, and its value is more sensitive to short-term increases in the cost of external finance than that of a Type 1 startup.

These results imply that, conditional on aggregate conditions and the quality of the entrepreneurial pool, an increase in the excess cost of finance will reduce the number of all startups and will reduce the number of Type 2 startups by relatively more than that of

Type 1 startups. Moreover, the results imply a financial accelerator channel that operates via the creation of new startups. By reducing the disposable income of entrepreneurs, a decline in GDP growth increases the need for external finance and amplifies the negative effects of financial shocks relatively more for Type 2 startups than for Type 1 startups.

In order to test these predictions, we identify Type 2 startups in the GEM dataset as those businesses for which the entrepreneur is expecting high future employment (relative to that for the average size of firms in its country/sector). A key part of our analysis is that we verify whether this ex-ante entrepreneurial selection of types is able to predict faster ex post firm growth. Conducting this type of test using only the GEM survey, which is a repeated cross-section, is unfeasible. Therefore, we match it at the 2-digit sector level with a sample obtained from the Sistema de Análisis de Balances Ibéricos (SABI) comprising all new firms founded since 2003 in Spain. The matched sample has 46 2-digit sectors and 226,954 firm-year observations. We link each firm in SABI with the share of startups with high growth potential in its sector in the year it was founded. We interpret this value as the probability that this firm is a high-growth firm. We find that the higher this ex ante probability is, the smaller the initial employment for new firms but the faster the employment growth over time: this faster employment growth results in the high-growth firms having a significantly larger size from six years of age onward. This result is robust to controlling for sector-year fixed effects and for the aggregate conditions at the time of the firms' entry, and therefore, it is not driven by sector- or time-specific factors. In other words, this finding provides a positive answer to our first question. The ex ante decisions of the entrepreneurs on the type of startup significantly affect the ex post ability of these businesses to create jobs.

Despite being limited to Spain, the matched firm-level dataset is sufficiently representative for our purposes. Spain is the country with the most extensive coverage in GEM, with more than 200,000 observations. Moreover, the Spanish economy was affected by large fluctuations in financing conditions during the sample period. Indeed, all the main results we later obtain from the entire GEM dataset are also confirmed when considering only the Spanish surveys.

After verifying the validity of our empirical measure of ex ante high growth potential, we provide an answer to our second question by testing the predictions of the model. Financial shocks are measured with fluctuations in the excess cost of external finance. Our preferred indicator is the Gilchrist and Zakrajsek (2012) bond spreads of financial institutions. Using data on European countries from Gilchrist and Mojon (2016), we compute the indicator for the US, Spain, Italy, France and Germany. Gilchrist and Mojon (2016) show that such spreads are good proxies for credit availability to households and firms and have strong predictive power for the real effects of financial crises. Therefore, they are ideal measures of the intensity of financial frictions affecting new startups. We also check that the results are robust to using an alternative measure of financial frictions, such as the financial distress indicators of Laeven and Valencia (2013) and of Romer and Romer (2017).

Our main results confirm the model's hypotheses. We find that conditional on GDP growth and individual characteristics, all startups are negatively affected by financial shocks but that high-growth startups are affected much more than low-growth ones. Moreover, we find a strong interaction between financial frictions and GDP growth: with lower GDP growth, the negative effect of financial shocks on startups with high growth potential becomes more amplified than the negative effect of financial shocks on low-growth startups.

We provide several robustness checks of these results. We might be overestimating the importance of financial shocks if the observed fluctuations in the cost of external finance are caused by aggregate productivity shocks that directly affect startup decisions. To control for this possibility, we compute the bond spreads predicted by the exogenous monetary policy shocks identified with high-frequency financial data by Jarocinski and Karadi (2018). These authors separately identify exogenous monetary policy shocks from shocks about new information from the Central Bank on the state of the economy. Therefore, such monetary policy shocks potentially affect the availability of credit and the bond spreads but are by construction orthogonal to contemporaneous shocks to investment opportunities. Using the predicted bond spreads instead of the actual spreads as an indicator of financial frictions confirms all our results.

As an additional test of our hypothesis, we consider two indicators often used in the literature to select the sectors more likely to face financial frictions: the external financial dependence indicator (Rajan and Zingales, 1998) and an indicator of intangibility (the share of intangible over total assets; see Falato et al., 2013, and Caggese and Perez, 2017). The model predicts that startups in sectors with higher indicators should be more negatively affected by financial frictions, and we confirm these predictions in the data.

Furthermore, our results are also confirmed when we control for variations in the real interest rate, which could affect differently the expected value of the different startups, when we exclude countries that did not experience the financial crisis, when we exclude selected sectors that might cause a spurious correlation, and when we include expectations about future business opportunities as an additional control variable.

Taken together, our results strongly support the view that financial frictions have different effects on the entry of firms with high growth potential and that this *composition of entry* channel is important for explaining slow recoveries after financial crises, which imply highly persistent output losses, as shown by Cerra and Saxena (2008). Abstracting from the general equilibrium effects on wages and prices, our results imply that an increase in the bond spread by one percentage point during a recessionary period changes the nature of firms created such that after ten years, there is 3.5% less employment in these firms.

The remainder of the paper is organized as follows. Section 2 outlines the related literature. Section 3 introduces a partial equilibrium model of the relationship between access to finance and entrepreneurial decisions. Section 4 describes the data. Section 5 conducts the empirical analysis and conducts tests of the model predictions. Section 6 presents some robustness checks. Section 7 concludes the paper.

2 Related literature

This paper is related to the large literature documenting the importance of financial constraints as a key factor influencing entry into entrepreneurship. Among others, Holtz-Eakin et al. (1994) and Blanchflower and Oswald (1998) show that consistent with the role of financial frictions in influencing startup business entry, financial wealth is an important determinant of entrepreneurial success. More recently, several authors emphasize the importance of housing wealth. Adelino et al. (2015) document that, controlling for demand factors, small businesses in areas with greater increases in housing prices experienced stronger growth in employment than did large firms in the same areas. Corradin and Popov (2015) show that housing wealth helps to alleviate credit constraints for potential entrepreneurs by enabling homeowners to extract equity from their property and invest it in their business. Schmalz et al. (2017) show that individuals affected by positive exogenous shocks to the collateral values of their properties are more likely to become entrepreneurs and, conditional on entry, use more debt, start larger firms, and remain larger in the long term. Robb and Robinson (2014) document that the most frequent source of financing of new firms is bank debt and that it is more extensively used in regions where supply is higher due to more home loans. Krishnan et al. (2014) show that firms that have better access to financing subsequently experience a higher growth in their productivity, especially if the firms were financially constrained. Hombert and Matray (2016) find that negative shocks to bank-firm lending relationships led to tighter financial constraints for small, innovative firms with more intangible projects and therefore negatively affected overall innovation activity. Deriving firm dynamic models in which financial constraints affect entrepreneurial entry, other authors show that such frictions are important to explain cross-industry and cross-country differences in aggregate productivity (among others, see Buera et al., 2011, Caggese and Cunat, 2013, Midrigan and Xu, 2014, and Cole et al., 2016).

We contribute to this literature by identifying the effects of financial conditions and their interaction with the business cycle on heterogeneous startup types. We provide new evidence that financial frictions not only affect the entry into entrepreneurship but also the type of business started, especially during recessions. These findings uncover a *composition of entry* channel that could contribute to explaining slow recoveries after financial crises; therefore, our paper is also related to studies of firm dynamics during the great recession. Clementi and Palazzo (2016) show that the sharp decline in the number of startups during the 2007-2009 recession might have contributed to the slow recovery, and Siemer (2018) emphasizes the importance of financial frictions in this decline. Pugsley and Sahin (2018) find that the decline in firm entry in the last decades contributed to a lower trend in employment growth and to the occurrence of jobless recoveries.

Our work is especially related to Sedlaceck and Sterk (2017), who show that not only did firm entry strongly decline during the 2007-2009 financial crisis but also that the startups that did enter during that period were significantly weaker in their potential to create jobs. In their model, these authors emphasize the importance of ex ante entry decisions. However, their empirical analysis focuses solely on firm-level data. Conversely, we analyze a rich cross-country survey of entrepreneurial choices and are able to study how financial factors affect the entrepreneurial decisions to create different types of businesses, while controlling for the quality of the entrepreneurial pool.

Entrepreneurial choices among heterogeneous individuals have also been extensively analyzed in the occupational choice and innovation literature (see, e.g., Poschke, 2013). Moreover, because of its focus on high-growth startups, our paper is related to the literature that emphasizes the importance of transformational entrepreneurs (Schoar, 2010) and to recent papers that identify the characteristics of these entrepreneurs (Brown et al., 2018; Azoulay et al., 2018). Other authors focus on the mobility of inventors and disruptive innovators and on the reallocation of highly skilled labor (see, among others, Acemoglu, Akcigit and Celik, 2014, and Akcigit, and Kerr, 2016).

Finally, our empirical analysis is related to those studies, in particular Braun and Larrain (2005), Kroszner, Laeven and Klingebiel (2007), and Dell'Ariccia, Detragiache and Rajan (2008), that use multi-country and multi-sector data to analyze the effect of financial factors on the cyclicality of economic activity. These studies use sector-level data, while we analyze the dynamics of heterogeneous startups by using entrepreneur-level information.

3 Model

In this section, we develop a stylized partial equilibrium model of the relationship between access to finance and heterogeneous startup decisions. The model has two key elements. First, potential entrepreneurs have insufficient wealth to finance their new startup, and external finance is costly, especially during financial crises. We introduce financial frictions in the model as an additional cost of borrowing, and in our empirical analysis, we identify it with the bond spreads of financial institutions. Gilchrist and Mojon (2016) show that such spreads are good proxies for household and firm credit availability. As described above, a large body of literature suggests that new entrepreneurs are financially constrained, and their need for external funds is confirmed in our dataset, where entrepreneurs finance on average around 50% of their startup costs with external financing sources (see Figure 12 in the Appendix). Note that the above-mentioned literature emphasizes the role of house prices. In our model, we assume that an increase in the cost of external finance increases the cost of borrowing for new entrepreneurs. A house prices channel could be introduced in the model by assuming that a higher cost of external finance increases the costs of mortgages and reduces house prices, and the collateral available to new entrepreneurs. This alternative channel of financial frictions would generate very similar results to those derived below, and therefore, we choose to keep the analysis simpler and not model this additional channel.

Second, potential entrepreneurs can choose different types of projects with different growth prospects. In Section 5.1, we confirm that our dataset is able to identify these different startup types.

Technology

Consider many risk-neutral entrepreneurs who can choose the type of startup j among N alternatives, with types indexed by $j = 1, 2, ..., N$. All types require the same initial sunk cost κ to operate. Once they begin to produce, every period, they face a liquidation probability d.[1] A continuing Type j firm generates output:

$$y_t = \left(\theta_t^j\right)^{\beta} l_t^{\alpha} \tag{1}$$

where l is labor input, $0 < \alpha < 1$, and $0 < \beta \leq 1$. One unit of labor costs an exogenous wage w.[2] Profits are:

$$\pi_t = \left(\theta_t^j\right)^{\beta} l_t^{\alpha} - w l_t \tag{2}$$

To keep the model tractable, we assume that wages are paid after earnings are realized and thus not subject to financial frictions and that $\beta = 1-\alpha$. Therefore, the labor demand that maximizes profits is:

$$l_t = \left(\frac{\alpha}{w}\right)^{\frac{1}{1-\alpha}} \theta_t^j \tag{3}$$

Substituting l_t in Equation 2, we express profits as a linear function of θ_t:

$$\pi\left(\theta_t^j\right) = \Psi \theta_t^j \tag{4}$$

$$\Psi \equiv \left[\left(\frac{\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}} - \left(\frac{\alpha}{w}\right)^{\frac{1}{1-\alpha}} w\right] > 0$$

Startup types differ in their expected productivity growth. We simplify the analysis and set $N = 2$:

Type 1 indicates a startup with low-growth potential and for which productivity θ_t^1 grows at an exogenous rate $g_t^1 = g^{med}$ in all periods. Starting a Type 1 business represents

[1] Given the partial equilibrium nature of the model, this can be interpreted as the probability of death of the entrepreneur or as the probability that the firm becomes obsolete and its productivity falls permanently to zero.

[2] θ_t^j can be interpreted literally as efficiency or as shorthand for quality improvements that increase demand. Similarly, $\alpha < 1$ can be interpreted as decreasing returns to scale or as shorthand for monopoly power.

the decision to provide mature and established products or services and/or products in well-known markets. This decision to start a Type 1 business has low risk and will result in immediate profits; however, the business also has low growth prospects.

Type 2 indicates a startup with high growth potential. $g_0^2 = g^{low} < g^{med}$ initially, but every year, with probability γ, g_t^2 permanently increases from g^{low} to $g^{high} > g^{med}$. Starting a Type 2 business represents the decision to provide a newer product or service and/or one in less well-known markets. The decision is riskier, and more time is required for the business to start generating revenues; however, the business has higher growth potential.[3]

Financing

The entrepreneur has an initial endowment of $a \leq \kappa$ and needs to borrow $b = \kappa - a$. In subsequent periods, debt can be repayed by using the flow of profits $\pi(\theta)$. One unit of debt implies a repayment of $\frac{1+r^b}{1-d}$ next period, which reflects the risk that the firm is liquidated before producing and is unable to repay the debt with probability d. We normalize the interest rate to zero, and therefore, r^b can be interpreted as the financial spread or excess cost of debt caused by financial frictions. Below, we derive the value of the business under frictionless finance and under financial frictions.

Value of the business without financial frictions

The access to finance is not a problem if either $a < \kappa$ but $r^b = 0$, meaning that the entrepreneur can borrow at the market interest rate, or $r^b > 0$ but $a = \kappa$, meaning that access to finance is costly but the entrepreneur can self-finance the startup cost.

In this case, the value of a new business with initial productivity θ_0 is given by the discounted sum of the future expected revenues net of κ. First, consider a Type 1 firm. In every period, it might liquidate with a probability d. If it does not liquidate, it generates profits $\Psi\theta_t$, where θ_t grows at the rate g^{med}. As shown in Appendix A, the net present

[3] The growth potential of Type 2 projects might also depend on different managerial and organizational strategies. For example, a restaurant owner might choose whether to manage a small traditional family restaurant or to attempt to develop a new restaurant chain.

value of profits for a Type 1 startup with initial productivity θ_0 is equal to:

$$V^1\left(\theta_0^1\right) = (1-d)\Psi \frac{\theta_0^1}{d-(1-d)\,g^{med}} \tag{5}$$

Conversely, the net present value of a Type 2 firm can be shown to be equal to (see Appendix A for details):

$$V^2\left(\theta_0^2\right) = (1-d)\Psi\Phi \frac{\theta_0^2}{1-(1-\gamma)(1-d)\left(1+g^{low}\right)} \tag{6}$$

where:

$$\Phi \equiv (1-\gamma) + \frac{\gamma}{d-(1-d)\,g^{high}}$$

The value of a Type j startup is thus given by $V^j(\theta_0^j) - \kappa$.

Value of the business with financial frictions

Financial frictions matter if the entrepreneur needs to borrow $b = \kappa - a > 0$ to start the firm and if the external financing is costly (r^b>0). We denote $C^j(\theta_0^j)$ as the net present value of these expected excess financing costs for a new business with initial productivity equal to θ_0^j. The value of a Type j startup is thus given by $V^j(\theta_0^j) - C^j(\theta_0^j) - \kappa$. In the presence of these frictions, the entrepreneur uses all earnings to repay b as quickly as possible, and the law of motion of debt is:

$$b_{t+1} = \left(\frac{1+r^b}{1-d}\right) b_t - \pi(\theta_t^j) \tag{7}$$

For a Type 1 firm, we first compute n^*, the expected number of periods necessary to repay the debt. Then, we compute b^*, the amount that could be repaid in the same n^* periods in the absence of the excess cost of finance (see equations 19-21 in Appendix A). The difference between b^* and b is by construction the net present value of revenues that pay for the excess cost of financing the startup:

$$C^1 = b^* - b \tag{8}$$

The calculation of C^2 is slightly more complicated because of the stochastic nature of productivity growth for Type 2 firms, but it is possible to show that it can be approximated to:

$$C^2 = \sum_{t=0}^{n^e} \left[(1-d)(1-\gamma)\right]^t r^b b_t + \frac{\gamma}{1-\gamma} \sum_{t=1}^{n^e} \left[(1-d)(1-\gamma)\right]^t C(b_t, g^{high}, \theta_t^2, r^b)] \tag{9}$$

where $C(b, g^{high}, \theta_0^2, r^b)$ is the excess cost of finance conditional on debt b, productivity growth g^{high}, initial productivity θ_0^2, and the interest rate premium r^b. See Appendix A for details. n^e is the expected number of periods needed to repay the debt, and b_t is the residual debt after t periods.

Calibration

We introduce heterogeneity across entrepreneurs by assuming that the initial productivity θ_0 is a function of their skills:[4]

$$\theta_{0i}^j = \phi_i^j E_i. \tag{10}$$

The generic entrepreneurial skill E is uniformly distributed across entrepreneurs, $E_i \in [1-e, 1+e]$, with $0 < e < 1$. ϕ_i^j denotes the skill of entrepreneur i specific to type j projects. The skills required to operate Type 2 firms, ϕ_i^2, are uniformly distributed over the interval $\phi_i^2 \in [\phi_{min}, 1]$. Conversely, the skills required to operate Type 1 firms are $\phi_i^1 = 1$ for all entrepreneurs. In other words, the draw of E_i determines one's chances of starting any type of firm, while the draw of ϕ_i^2 determines the probability of starting a Type 2 firm rather than a Type 1 firm.

[4] An alternative assumption to consider heterogeneous growth rates of productivity across entrepreneurs would have similar implications.

For the analysis in the next sections, we consider the following calibration of the model's parameters. The probability of death d is equal to 0.05, yielding an average firm duration of 20 years. g^{med} is equal to 3%; therefore, the employment of Type 1 firms grows on average at 3% every year, consistent with the median employment growth rate of US firms.[5] For Type 2 firms, g^{low} is normalized to zero. Moreover, $g^{high} = 4\%$, and $\gamma = 20\%$, so that their resulting expected employment growth relative to Type 1 firms roughly matches the relative employment growth of the *high growth* startups we identify from matching the GEM and SABI datasets (see Section 4.1 for details). $\alpha = 0.6$ matches the labor share of output. The initial sunk cost κ is normalized to one, and the wage w is set equal to 1.2. As in Midrigan and Xu (2014), this value implies that profits for the average firm in the industry are four times larger than κ. The two remaining parameters, which determine the heterogeneity in startup values across entrepreneurs, are e and ϕ_{min}. We set $\phi_{min} = 0.2$, which roughly matches the high-growth to low-growth startup ratio of 0.5 that we find in the data (see Section 4.1 for details). The remaining parameter e determines the fraction of individuals choosing to be entrepreneurs: this is not the main focus of this exercise and does not significantly affect the rest of the analysis. We consider a benchmark value of $e = 0.7$, which generates a realistic sensitivity of overall entrepreneurship rates to financial frictions.

3.1 Access to finance and startup decisions

In this section, we analyze how entrepreneurial startup decisions are affected by two key variables: the entrepreneur's internal finances a and the excess cost of external finance r^b.

No financial frictions

If there are no financial frictions ($r^b = 0$ or $\kappa = a$), then it follows that $C^1 = C^2 = 0$,, and the entrepreneurs decide which business to start based exclusively on the value of V:

Proposition 1. *In the absence of financial frictions, there exist a threshold* $\overline{\phi}(E_i)$

[5]Source: the authors' own calculations by using Compustat data.

such that entrepreneurs with skills ϕ_i^2 above the threshold prefer a Type 2 startup and entrepreneurs below the threshold prefer a Type 1 startup.

The proof of the proposition is straightforward. Our calibration ensures that for a given value of E_i, $V^2 < V^1$ for entrepreneurs with $\phi_i^2 = \phi_{min}$, while $V^2 > V^1$ for the maximum value $\phi_i^2 = 1$. Moreover, equations 5, 6 and 10 imply that V^2 increases linearly in ϕ_i^2 for a given value of E_i. Thus, entrepreneurs with $\phi_i^2 > \overline{\phi}(E_i)$ will start a Type 2 firm if their generic skill E_i is sufficiently high that $V^2{>}0$ and will not start any firm if $V^2 \leq 0$. Entrepreneurs with $\phi_i^2 \leq \overline{\phi}(E_i)$ will start a Type 1 firm if $V^1{>}0$ and will not start any firm if $V^1 \leq 0$.

Financial frictions

If there are financial frictions, then both C^1 and C^2 are positive. Therefore, each entrepreneur selects the project with the highest net value $V^j - C^j - \kappa$. Panel I of Figure 1 shows that both C^1 and C^2 are increasing in the cost of external finance r^b.[6] Increasing financing costs reduces the net value of the startups and increases the minimum generic skills E required to start a business:

Proposition 2*. Conditional on financial wealth a, an increase in the cost of external finance r^b will reduce the frequency of all startups.*

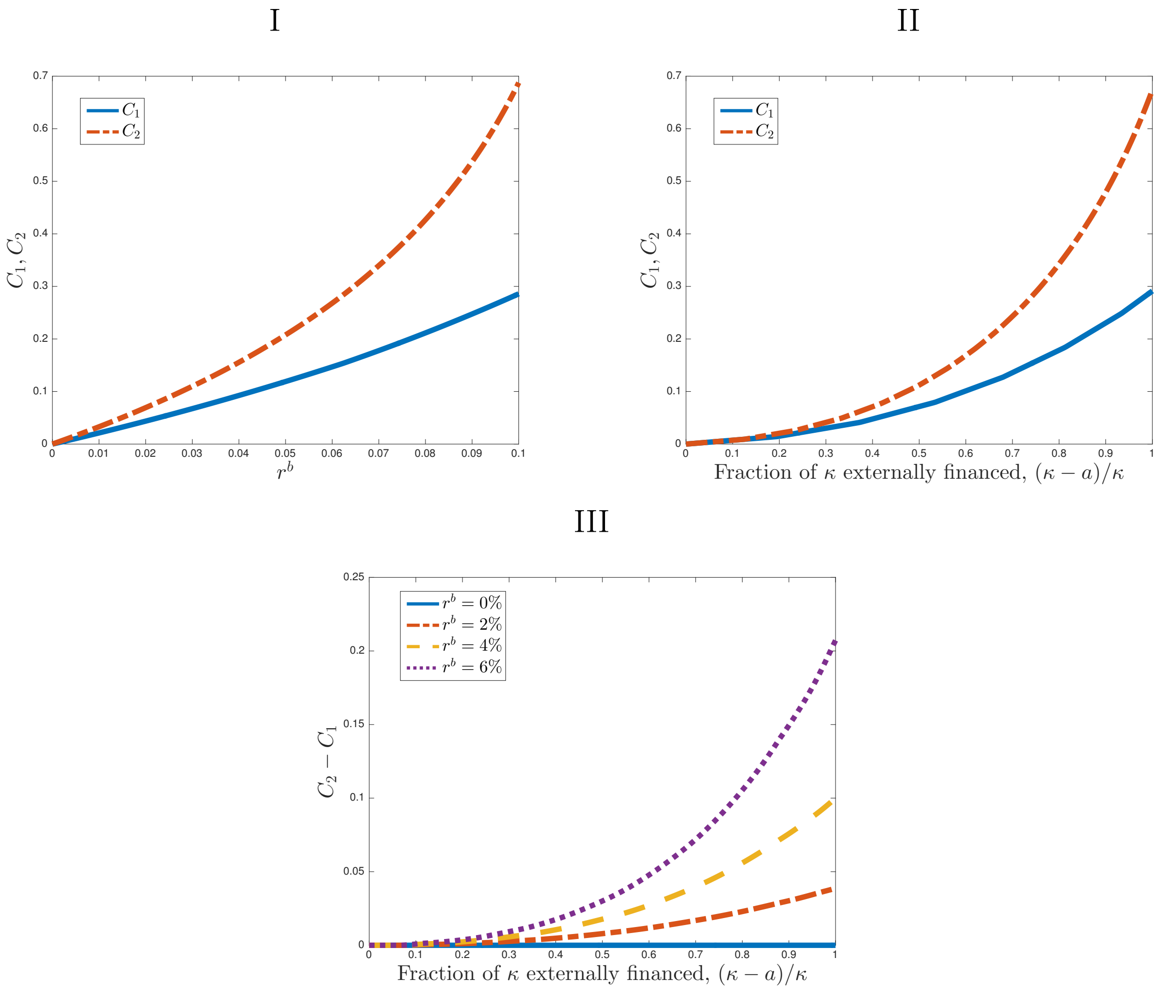
The finding that financial frictions reduce firm entry is not new in the literature. Therefore, the most novel part of our analysis is the derivation and testing of the predictions regarding the differential effects on heterogeneous startup types. In particular, Panel I of Figure 1 shows that C^2 is more sensitive to financial frictions than C^1, yielding the following prediction:

Proposition 3*. Conditional on financial wealth a, an increase in the cost of external finance r^b will reduce the number of Type 2 startups relatively more than that of Type 1 startups.*

The intuition is that when both startups have similar net present value, a Type 2

[6] This figure is generated by choosing ϕ^2 and E so that $V^1 = V^2$; i.e., entrepreneurs are indifferent between startups types when there are no financial frictions.

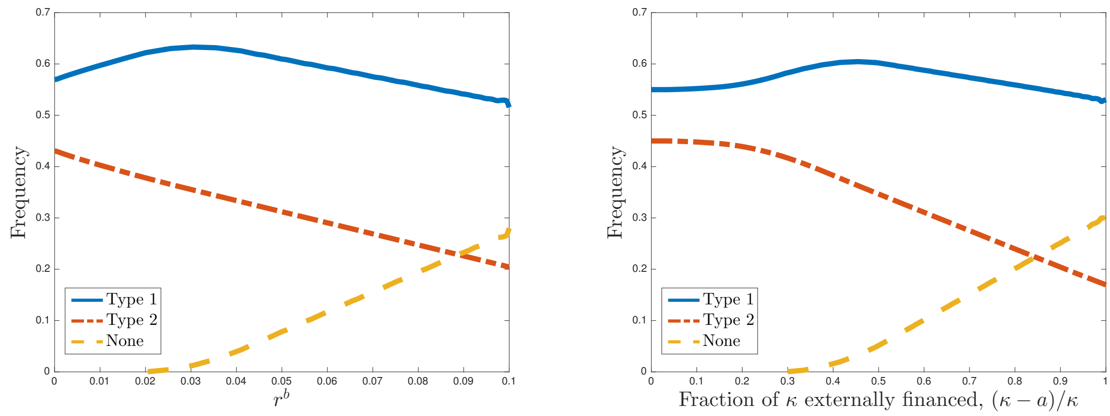
Figure 1: Properties



startup generates revenues further in the future and takes longer to generate sufficient earnings and repay the initial debt. As a consequence, with a Type 2 startup, the entrepreneur has to pay the high external financing costs for a longer period than with a Type 1 startup of similar value.

Proposition 3 is illustrated in Figure 2, which shows the frequency of each type of startup (or of not starting any business) among all potential entrepreneurs. In the left panel, as r^b on the x-axis increases from zero to a positive value, the frequency of Type 1 startups initially increases because some entrepreneurs with a value of ϕ^2 just above the threshold $\overline{\phi}$ switch from Type 2 to Type 1 startups. With a further increase in

Figure 2: Predicted frequencies of startup types



Notes: $(\kappa - a)/\kappa$ is set to 0.75, and r^b is set to 0.075 to generate the frequencies in the left and right panel, respectively.

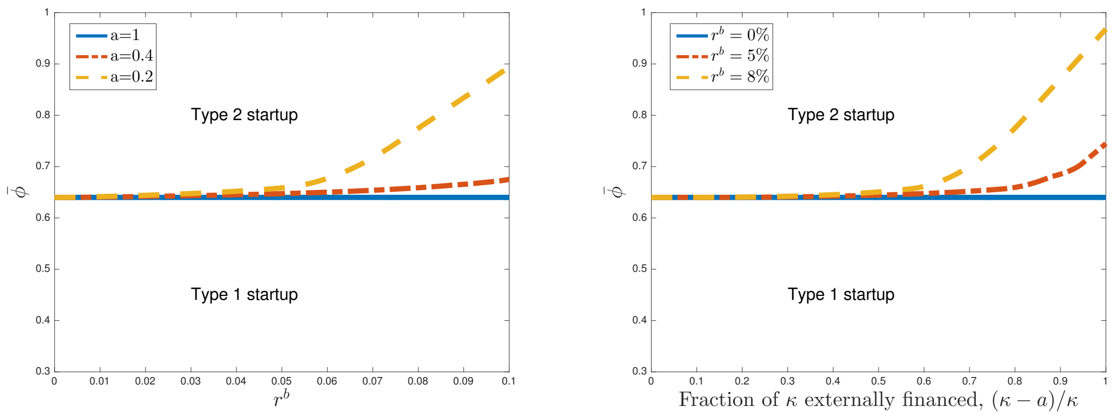
financing needs, the cost becomes so high that the entrepreneurs at the lower end of the distribution of E stop starting businesses; therefore, the frequency of Type 1 startups begins to decrease, although much less strongly than that for Type 2 startups, as stated in Proposition 3. The right panel shows a similar pattern for an increase in financing needs $\kappa - a$.

Thus far, we have considered in isolation changes in external financing needs and in the cost of external finance. Panel III of Figure 1 shows how they interact with one another: an increase in financing needs $\kappa - a$, which has no effect if $r^b = 0$, will instead progressively increase the funding costs of Type 2 more than those of Type 1 projects as r^b increases. The implications are summarized in the following proposition.

Proposition 4. An increase in external financing needs $(\kappa - a)$ *increases the negative effects of* r^b *relatively more on Type 2 startups than on Type 1 startups.*

Proposition 4 is illustrated in the left panel of Figure 3, which depicts the choice between startup types as a function of the cost of external finance r^b. The line is flat when financial wealth a is equal to 1, which is also the value of κ, such that financing needs are $\kappa - a = 0$. In this case, the threshold $\overline{\phi}$ is constant because the excess cost

Figure 3: Interaction between financing needs and the cost of external finance.



Each line represents the threshold value $\overline{\phi}$ given r^b and a. Entrepreneurs with $\phi^2 > \overline{\phi}$ prefer Type 2 projects to Type 1 projects. Both panels are generated for entrepreneurs with average skills of $E = 1$.

of finance is irrelevant to the choice of type of project, as stated in Prediction 1. The slope is slightly positive when financing needs are moderate (a=0.4, 60% of κ is financed with debt) and becomes very steep when financing needs are high (a=0.2, 80% of κ are financed with debt). The right panel of Figure 3 considers the symmetric case of varying $\kappa - a$ for given levels of r^b.

3.2 Predictions

In the empirical section, we directly measure r^b with several alternative indicators of financing conditions. Therefore, Propositions 2 and 3 imply the two following predictions, which we can test while controlling for aggregate business conditions as well as for individual entrepreneurial characteristics.

Prediction 1. Conditional on GDP growth and individual characteristics, an increase in the cost of external finance will reduce the frequency of all startups.

Prediction 2. Conditional on GDP growth and individual characteristics, an increase in the cost of external finance will reduce the number of Type 2 startups relatively more than that of Type 1 startups.

Moreover, following the financial accelerator literature, we assume that financing needs

$\kappa - a$ are negatively correlated to GDP growth. We interpret a as funds that are either accumulated from previous periods or derived from current earnings. Intuitively, during booms, individuals with entrepreneurial abilities have on average larger personal financial resources because they are more likely to be working and/or have a larger income stream than they would have during recessions.[7] Therefore, we can also test the following prediction.

Prediction 3. A decline in GDP growth increases the negative effects of r^b relatively more on Type 2 startups than on Type 1 startups.

Since the model is highly stylized, it is useful to discuss how other unmodeled factors might affect these predictions. Financial frictions are introduced as a wedge between the real interest rate and the borrowing rate. This type of wedge is tightly related to the bond spread that was used in the empirical section of this paper and that is widely used in the literature as a measure of the intensity of financial frictions. An alternative way of modeling these frictions would be to introduce collateral constraints or other forms of credit rationing and shocks that generate unexpected liquidity needs. This alternative framework would generate similar implications as those from the current model. High-growth startups would be more vulnerable to credit constraints that might force them to liquidate prematurely because they could not obtain financing after experiencing negative liquidity shocks. Therefore, tighter borrowing constraints would negatively affect high-growth startups more than they would affect low-growth ones.

One restrictive assumption of the model is that neither the riskless interest rate nor the law of motion of productivity θ_t^j is correlated with the business cycle. Given their different inter-temporal profiles, the two startup types would be affected differently by

[7] One might argue that this assumption is restrictive because the accumulation of financial wealth is very persistent over time and therefore less tightly correlated with the business cycle than is income. Nonetheless, we believe that this assumption is without loss of generality. On the one hand, empirical models of household precautionary saving show that households exhibit buffer stock behavior whereby their net financial wealth is highly sensitive to the income stream in the current and recent periods (e.g., Carroll, 2001). On the other hand, in our empirical analysis, we control for, among other things, the income group of the household within the country. These income groups are likely correlated with long-term household wealth, and thus, we control for the effects of wealth unrelated to business cycle fluctuations.

temporary fluctuations in the interest rate. In Appendix tables 33 and 34, we show that all the results are robust to controlling for country-specific riskless interest rates.

With respect to productivity, one alternative possibility is that the growth potential of projects is procyclical and that the initial value of θ_t^j, its growth rate g_t^j, and the probability γ are positively related to GDP growth. This is likely to reinforce the procyclicality of startups but should not affect predictions 1-3, which focus on the effects of changes in financing costs conditional on GDP growth.

More generally, it is possible that the observed fluctuations in the cost of external finance are caused by aggregate productivity shocks that directly affect startup decisions. In this case, the estimated effect of financial frictions would be biased upwards. To control for this possibility, in Section 5.4, we consider a robustness check in which we use financial shocks predicted by exogenous monetary policy shocks identified with high-frequency financial data (Jarocinski and Karadi, 2018).

Another important element excluded from the model is the consideration that financial frictions might differ across projects. Several theoretical and empirical papers argue that such frictions are stronger for Type 2 firms. These are firms that propose more innovative projects, are riskier and are more likely to be subject to asymmetric information and other financial frictions than Type 1 firms. On the one hand, in the model, this feature can be introduced by assuming that the excess cost of finance r^b is larger for Type 2 startups, and this assumption would of course reinforce the results described above. On the other hand, in Section 5.5, we exploit this feature of the model by considering sectorial indicators of the intensity of financial frictions and use them to provide additional testable predictions.

4 Data

4.1 GEM dataset

Our main data source is the GEM, the most comprehensive cross-country survey on entrepreneurial activity currently available (Reynolds et al., 2005). The GEM includes random samples of adult individuals from over 100 countries, with sample sizes ranging from approximately 1000 in some small countries to over 200,000 in Spain. The representativeness of this sample is confirmed by Poschke (2018), who shows that the firm size distribution obtained from GEM survey responses from entrepreneurs matches remarkably well the distribution obtained from administrative data sources. The period of the sample used for our analysis is 2002-2013.[8] As data on many of the smaller countries are available for only a few years, we clean the data by dropping countries with observations in fewer than nine years. This leaves 26 countries in our sample, with five (Argentina, Brazil, China, Latvia, and Peru) being non-OECD countries, which we also drop.[9] Thus, our final sample includes 21 countries and approximately one million individual observations. We use the following two survey questions to identify individuals starting a business (nascent entrepreneurs):

1. "*Over the past twelve months, to help start a new business, have you participated in any undertaking, such as looking for equipment or a location, organizing a startup team, working on a business plan, beginning to save money, or any other similar activity?*"

2. "*Will you personally own all, part, or none of this business?*"

An individual is classified as starting a business if he/she answers "yes" to the first question and "all" or "part" to the second question. Thus, a nascent entrepreneur must

[8] The survey began in 1999, but the first three years have fewer observations and variables; therefore, we include only the years 2002-2013.

[9] We eliminate these developing countries to limit cross-country heterogeneity in the data. However, their inclusion does not significantly change the results.

have been active in establishing a new business during the last year and own at least part of this business. Some studies (e.g., Koellinger and Thurik, 2012) impose the additional restriction that the business must not have paid salaries or wages for more than three months. However, we believe that this might lead to the exclusion of too many new nascent businesses; therefore, we relax this restriction.[10]

There are several additional questions regarding the kind of business an individual is starting. In particular, two questions directly attempt to identify businesses with the potential to grow. The first asks about the expected size of the firm five years into the future. The second asks whether the startup will introduce innovative products or services. The first question is more directly related to our model and more generally to the potential of new startups to create jobs. Therefore, we use this question to identify our benchmark category of high-growth startups. We classify a startup as having "*high growth potential*" if the number of employees expected by the entrepreneur in 5 years is larger than the average size of firms (as measured by the number of employees) that are at least 5 years old in the same 2-digit sector and country. A total of 34% of all startups are classified in this category.[11] All remaining startups are classified as having low-growth potential. Figure 8 and Table 17 in Appendix B show the sectorial distribution of startups.[12]

The question regarding future employees is intended to capture the expectations of the growth potential of the new firm. However, in practice, it might also reflect expectations

[10]Approximately 27% of nascent entrepreneurs in our sample report having paid salaries or wages for more than three months. The results remain qualitatively unaffected when we exclude them. Regarding demographic differences, the individuals starting a business were found to be somewhat younger (37 vs. 40 years), to more often have a post secondary education (46% vs. 40%) and to be female with a probability of 35%. Moreover, 89% of business starters are employed, and 12% already own an established business, whereas these percentages are 81% and 6%, respectively, among the remaining respondents.

[11]Alternatively, we define only those startups as high-growth for which the expected number of employees is twice as high as the average of firms at least 5 years old. This leads to a share of high-growth startups around 18%. The results are qualitatively robust to using this stricter definition.

[12]The survey provides information on the size of the initial startup investment. This is reported in Figure 12 in the Appendix. It also provides information on the share financed by the entrepreneur and the share provided by external sources. However, the entrepreneurs' own funds are not always derived from their savings: these are often the funds borrowed by the entrepreneur rather than by the firm (see literature review in Section 2). Therefore, this information is not useful for distinguishing between the entrepreneur's own savings and external financing.

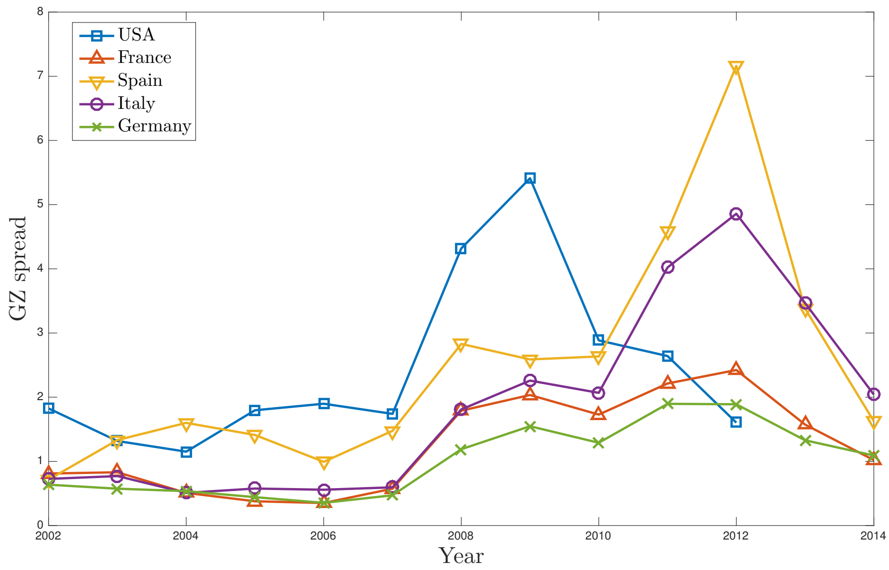
about the economy. For example, it could be that entrepreneurs are pessimistic during downturns and systematically underestimate the growth potential of their new firms. Alternatively, it could be that during downturns, entrepreneurs expect lower growth because all businesses, regardless of their nature, grow slowly. Both possibilities would negatively bias our measure of high-growth startups during such periods.

It is therefore important to verify that our indicator of high-growth startups provides information about the nature of the new business and not just about expectations of current and future market conditions. We verify this in two ways. First, in Section 6, we show that all our main results are confirmed when we control for a GEM survey variable that captures the expectations of the entrepreneurs about the state of the economy. Second, In Section 5.1, we verify the correlation between the probability to be a high-growth startup and the ex post growth of firms. This analysis is computed on firm-level data and controls for sector-specific year fixed effects, as well as for the state of the economy in the year the firm was born. We find that even conditional on this extensive set of controls, our measure of high growth potential startups predicts significantly faster firm employment growth in the future.

Finally, as an additional robustness check, in Section 6.3, we use additional survey questions from the GEM to identify entrepreneurs who plan to offer a product or service that is considered new by the potential customers and/or that embodies new technologies.[13] These startups, which we call *innovative*, might grow faster in the long run because new products or services have the potential to capture larger market shares. The regressions for this alternative classification broadly confirm our main results.

[13] We classify a startup as innovative if an entrepreneur responds "All" to the question "Will all, some, or none of your potential customers consider this product or service new and unfamiliar?" and "Less than a year" to the question "How long have the technologies or procedures required for this product or service been available?".

Figure 4: GZ spread by country



Notes: The figure plots the interest spread between the bonds of financial institutions and the risk-free rate based on Gilchrist and Zakrajsek (2012).

4.2 Business cycle and financial crisis data

In our empirical analysis, we use yearly data on GDP per capita from the Penn World Tables and calculate yearly real GDP growth rates (details are in Appendix B.2).[14]

The key variable to test our predictions is the excess cost of external finance r^b. We consider three empirical indicators related to it. The first is a country-year level financial crises dummy, which is based on systemic banking crises data from Laeven and Valencia (2013). According to their measure, 14 countries in our sample suffered a financial crisis, lasting from 2007 to 2013 in the US and the UK and from 2008 to 2013 in the remaining countries. There were no financial crises in Chile, Croatia, Finland, Japan and Norway.

Second, we consider a more detailed indicator of stress in the financial sector: the Gilchrist and Zakrajsek (2012) (GZ) bond spread of financial institutions. Using the data from European countries reported by Gilchrist and Mojon (2016), we compute the

[14]Alternatively, we used the deviation from the GDP trend as an indicator of business cycle conditions and obtained qualitatively similar results.

Table 1: Percentage of individuals starting a firm

	All	Low growth	High growth
Full	2.40	1.29	1.11
No Fin. crisis	2.81	1.47	1.35
Fin. crisis	1.84	1.06	0.78
% Difference	-34.52	-27.89	-42.22

indicator for the US, Italy, France and Germany (details are in Appendix B.4). Gilchrist and Mojon (2016) show that such spreads are good proxies for credit availability to households and firms and have strong predictive power for the real effects of financial crises. Therefore, they are ideal measures of the intensity of financial frictions affecting new startups. Figure 4 shows the evolution of the measure by country over the sample period. It spikes in 2009 in the US and in 2012 in Spain and Italy, while it is only moderately elevated between 2008 and 2013 in France and Germany.

Third, in Section 6.1, we consider, as an alternative, the financial distress indicator of Romer and Romer (2017) (RR). On the one hand, the GZ spread is conceptually more tightly related to the excess cost of finance in the model. The RR indicator is explicitly designed to capture other factors of financial distress beyond high bond spreads. On the other hand, these other factors might presumably be important for the new firms' access to finance, and the RR indicator has the additional advantage of being available for almost all the countries in our dataset (details are in Appendix B.5). As expected, for the subset of countries with both indicators, the GZ spread and the RR indicator are tightly correlated, with the correlation coefficient being approximately 0.79.

Descriptive statistics are shown in Table 1. In terms of unconditional averages, the percentage of individuals starting a business is 34% lower during the financial crisis. The drop is larger for individuals starting new firms expecting high employment growth, the number of which falls by 42%. Accordingly, at approximately 28%, the drop is smaller in the complementary category of firms expecting low employment growth.

4.3 Firm-level dataset from SABI

The GEM dataset provides extensive information on the individuals starting new firms, but its repeated cross-sectional structure does not allow us to follow the performance of the individual firms over time. Therefore, we complement our data with a panel of Spanish firms from the SABI data, which contains the number of employees for nearly the entire universe of firms that were established in 2003 or later.

5 Empirical analysis

5.1 Firm-level analysis

In this section, we analyze the firm dynamics in the Spanish SABI dataset to verify whether the startups we identify with high growth potential are informative of the ex post ability of firms to create jobs. In particular, we are interested in identifying the growth potential deriving from the nature of the businesses and not simply caused by the market conditions that prevailed when the firms were created. The analysis of the Spanish data is sufficiently representative of the whole sample. Spain is the country with the largest coverage in the GEM survey, with approximately 235,000 observations in total and at least 16,000 yearly observations from 2003. Indeed, all the main results we later obtain on the entire GEM dataset are also confirmed when considering only the Spanish surveys (see Appendix C for details).

We cannot link the GEM and SABI datasets at the firm level, but we can do so at the industry level. Using the GEM data, we compute the variable $Share\_growth_{s,t}$, i.e., the share of *high-growth* startups in a 2-digit sector s in year t in Spain. This is computed for a total of 46 2-digit sectors listed in Table 17 in Appendix B.6. Then, we match this variable with the SABI data. Of the 344,869 firms in our SABI sample, we can match 226,954 to sectors of startups identified in the GEM, of which 186,341 provide data on employment. Therefore, for this subset of firms with employment data, we have the associated value of $Share\_growth_{s,t}$ in their sector and year of creation. For instance, if the share of startups classified as high growth in the retail trade sector in 2005 in Spain is 30% according to the GEM, this value is matched to all retail trade firms born in 2005 in the SABI data. We interpret this percentage as the likelihood of a firm being *high growth*. To ensure that we focus on entrepreneurial startups only, we eliminate subsidiaries of other companies and companies primarily owned by foreign shareholders. Furthermore, we eliminate companies that have more than 100 employees during the first year of existence (443 in total). Then, we estimate the following model:

$$Employment_{i,s,t} = \beta_0+\sum_{k=0}^{10}\beta_1^k age_{i,s,t}^k+\sum_{k=0}^{10}\beta_2^k age_{i,s,t}^k \cdot Share\_growth_{i,s}+\sum_{k=0}^{N}\gamma_k X_{i,s,t}^k+\varepsilon_{i,s,t}, \quad (11)$$

The dependent variable $Employment_{i,s,t}$ is either the employment level or the employment growth of firm i in sector s in year t. $Share\_growth_{i,s}$ is the share of high-growth startups in sector s in the year firm i was founded, and $age_{i,s,t}^k$ is a dummy variable equal to one if the firm is k years old in year t. Among the N control variables $X_{i,s,t}^k$, we include year and sector dummies and GDP growth in the year the firm was born interacted with the age dummies. A positive value of the coefficient β_2^k, which multiplies the product of $Share\_growth_{i,s}$ and $age_{i,s,t}^k$, means that the higher the probability of being *high growth* is, the faster the employment growth or the higher the employment level of firm i at k years of age.

The regression results are shown in Table 2 (β_2 coefficients only; the full set of β_1 coefficients is shown in Appendix Table 18). In columns 1 and 4, we control for sector and year dummies. Hence, we measure the age profile of firm employment growth (in column 1) and level (in column 4), controlling for all business cycle factors and sector-specific factors. In the first column, in their first two years of existence, employment growth is significantly lower for likely *high-growth* firms than for other firms. However, it becomes significantly higher from age four onward, except for the last two coefficients for ages 9 and 10, which are no longer significant. In column 4, the log employment level shows dynamics consistent with the findings of column 1. For newborn firms, a higher

Table 2: Share of high-growth startups at firm creation and employment from SABI

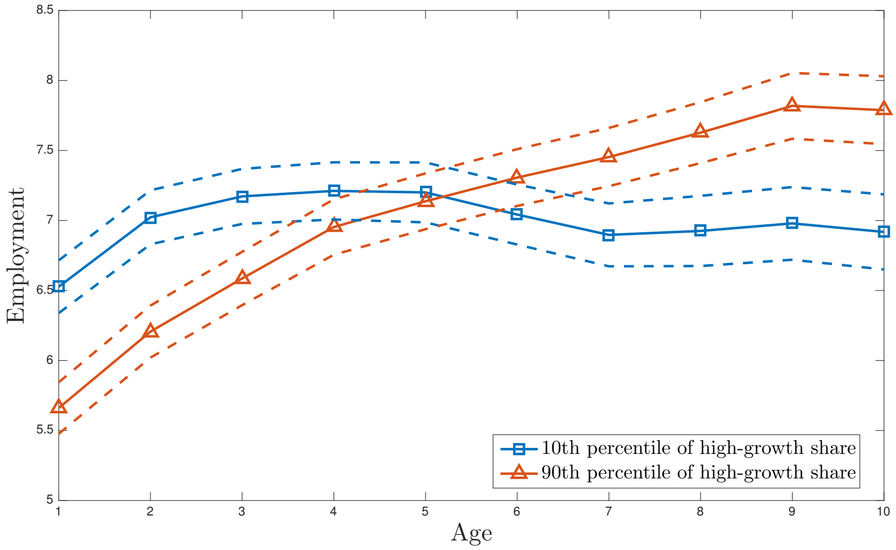
	(1) Empl. growth	(2) Empl. growth	(3) Empl. growth	(4) Employment	(5) Employment	(6) Employment
Age 0 x share				-1.284***	-0.588***	-0.681***
				(0.1081)	(0.1363)	(0.1384)
Age 1 x share	-0.315***	-0.206***	-0.225***	-1.812***	-0.959***	-0.910***
	(0.0355)	(0.0399)	(0.0406)	(0.1081)	(0.1238)	(0.1271)
Age 2 x share	-0.040**	0.060***	0.071***	-1.708***	-0.821***	-0.766***
	(0.0164)	(0.0201)	(0.0205)	(0.1230)	(0.1346)	(0.1353)
Age 3 x share	0.010	0.033**	0.037**	-1.227***	-0.639***	-0.589***
	(0.0132)	(0.0153)	(0.0155)	(0.1296)	(0.1412)	(0.1451)
Age 4 x share	0.057***	0.020	0.022	-0.540***	-0.418***	-0.407**
	(0.0129)	(0.0145)	(0.0147)	(0.1408)	(0.1539)	(0.1606)
Age 5 x share	0.026**	-0.001	0.001	-0.133	-0.379**	-0.369**
	(0.0132)	(0.0148)	(0.0149)	(0.1561)	(0.1709)	(0.1794)
Age 6 x share	0.052***	0.037**	0.034**	0.553***	0.084	0.117
	(0.0144)	(0.0158)	(0.0160)	(0.1744)	(0.1911)	(0.1951)
Age 7 x share	0.070***	0.047***	0.047***	1.163***	0.511**	0.533***
	(0.0146)	(0.0158)	(0.0158)	(0.1879)	(0.2049)	(0.2055)
Age 8 x share	0.065***	0.027	0.026	1.469***	0.618**	0.616**
	(0.0154)	(0.0179)	(0.0182)	(0.2373)	(0.2587)	(0.2521)
Age 9 x share	0.019	0.061***	0.059***	1.756***	0.904***	0.861***
	(0.0180)	(0.0207)	(0.0209)	(0.2648)	(0.2865)	(0.2861)
Age 10 x share	-0.020	0.057**	0.063**	1.819***	1.064***	0.979***
	(0.0227)	(0.0242)	(0.0246)	(0.2946)	(0.3155)	(0.3190)
Year FE	Yes	No	No	Yes	No	No
Sector FE	Yes	No	No	Yes	No	No
Year-sector FE	No	Yes	Yes	No	Yes	Yes
Age-growth interactions	No	No	Yes	No	No	Yes
Observations	706578	706578	706578	947696	947696	947696
R-squared	0.110	0.113	0.113	0.149	0.150	0.150

Notes: In columns 1-3, the dependent variable is the yearly employment growth of firms established in 2003 or later; 0.1% of the tails are winsorized. In columns 4-6, the dependent variable is the log employment level. *share* is the share of high-growth startups measured from the GEM data and that are in the 2-digit sector to which the firm belongs in the year it was born. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

share of high-growth startups is related to a smaller initial size. The correlation instead becomes positive from six years old onward, growing stronger as firms become older.

In columns 2 and 5, we control for sector-specific year effects. These dummies absorb factors that are common to all firms in each sector and year. While the estimated coefficients fall slightly in magnitude, all the results are confirmed, particularly the fact that likely high-growth firms are initially smaller but become significantly larger over time (column 5). In columns 3 and 6, we include an additional control variable, namely, the growth rate of GDP in the year the firm was born interacted with age. As explained above, this controls for the possibility that the share of high-growth startups does not

Figure 5: Predicted employment by age from SABI



Notes: The figure shows the predicted firm employment based on column 4 of Table 2 for different values of *share*. The 10th percentile is 18%, and the 90th percentile is 66%. The dashed lines show 90% confidence intervals.

measure the nature of the new businesses but rather the expectations related to the economy when the firm was born. Also in this case all the results are confirmed.[15].

Based on the specification in column 4, Figure 5 plots the paths of predicted firm employment over age depending on the share of high-growth startups in the sector and year the firm was born. We show the paths for the 10th and 90th percentiles of the share, which are 18% and 66%, respectively. Firms born when the share of high-growth startups was lower are predicted to be larger initially. However, after two years, firms born when the share was higher are predicted to grow faster and to eventually overtake the other firms after six years.

Note that these findings might be affected by selection. Perhaps high growth firms do not grow faster on average but have more volatile growth rates; in this case, our estimates may not capture the low growth rates of the firms that exit from the market. We believe that if present, such selection effects would imply a different interpretation of

[15] We also ran alternative specifications excluding firms born during the great recession (2008 and 2009) and obtained similar results.

the nature of these high-growth firms but would not necessarily reduce their importance since more innovative and riskier firms are more likely to introduce frontier technologies that are important for aggregate employment and productivity growth. Nonetheless, we can test the importance of selection effects by running a regression similar to Equation 11, where the dependent variable is $Exit_{i,s,t}$, a binary variable equal to 1 if firm i exits from the market in year t and zero otherwise. The estimation results (see the first two columns of Table 19 in Appendix C) show that the likelihood of being a high-growth firm slightly increases the exit risk only in the first year of existence, while it reduces such risk for firms between 2 and 8 years old. Thus, high-growth firms seem less risky than low-growth ones, and selection effects should not play a important role for the previous results. Importantly, columns 3 and 4 of Table 19 show that the likelihood of being an innovative startup (as defined in Section 4.1) instead generally increases its exit risk in the first four years of existence. Since innovative firms should be on average riskier, this finding is plausible and confirms the reliability of our matching between the entrepreneurship information in GEM and the firm-level information in SABI.

Overall, these results show a clear and statistically significant pattern: firms more likely to have been derived from a high-growth startup are initially smaller but have more potential to grow and become larger in the medium/longer term than do firms more likely to have been derived from low-growth startups. This finding does not seem to be driven by selection effects and is consistent with the behavior of Type 2 firms in the model, as well as with our claim that the high-growth startups indicator constructed in the GEM dataset is a valid measure of the intrinsic growth potential of these new firms and does not just capture market-related factors.

5.2 Individual-level analysis: estimation strategy

In this section, we test the predictions of the model by estimating how the propensity to start a business is related to financial conditions. Our baseline is the following probit model:

$$Pr(start_{i,j,t} = 1 | X_{i,j,t}) = \Phi(\beta_0 + \beta_1 bus_{j,t} + \beta_2 fin_{j,t} + \sum_{k=0}^{N} \gamma_k X^k_{i,j,t} + \varepsilon_{i,j,t}), \quad (12)$$

where $start_{i,j,t} = 1$ is a dummy indicating that individual i in country j in year t is starting a firm. $bus_{j,t}$ is a variable capturing the state of the business cycle in country j at time t, for which we use the real GDP growth rate in terms of purchasing power parity. $fin_{i,j,t}$ is the variable measuring shocks to the cost of external finance, for which we consider the three alternative measures described in detail in section 4.2. $X^k_{i,j,t}$ is a vector of N control variables including country dummies, sex, age and educational level.[16] We weight observations by using the weight variable included in the GEM.[17]

We estimate these models with a dummy for the start of any business as the dependent variable, as well as with dummies for starts in subcategories only. Because we control for individual characteristics, our analysis identifies how the propensity to start different types of businesses is affected by shocks to the cost of finance conditional on the quality of the potential entrepreneurial pool and the business cycle. Prediction 1 implies that β_2 should be negative when the dependent variable is all startups. Prediction 2 implies that β_2 should be more negative for high-growth startups than for low-growth ones. Furthermore, in order to test Prediction 3, we estimate a model that includes the interaction $bus_{j,t} \cdot fin_{j,t}$:

$$Pr(start_{i,j,t} = 1 | X_{i,j,t}) = \Phi(\beta_0 + \beta_1 bus_{j,t} + \beta_2 fin_{j,t} + \beta_3 bus_{j,t} \cdot fin_{j,t} + \sum_{k=0}^{N} \gamma_k X^k_{i,j,t} + \varepsilon_{i,j,t}). \quad (13)$$

Prediction 3 implies that β_3 should be positive, indicating stronger negative effects of financial frictions when GDP growth is lower. Furthermore, it implies that β_3 should be larger in absolute value for the high-growth startups than for the low-growth ones.

[16] In Section 6, we present the results with dummies for the income level (three categories). The information on the actual income level of respondents is not available in the GEM data. Instead, the GEM contains a variable that indicates whether a person in a specific year and country is in the lowest 33%, the middle 33% or the upper 33% of the income distribution of all respondents. Thus, by construction, this variable cannot control for income differences in the pool of entrepreneurs over time or across countries. We therefore choose not to include it as a control variable in the baseline regressions.

[17] According to the description of the GEM, the weights are "developed such that proportions of different subgroups (gender and age, for example) match the most recent official data descriptions of the population of a country." Our results are robust to not weighting the observations.

This estimation strategy requires that cyclical fluctuations and financing conditions are not perfectly correlated in the data, and we find that this is the case in our sample. The correlation between the GZ spread and GDP growth is -0.39, and that between the RR indicator and GDP growth is -0.40: the correlations are thus low enough that their effects can be separately identified. This is shown in detail in Appendices B.4 and B.5, where we report the scatterplots between GDP growth (deviations from country averages) and the values of the two indicators. These plots show a clear negative relation, which, however, is far from perfect due to many observations with high levels of financial frictions and medium or moderately high values of GDP growth.

Furthermore, although GDP growth and our financial shocks indicators provide independent sets of information, it is still possible that these financial shocks are themselves caused by investment opportunity shocks that directly affect startup decisions. In section 5.4, we control for this possibility and show that the results are confirmed when using predicted financial shocks that are orthogonal to investment opportunities shocks.

5.3 Individual-level analysis: baseline results

In Table 3, we show the results of the baseline probit model (12). In columns 1-3, the Laeven and Valencia (2013) financial crisis dummy is used as the indicator of financial shocks. In the first column, the dependent variable is any type of startup. The coefficient of GDP growth is positive and significant, indicating that startup creation is procyclical. On the other hand, the financial crisis indicator has a marginally significant negative effect on the probability to create any kind of startup. As shown in column 2, when we use an indicator for low-growth startups as a dependent variable, the effect of a financial crisis becomes insignificant. However, as seen in column 3, when we use an indicator for starting a high-growth startup, the effect of a financial crisis is significant and more negative than it is for all startups. In columns 4-6, we replace the financial crisis dummy with the bond spread of financial institutions (GZ spread). The *GDP growth* coefficient is

Table 3: Financial crisis, GZ spread and probability of starting a firm

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	1.723*	1.875**	0.942	3.242***	3.041***	2.729***
	(0.9961)	(0.8262)	(1.0024)	(0.6766)	(0.6986)	(0.4148)
Fin. crisis	-0.085*	-0.063	-0.110**			
	(0.0489)	(0.0437)	(0.0490)			
GZ spread				-0.027*	-0.018	-0.039***
				(0.0151)	(0.0155)	(0.0106)
Observations	894126	894126	894126	370280	370280	370280
R-squared	0.060	0.045	0.075	0.037	0.035	0.034

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

larger and more significant in columns 4-6 than in the first three columns. The difference is explained by the difference in the sample selection. The specification in the last three columns is estimated on a smaller subset of countries (the US, Spain, France, Germany and Italy), for which startups are more procyclical over the whole sample period. Again, we find that the coefficient of the financial frictions indicator is significant and more negative for high-growth startups and insignificant for low-growth startups.

Overall, the results in Table 3 confirm both Predictions 1 and 2. In terms of the marginal effects at the mean, during the financial crisis, low-growth startups are reduced by an (insignificant) 16% and high-growth startups by 31%. An increase in the GZ spread by one point decreases high-growth startups by 11% (versus 4.5% in the complementary group).[18]

In Table 4, we show the results of estimating model (13) with the interaction term between GDP growth and the financial shocks indicator. The GDP coefficient becomes insignificant in the first three columns. However, the *Financial crisis* dummy, which now indicates the effect conditional on GDP growth being zero, is more negative and more sig-

[18] We compute these semi-elasticities with the control variables being evaluated at their means. Since the covariates are demographic characteristics and not systematically correlated with the cycle, fixing them should not pose a major problem for the interpretation of the marginal effect of a financial crisis. We have verified that the marginal effect is not sensitive to the particular values at which the controls or GDP growth are evaluated.

Table 4: Financial crisis, GZ spread and probability of starting a firm

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	0.663	0.954	-0.012	5.447***	4.418**	6.182***
	(0.7114)	(0.6452)	(0.5684)	(1.9928)	(1.7647)	(1.6892)
Fin. crisis	-0.162***	-0.129***	-0.185***			
	(0.0516)	(0.0412)	(0.0628)			
Fin. crisis x GDP growth	4.679***	3.886***	5.093**			
	(1.7898)	(1.3950)	(2.5150)			
GZ spread				-0.020	-0.013	-0.033*
				(0.0197)	(0.0189)	(0.0173)
GZ spread x GDP growth				2.450	1.532	3.829**
				(1.6126)	(1.3356)	(1.5513)
Observations	894126	894126	894126	370280	370280	370280
R-squared	0.062	0.046	0.077	0.039	0.035	0.039

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

nificant than that in Table 3. Finally, the interaction between *GDP growth* and *Financial crisis* is positive and statistically significant. In general, a positive interaction coefficient indicates the greater cyclicality of startups during the financial crisis period. Moreover, since GDP growth was slower during the financial crisis than during the previous period, the positive interaction coefficient can also be interpreted as showing a significant slowdown in startups during the financial crisis for those countries that experienced larger contractions in GDP. Columns 2 and 3 show that both the financial crisis coefficient and its interaction with GDP growth are larger for the likely high-growth startups than for the complementary group, confirming Prediction 3. In terms of the marginal effects at the mean, a one-percentage point decrease in GDP growth during the financial crisis reduced high-growth startups by 14% (versus 10% in the complementary group) more than did a one-percentage point GDP decrease in the period outside the financial crisis.

In columns 4-6, the *GZ spread* coefficient, which again measures the effect conditional on zero GDP growth, is negative but not statistically significant, except for the high-growth startups in column 6. This result is consistent with the model, which predicts that the excess cost of finance has a significant effect on startup decisions only when the potential entrepreneurs' own financial wealth is very low. This might happen to many entrepreneurs during downturns, while it is less likely to happen to them during periods of flat or growing GDP. Importantly, the interaction term *GZ spread x GDP growth* is large and statistically significant for the startups with high growth potential. In other words, a worsening of GDP growth increases the negative effect of *GZ spread* much more for high-growth startups than for the complementary sample, consistent with Prediction 3. When GDP growth is zero, the marginal effect of an increase in the *GZ spread* is -9.3% for high-growth startups. With a fall in GDP growth by one percentage point, this marginal effect is reinforced by an additional -10.8%. Conversely, in the case of low-growth startups, the marginal effect is only -3.3% when GDP growth is zero and decreases by an additional -3.8% when GDP growth falls by one percentage point.

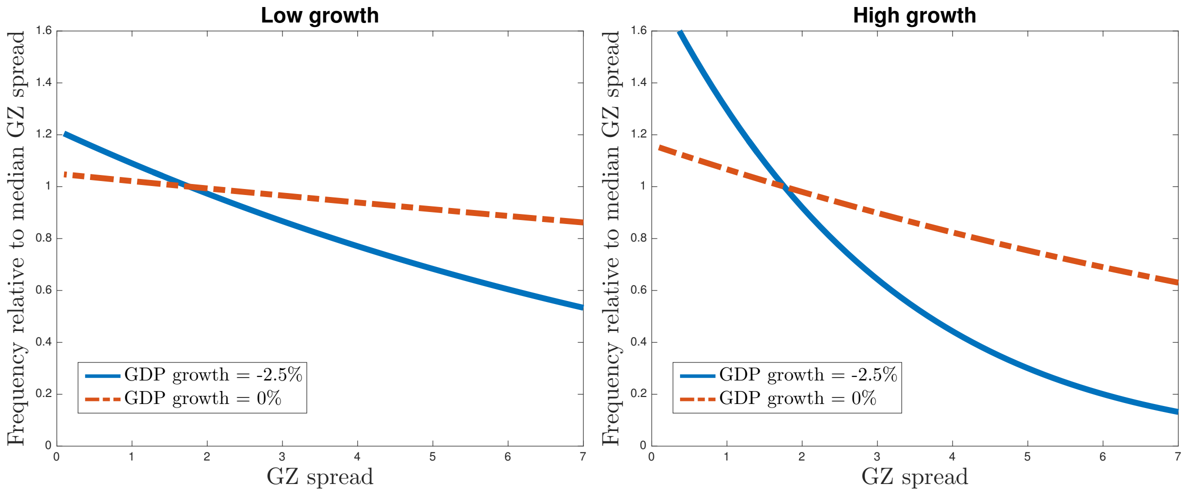
To relate these results to the model more clearly, we use the estimated coefficients to compute the marginal effects of *GZ spread*, conditional on a given value of GDP growth, for high-growth startups and the complementary group, as depicted in Figure 6. The solid line represents a contractionary period (GDP growth equals -2.5%), and the dashed line represents a period of zero growth. The lines are normalized to 1 for the median value of the GZ spread. For example, a value of the y-axis of 1.2 implies that the probability to start a business in the respective category is 20% higher than when the GZ spread is at its median value.[19]

Figure 6 is useful because it provides a graphical test of the predictions. Prediction 1 is satisfied if all the lines are decreasing in the GZ spread. Prediction 2 is satisfied if the lines are steeper for the likely high-growth-potential startups than for the complementary group. Finally, Prediction 3 is satisfied if the difference in slope between the solid line and dashed line is larger for the high-growth startups than for the complementary group.

Figure 6 is consistent with all the predictions. In terms of significance, a Wald test confirms that the negative slope of the solid line for the high-growth startups is signif-

[19]Without the normalization, the dashed lines would always lie above the solid line, as more startups are created when GDP growth is higher.

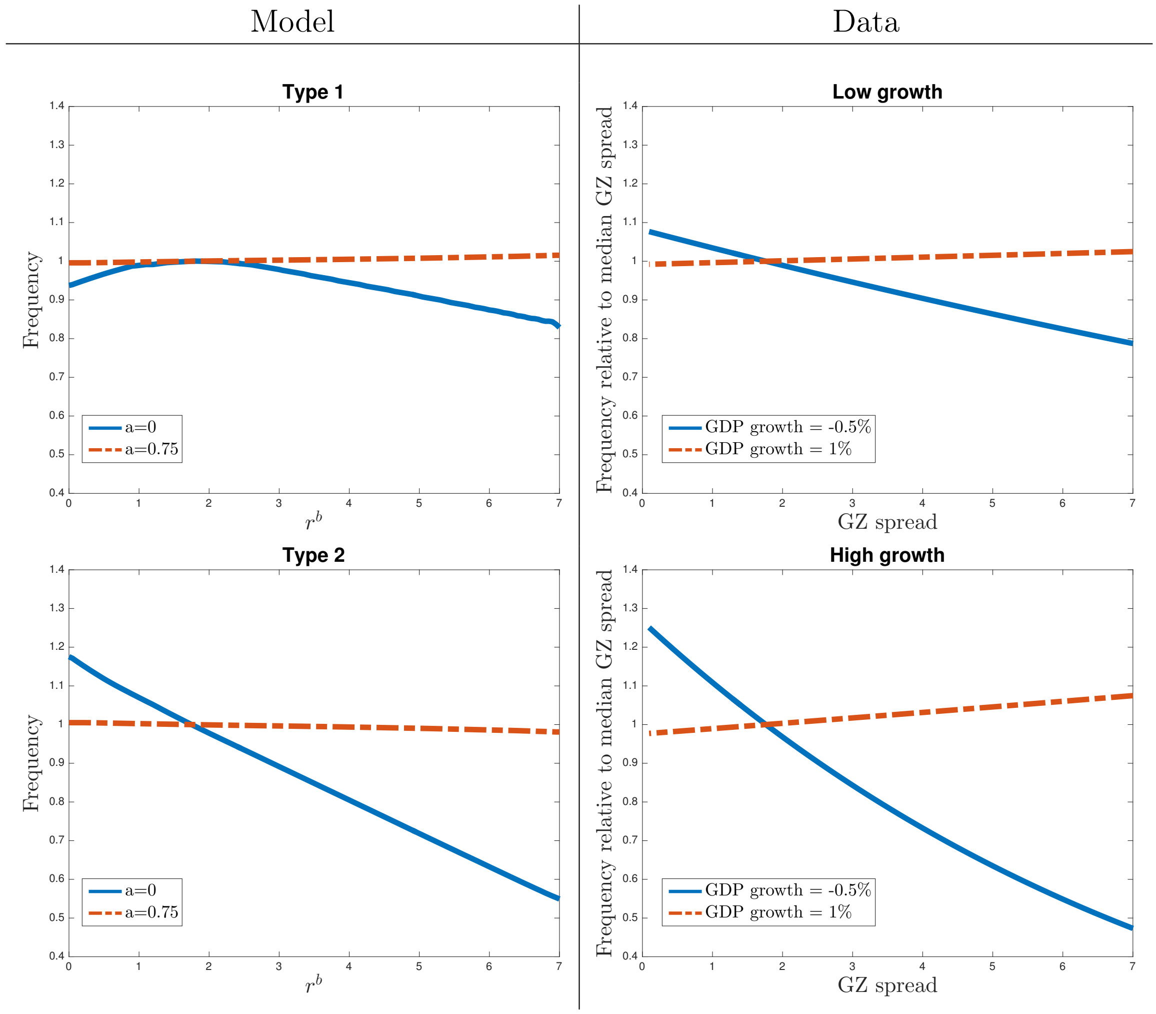
Figure 6: GZ spread and probability of starting a firm



icantly different from zero and significantly steeper than the dashed line, while for the low-growth firms, we cannot reject the hypothesis that the solid line has the same slope as the dashed one. Since these are separate regressions, we cannot test whether the slopes of these lines are different across these two graphs. Therefore, in Appendix Table 32, we estimate a two-step Heckman selection model, where the first step determines the probability of starting any type of business and the second step determines the specific type. This approach allows us to test and confirm that the interaction term *GZ spread x GDP growth* is significantly larger for high-growth startups than for the complementary group.

We further illustrate graphically the correspondence between the model and data in Figure 7, which compares the lines predicted by the model for Type 1 and Type 2 startups in the left part of the figure with those estimated in the data in the right part of the figure. For the model part, we choose two values of the entrepreneur's initial endowment to represent the case in which the complete initial sunk cost κ has to be financed externally ($a = 0$) and the case in which only 25% has to be financed externally ($a = 0.75$). The other parameters are those defined for the benchmark calibration described in Section 3.1. We then calculate the frequencies of Type 1 and Type 2 startups for each of the two cases and normalize them by their frequency at $r^b = 2\%$, which corresponds to the

Figure 7: Comparison of model and empirical predictions



median GZ spread in the data. These are represented in the "model" part of Figure 7.[20]

For the empirical data part, we predict the relative frequencies by using the probit estimation results fixing GDP growth at different values. In the top right graph, we choose the values for GDP growth that match the Type 1 startups in the top left graph. The GDP growth rates that closely replicate the model predictions with $a = 0$ and $a = 0.75$ are -0.5% and 1%, respectively.[21]

Then, in the bottom right graph, we predict the high-growth startups' frequencies by

[20] These lines are essentially normalized versions of the lines shown in the left Panel of Figure 2.

[21] Hence, according to the data, when GDP growth reaches 1%, financial frictions barely matter for startup decisions. Note that the functional form of the probit regressions does not allow the replication of the inverse u-shaped solid line predicted by the model for values of r^b around 0% to 3%.

using these same GDP growth values. These frequencies are fairly close to the model's Type 2 startup predictions shown in the bottom left plot. The effect of the GZ spread on the relative frequencies is only somewhat stronger in the data than in the model. In other words, the greater sensitivity to the cost of finance of Type 2 relative to Type 1 startups in the model matches well the greater sensitivity of the high-growth startups in the data.[22]

5.4 Using exogenous monetary policy shocks to predict credit spreads

The previous sections show that high-growth startups are more negatively affected by financial shocks than their complementary startup types. Moreover, this negative differential is amplified during downturns. Appendix C shows that these results are confirmed when analyzing only the Spanish GEM surveys, and the analysis from SABI data for Spain has shown that the percentage of high-growth startups predicts the faster employment growth of young firms. Thus, we find support for the prediction that businesses with high growth potential are more difficult to start after negative financial shocks.

In the model, the causality goes from financial frictions to startup decisions for given investment opportunities. However, it might be that in the data negative investment opportunity shocks not captured by the control variables affect both financial frictions and startup choices at the same time. In Section 6.4, we show that the results are robust to adding year fixed effects, which control for any shock common to all countries. Here, we provide additional evidence in support of the causal link from financial frictions to startup decisions. More specifically, we repeat the estimation of models (12) and (13) by using

[22]The positive slope of the dashed line in the bottom right graph comes from the fact that in the regression model, the marginal effect of the GZ spread increases linearly with GDP growth. Therefore, we cannot capture the non-linearity due to which financial frictions do not matter for high values of GDP growth. As a result, the slope eventually "overshoots" and becomes positive, thereby linearly extrapolating the strong relationship between the marginal effect of the GZ spread and GDP growth that exists for low values of GDP growth. Indeed, the coefficient of the interaction between the GZ spread and GDP growth is zero for all types of startups if we run the probit model by using only observations with GDP growth larger than 1%.

the GZ spread predicted by monetary policy shocks from Jarocinski and Karadi (2018). These authors follow well-established literature that uses high-frequency financial-market surprises around key monetary policy announcements to identify unexpected variations in monetary policy (e.g.. see Campbell, Evans, Fisher and Justiniano, 2012; Gertler and Karadi, 2015; Nakamura and Steinsson, 2018; Paul, 2017; Corsetti, Duarte and Mann, 2018). The innovative aspect of Jarocinski and Karadi's (2018) approach is that for both the US and the EU, they are able to separately identify exogenous monetary policy shocks from shocks about new information from the Central Bank regarding the state of the economy.

Therefore, these monetary policy shocks potentially affect the availability of credit and the bond spreads but are by construction orthogonal to contemporaneous shocks to investment opportunities. Our identifying assumption is that monetary policy shocks in year t affect entrepreneurial decisions in that year only through their effect on credit availability, as measured by the bond spreads. Both the monetary policy shocks and the bond spreads are available at the monthly level, and we regress the bond spreads in year t and month j on the monetary policy shocks in year t from month 1 to month j.[23] Since the nature of monetary policy changed substantially in the period after the financial crisis, we allow the estimated coefficients to be different pre- and post-2007. In our analysis thus far, we have considered the bond spreads of financial institutions as a measure of financial frictions faced by firms because Gilchrist and Mojon (2016) show that such spreads are good proxies for credit availability to households and firm, while variations in corporate bond spreads might be affected more by other factors. However, the fluctuations in corporate bond spreads predicted by monetary policy shocks are independent from these other factors, and we can therefore consider both the predicted corporate and financial bond spreads as measures of financial frictions. The results of these regressions are shown in Appendix C, Table 22 (we report only the first four lags due to space constraints).

[23]We exclude lagged monetary policy shocks from previous years because they might affect startup decisions indirectly through their delayed effect on economic activity. Bond spreads in each EU country is predicted using EU monetary policy shocks, while US bond spreads are predicted using US shocks.

Table 5: Predicted GZ spread and probability of starting a firm

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	3.615***	3.279***	3.285***	3.697***	3.335***	3.429***
	(0.1414)	(0.1614)	(0.2160)	(0.1578)	(0.1795)	(0.2440)
Predicted GZ spread	-0.102***	-0.058***	-0.188***	-0.094***	-0.051**	-0.175***
	(0.0191)	(0.0212)	(0.0322)	(0.0204)	(0.0229)	(0.0338)
Pr. GZ spread x GDP growth				0.586	0.383	1.128
				(0.4966)	(0.5418)	(0.8837)
Observations	399494	399494	399494	399494	399494	399494
R-squared	0.040	0.037	0.038	0.040	0.037	0.038

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. The standard errors are obtained by using a bootstrap procedure with 5000 replications. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 6: Predicted GZ spread and probability of starting a firm (without US)

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	3.417***	3.111***	3.103***	3.566***	3.259***	3.184***
	(0.1441)	(0.1649)	(0.2182)	(0.1454)	(0.1668)	(0.2194)
Predicted GZ spread	-0.221***	-0.160***	-0.302***	-0.187***	-0.130***	-0.271***
	(0.0234)	(0.0265)	(0.0381)	(0.0239)	(0.0270)	(0.0389)
Pr. GZ spread x GDP growth				6.310***	5.219***	7.141***
				(0.7065)	(0.7810)	(1.2117)
Observations	348905	348905	348905	348905	348905	348905
R-squared	0.035	0.030	0.039	0.036	0.030	0.040

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. The standard errors are obtained by using a bootstrap procedure with 5000 replications. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Monetary policy shocks have predictive power on the bond spreads, especially after 2007 and especially for corporate bonds. The correlation of the resulting predicted monthly corporate GZ spread with the actual spread is equal to 0.62 (0.44 when subtracting country means). We then compute the yearly averages of the predicted monthly spreads to replace the actual spread in the estimations. Figure 11 in the Appendix plots the series of predicted GZ spreads.

Table 5 presents the results for the full sample, and, to reduce the heterogeneity in the nature of the monetary policy shocks, Table 6 only considers the subsample of EU countries. The indicator of financial frictions is the predicted corporate bond spread. The variance-covariance matrix is estimated with a bootstrap procedure with 5000 replications. The coefficients of the predicted GZ spread and its interaction with GDP growth confirm the main result that high-growth startups are more penalized by exogenous increases in the cost of credit than are low-growth startups, especially in periods of low GDP growth. In fact, the coefficients are often larger and more significant than in the baseline in Tables 3 and 4. In Tables 23-26 in Appendix C, we consider alternative specifications: one that considers only Spain, one that uses predicted financial spreads, and one in which the monetary policy coefficients are restricted to be the same before and after 2007. These results also broadly confirm the main findings.

Given our identification strategy, the results in tables 5-6 can be interpreted as the effects of exogenous changes in the cost of finance, and we can combine them with the analysis on SABI in Section 5.1 to predict their implications for employment growth (abstracting from the general equilibrium effects on wages and interest rates). To do so, we perform a simple back-of-the-envelope calculation to quantify the effects of financial frictions and business cycle conditions on the size of new firms emerging via the *composition of entry* channel, i.e., the decline in the share of high-growth startups. First, we compute the changes in this share based on the changes in the probabilities of the two types of startups predicted by the regression models of columns 5 and 6 in Table 6. We then use column 4 of Table 2 and, in particular, the coefficient of the *Age 10 x share* interaction to calculate the implied effect on employment per firm ten years after firm creation.

We consider two scenarios: a boom period with GDP growth equal to 3% and a recession period with GDP falling by 3%.[24] In the first scenario, we obtain from Table 6 that a one percentage point increase in the GZ spread decreases the share of high-growth startups from 29.6%, which is the prediction at the average spread, to 25.1%. If the spread increases by three percentage points, the share decreases to 17.4%. This finding implies that on average, in response to a one percentage point and three percentage point

[24]These growth rates roughly correspond to those of Spain before and during the Great Recession.

increase in the GZ spread, new firms grow slower and, after 10 years, have 1.1% and 2.9% fewer employees, respectively. In the recession scenario, the effects of an increase in the spread are considerably larger. A one-percentage point increase implies a decrease in the high-growth startup share from 28.7% to 16.0%, whereas a three percentage point increase implies a decrease to 2.7%. As a result, in response to a one percentage point and a three percentage point increase in the GZ spread, the predicted firm employment after 10 years decreases by 3.5% and 6.6%, respectively.

Note that this *composition of entry* channel leads to a lower employment level after approximately six years, as this is the point at which the coefficients of the age-share interactions become positive. However, the effect on employment growth is felt already after 3-4 years, as seen in the first column of Table 2.

5.5 Industry-level measures of financial frictions

In the model, we assume that Type 1 and Type 2 startups have different patterns of productivity growth but need to finance the same initial investment κ and face the same excess cost of external finance r^b. An alternative approach to test the link between finance and startup type is to instead select projects that differ in terms of κ and r^b. We identify differences in κ in the data with the Rajan and Zingales (1998) external financial dependence (EFD) indicator, which measures the fraction of investment needs not covered by internally generated funds. The hypothesis is that the different technological features of the industries determine the different financing needs of firms. In high-EFD Industries, firms require on average more external financing to fund their investment, and thus, it is plausible that in such industries, startups have a larger value of κ than do other industries. Predictions 1-3 can clearly be extended to this case: a higher value of κ means that startups need higher initial financing and are more affected by changes in r^b. Therefore, startups in high-EFD industries are likely to be more sensitive to changes in the excess cost of finance than startups in low-EFD industries. To investigate this hypothesis, we repeat our estimations considering only starts in the manufacturing sector. We use data on industry-level financial dependence from Kroszner et al. (2007), and we identify the manufacturing startups with low- and high-external financial dependence (*low EFD* and *high EFD*). The details can be found in the Appendix B.6.

Furthermore, we identify differences in r^b with differences in asset tangibility. The corporate finance literature has shown that the tangibility of assets is an important factor for firms to obtain loans (see, e.g., Almeida and Campello, 2007). More tangible assets have more collateral value, which can be pledged to obtain loans with low excess cost r^b. Therefore, industries with a higher share of intangible assets should have less pledgeable collateral and higher values of r^b, especially in periods of financial stress and high external finance costs. We match the Compustat SIC classification with the 2-digit sectors in the GEM dataset, and we assign to each GEM sector the intangible capital share computed in Caggese and Perez (2017). We then calculate the median values and classify a sector as having a high (low) intangible share if its value is above (below) the median.

5.5.1 Intangible assets

In this section, we analyze the behavior of startups classified according to the amount of intangible assets. Analogously to Table 1, Table 7 compares the percentages of startups inside and outside the period of the financial crisis but considers only the sectors for which the measure of intangibility is available.[25] The percentage difference between the two periods is very similar to that reported in Table 1 for all startups. When comparing the drop in startups between low- and high-intangible sectors, we see that it is much larger for the latter than for the former, in line with our expectations.

Table 8 shows the regression results of Equation (13). In the first three columns, we find that the financial crisis dummy and the interaction with GDP growth are much larger and more significant for the high-intangible sectors, which is consistent with the

[25] We can classify only a subset of all startups (approximately 54%) because the information on the intangible share is not available for all sectors in the GEM data. We have verified that the main results shown in Table 4 also hold in this subsample.

Table 7: Percentage of individuals starting a firm (sectors with tangibility information)

	All	Low intan.	High intan.
Full	2.07	1.42	0.65
No Fin. crisis	2.42	1.63	0.79
Fin. crisis	1.58	1.13	0.45
% Difference	-34.71	-30.67	-43.04

Table 8: Financial crisis, GZ spread and probability of starting a firm

	(1) All	(2) Low intan.	(3) High intan.	(4) All	(5) Low intan.	(6) High intan.
GDP growth	0.637	0.535	0.617	5.520**	4.082***	7.337**
	(0.6258)	(0.5267)	(0.6626)	(2.1986)	(1.5757)	(3.2281)
Fin. crisis	-0.163***	-0.108***	-0.252***			
	(0.0545)	(0.0403)	(0.0858)			
Fin. crisis x GDP growth	4.446**	3.155**	6.663**			
	(1.8791)	(1.5244)	(2.7952)			
GZ spread				-0.020	-0.017	-0.026
				(0.0221)	(0.0155)	(0.0362)
GZ spread x GDP growth				2.660	2.067*	3.415
				(1.7119)	(1.1846)	(2.6825)
Observations	894126	894126	894126	370280	370280	370280
R-squared	0.062	0.057	0.063	0.039	0.028	0.053

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

predictions of the model. The last three columns show the results of Equation (13) by using the *GZ spread*. Compared to the firms in the complementary sample, high-intangible firms are more sensitive to financial conditions, especially during downturns. However, the interaction coefficient in column 6 is not statistically significant.

5.5.2 External financial dependence

Table 9: Percentage of individuals starting a manufacturing firm

	All	Low EFD	High EFD
Full	0.24	0.14	0.10
No Fin. crisis	0.29	0.17	0.13
Fin. crisis	0.16	0.10	0.06
% Difference	-44.83	-41.18	-53.85

In this section, for the smaller sample of manufacturing startups (approximately 5%

of the total), we analyze the behavior of startups classified according to external financial dependence. Table 9 shows the percentages of individuals starting manufacturing firms.

Table 10: Financial crisis, GZ spread and probability of starting a firm

	(1) All	(2) Low EFD	(3) High EFD	(4) All	(5) Low EFD	(6) High EFD
GDP growth	0.463	1.650**	-0.973	5.544**	3.856*	7.181***
	(1.3067)	(0.8264)	(1.7349)	(2.1866)	(2.1633)	(1.7670)
Fin. crisis	-0.163***	-0.089**	-0.240**			
	(0.0605)	(0.0385)	(0.0941)			
Fin. crisis x GDP growth	5.104**	2.672*	7.688**			
	(2.1893)	(1.5425)	(3.3112)			
GZ spread				0.003	-0.016	0.025
				(0.0223)	(0.0194)	(0.0275)
GZ spread x GDP growth				1.407	0.406	2.693***
				(1.5613)	(1.8754)	(0.7913)
Observations	894126	894126	894126	370280	370280	370280
R-squared	0.057	0.047	0.067	0.032	0.032	0.032

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

The decline in the probability of starting any type of manufacturing firm between crisis and non-crisis periods is -45%, whereas the drop is -54% in the high-EFD sectors and -41% in the low-EFD sectors.

Table 10 shows the regression results. We find that the financial crisis dummy and the interaction with GDP growth are much larger and more significant for the high-EFD sectors than for the low-EFD sectors. Using the *GZ spread*, we also find large differences in the interaction term between the two categories, in line with the predictions of the model.

6 Robustness checks

In this section, we complement the analysis with a number of robustness checks. We use an alternative measure of financial frictions, and we make use of additional information from the GEM survey to include further control variables and consider an alternative

Table 11: Romer and Romer financial distress indicator and probability of starting a firm

	(1) All	(2) Low growth	(3) High growth
GDP growth	2.812**	2.615**	2.509
	(1.4047)	(1.1194)	(1.6171)
RR indicator	-0.012	-0.011	-0.012*
	(0.0088)	(0.0089)	(0.0070)
RR indicator x GDP growth	0.667**	0.528*	0.789*
	(0.3384)	(0.2876)	(0.4229)
Observations	731881	731881	731881
R-squared	0.043	0.039	0.041

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

method to identify high-growth startups.

6.1 Alternative indicator of financial frictions

Table 11 replicates the last three columns of Table 4 and Figure 6. As an alternative measure of financial frictions, the RR financial distress indicator of Romer and Romer (2017) is used. As argued above, the RR indicator is explicitly designed to capture both high bond spreads and other factors of financial distress that might be important for the new firms' access to finance, and it is available for our full sample. The results show that the predictions of the model are confirmed with this alternative measure. The coefficient of the RR indicator is significantly negative only for high-growth startups, and the interaction is larger in magnitude for this group than for the complementary sample.

6.2 Additional control variables

In the previous sections, we found a strong negative effect of financial conditions on *high-growth* startups, which are defined as firms with entrepreneurs who expect that their firm will become larger in size than the other firms in their country/industry in 5 years' time. The designers of the survey included this question on the expected future size of the firm precisely to capture startups with high growth potential. However, as argued above, the answer might be affected by expectations on the future state of the economy. Our analysis on firm-level data (see Figure 5) confirms that our *high-growth* indicator provides relevant information on the intrinsic growth potential of these startups and is not simply a measure of expectations on the economy. Moreover, since our results are conditional on GDP growth, to the extent that these expectations are correlated with current growth, our estimated effect of financial frictions on high-growth startups is robust to this problem. Nonetheless, we can further check the robustness to this potential problem because the GEM surveys contain a question on expectations of future business opportunities. The exact question is "*In the next six months, will there be good opportunities for starting a business?*", which can be answered with *Yes*, *No* or *Don't know*. We exclude respondents with the answer *Don't know* and include in the analysis the variable *Opportunity expectations*, which is equal to 1 for *Yes* and 0 otherwise. Although the time horizon of this expectations variable is relatively short, we should expect that if the results of the *high-growth* startups are entirely driven by future expectations of the economy, they should be at least partially absorbed by the inclusion of this variable.

Table 12 repeats the analysis in Table 4 after adding the *Opportunity expectations* variable, Its coefficient is positive and strongly significant in all specifications. The other coefficients are quantitatively very similar to the baseline results in Table 4 and only slightly less significant; we still find that the stronger negative effect of a financial crisis and of high bond spreads is concentrated among the *high-growth* startups. The fact that the baseline results still hold when controlling for expectations confirm that our categorization of startups reflects the nature of the new businesses rather than just general expectations about the economy.

In Table 13, we add as a control variable the share of firm exits for each sector/country/year observation. This variable captures the possibility that new startups are driven by the presence of serial entrepreneurs who seek to start a new business. We find this variable to be generally not statistically significant and that it does not affect the previous results.

Table 12: Financial crisis, GZ spread and probability of starting a firm

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	0.175	0.533	-0.469	4.123	3.050	5.208**
	(0.6646)	(0.5870)	(0.5285)	(2.8590)	(2.6752)	(2.2067)
Fin. crisis	-0.207***	-0.171***	-0.218***			
	(0.0563)	(0.0460)	(0.0652)			
Fin. crisis x GDP growth	4.002*	3.109*	4.651*			
	(2.1734)	(1.7523)	(2.7785)			
GZ spread				-0.033	-0.027	-0.042*
				(0.0285)	(0.0281)	(0.0221)
GZ spread x GDP growth				1.793	0.815	3.411*
				(2.3642)	(2.1263)	(1.9965)
Opportunity expectations	0.448***	0.403***	0.401***	0.452***	0.412***	0.400***
	(0.0315)	(0.0306)	(0.0304)	(0.0581)	(0.0553)	(0.0502)
Observations	724666	724666	724666	328196	328196	328196
R-squared	0.084	0.065	0.094	0.065	0.059	0.058

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

In Table 14, we add income categories (for the definition, see footnote 16). We find that being in a higher income category increases the probability of starting a new firm, but again in this case, the inclusion of these additional control variables does not significantly change the results obtained previously, and if anything, it makes them slightly stronger.

Table 13: Financial crisis, GZ spread and probability of starting a firm

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	0.846	1.106**	0.175	5.489***	4.494***	6.211***
	(0.5810)	(0.5359)	(0.4685)	(1.6142)	(1.3957)	(1.4898)
Fin. crisis	-0.165***	-0.132***	-0.187***			
	(0.0478)	(0.0377)	(0.0603)			
Fin. crisis x GDP growth	4.562**	3.785***	4.972*			
	(1.8739)	(1.4442)	(2.6347)			
GZ spread				-0.020	-0.013	-0.033*
				(0.0200)	(0.0192)	(0.0175)
GZ spread x GDP growth				2.468*	1.560	3.844***
				(1.4484)	(1.1886)	(1.4622)
Share of exits	2.638	2.243	2.396	0.426	0.809	0.264
	(2.4544)	(2.2407)	(2.0532)	(4.4940)	(4.8303)	(2.9459)
Observations	894126	894126	894126	370280	370280	370280
R-squared	0.062	0.046	0.078	0.039	0.035	0.039

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 14: Financial crisis, GZ spread and probability to start a firm

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	0.582	0.901	-0.126	5.350***	4.334**	6.079***
	(0.6673)	(0.6105)	(0.5206)	(2.0213)	(1.7920)	(1.6990)
Fin. crisis	-0.193***	-0.148***	-0.232***			
	(0.0510)	(0.0403)	(0.0598)			
Fin. crisis x GDP growth	4.756***	3.915***	5.244**			
	(1.8002)	(1.3862)	(2.5267)			
GZ spread				-0.027	-0.018	-0.043**
				(0.0195)	(0.0189)	(0.0170)
GZ spread x GDP growth				2.422	1.502	3.840**
				(1.6770)	(1.3905)	(1.6037)
Middle income	0.093***	0.075***	0.106***	0.158***	0.125***	0.183***
	(0.0333)	(0.0267)	(0.0400)	(0.0234)	(0.0247)	(0.0217)
High income	0.133***	0.073***	0.202***	0.145***	0.104***	0.193***
	(0.0177)	(0.0193)	(0.0209)	(0.0249)	(0.0260)	(0.0301)
Observations	894126	894126	894126	370280	370280	370280
R-squared	0.064	0.047	0.081	0.042	0.037	0.044

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 15: Financial crisis, GZ spread, RR indicator and probability of starting a firm

	(1) Not inn	(2) Inn	(3) Not inn	(4) Inn	(5) Not inn	(6) Inn
GDP growth	1.009*	-0.094	4.892***	5.274***	2.718**	2.300*
	(0.5543)	(0.6750)	(1.8046)	(1.9308)	(1.2650)	(1.3715)
Fin. crisis	-0.140**	-0.156***				
	(0.0624)	(0.0349)				
Fin. crisis x GDP growth	4.095**	4.552***				
	(1.8890)	(1.2877)				
GZ spread			-0.030	0.010		
			(0.0199)	(0.0241)		
GZ spread x GDP growth			2.042	2.829**		
			(1.6368)	(1.1941)		
RR indicator					-0.010	-0.012
					(0.0083)	(0.0137)
RR indicator x GDP growth					0.542	0.793***
					(0.3498)	(0.2464)
Observations	894126	894126	370280	370280	731881	731881
R-squared	0.046	0.087	0.038	0.028	0.040	0.036

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

6.3 Classifying startups by innovativeness

In this subsection, we alternatively classify startups as non-innovative and innovative ones. The latter might grow faster in the long run because new products or services have the potential to capture larger market shares. However, these firms supplying novel products are also likely to be associated with higher risk and are thus additionally affected by risk premia fluctuations, which are not captured by our model. Nevertheless, we expect results similar to those in our baseline classification because in most countries, risk premia increased during the period of the financial crisis and the financing of innovative and potentially risky projects became more difficult.

Table 15 shows the regression results for the two complementary subgroups of non-innovative and innovative startups. The results obtained by using the previous classification of low- and high-growth startups are broadly confirmed, regardless of the indicator of financial frictions used. The effect of the interaction term is always significant and larger for innovative startups than for non-innovative startups.

6.4 Additional robustness checks

In tables 20, 21 and 23 in Appendix C, we show that the negative effect of bond spreads on high-growth startups is confirmed when considering only the Spanish GEM surveys. In Table 27, we replicate all the regressions after excluding the countries that did not experience a systemic banking crisis (according to Laeven and Valencia 2013). Thus, the crisis dummy is identified by comparing the crisis period with the pre-crisis period only for countries that experienced the crisis. In Table 28, we exclude the construction sector. We do this because in most countries, the collapse of this sector caused the banking crisis, rather than vice versa. Both of these robustness checks confirm the results shown above. In Table 29, we exclude startups that have already paid some wages and thus might have been established before, and once again, we confirm the previous results.

In Table 30, we estimate the baseline model when additionally including year fixed effects, which control for any time-varying factor common to all countries. As expected, representing a common shock to almost all countries in our dataset, the financial crisis dummy becomes insignificant. Nonetheless, the main results regarding the interaction between financial frictions and GDP growth are confirmed.

In Table 31, we replace the financial crisis indicator with an indicator for the Great Recession. This is a dummy equal to one if a country suffered two subsequent quarters with negative economic growth during the period 2008-2010. We find that the interaction term is strongly significant and larger for high-growth startups. This finding implies that these startups declined more during the great recession in countries that experienced a larger contraction in GDP during that period.

In Table 32, we estimate a two-step Heckman selection model. The first-stage selection equation determines the probability of starting a business and in addition to GDP growth, includes the indicator for financial frictions and their interaction; it also includes the additional control variables of sex, education, age and country dummies. The second-stage equation estimates the effects of GDP growth and financial frictions on the type of business created. This specification allows us to disentangle the effect of demographics on the likelihood of opening a business from the effect of financial conditions on starting a business with high growth potential. The results of the second stage shown in the table confirm that startups with high growth potential are less frequent during a financial crisis and are significantly more sensitive to financing conditions than are the other startups.

In Tables 33 and 34, we include the country-specific riskless interest rate and its interaction with GDP growth as regressors.[26] The tables show that, probably because it is a leading empirical indicator of the business cycle, the riskless rate generally has a positive relation with startups. Column 6 shows that the interaction of the riskless rate with GDP growth is negative and significant for high-growth startups. The magnitude of the estimated coefficients implies that riskless rates are always positively correlated with high-growth startups but more so during downturns than during upturns. Importantly, our main results are confirmed, and compared to the findings in our baseline estimation in Table 4, the coefficients of the interaction between the GZ spread and GDP growth become somewhat larger in absolute value and gain significance.

7 Conclusion

In this paper, we investigate whether financial frictions differentially affect startups with high growth potential. Our stylized model predicts that at the margin, a high-growth-potential startup is less profitable in the short term and more profitable in the long term. We use survey-level information from the GEM dataset to identify high-growth startups in the data. For the case of Spain, which has very extensive coverage in the GEM dataset, we use firm-level data from SABI to confirm that high-growth startups are more likely to grow faster and employ more people in the long term than are other

[26] We obtain the series of 3-month nominal interest rates (computed by the OECD by using either treasury bills or money market rates), and we subtract the inflation rates to obtain the real rates.

startups. The model predicts that high-growth startups are more negatively affected by increases in the cost of external finance, especially when GDP growth is low, and our empirical results confirm these predictions. Importantly, we find additional evidence that is consistent with a financial accelerator story. The access to finance matters, especially for startups in sectors with a high share of intangible assets and in sectors with a high dependence on external financing. Taken together, our results support the view that this *composition of entry* channel is important for explaining slow recoveries after financial crises. The policy implication of our analysis is that credit subsidies specifically targeted at high-growth startups should be effective at countering the negative long-term effects of financial crises.

References

Acemoglu, Daron, Ufuk Akcigit, and Murat Alp Celik, "Young, Restless and Creative: Openness to Disruption and Creative Innovations," NBER Working Papers 19894, National Bureau of Economic Research, Inc February 2014.

Adelino, Manuel, Antoinette Schoar, and Felipe Severino, "House prices, collateral, and self-employment," *Journal of Financial Economics*, 2015, *117* (2), 288–306.

Akcigit, Ufuk and William R. Kerr, "Growth through Heterogeneous Innovations," CEPR Discussion Papers 11660, C.E.P.R. Discussion Papers November 2016.

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

Azoulay, Pierre, Benjamin Jones, J. Daniel Kim, and Javier Miranda, "Age and High-Growth Entrepreneurship," NBER Working Papers 24489, National Bureau of Economic Research, Inc April 2018.

Blanchflower, David G. and Andrew J. Oswald, "What Makes an Entrepreneur?," *Journal of Labor Economics*, 1998, *16* (1), 26–60.

Braun, Matías and Borja Larrain, "Finance and the Business Cycle: International, Inter-Industry Evidence," *Journal of Finance*, June 2005, *60* (3), 1097–1128.

Brown, J. David, John S. Earle, Mee Jung Kim, and Kyung Min Lee, "High-Growth Entrepreneurship," IZA Discussion Papers 11662, Institute of Labor Economics (IZA) July 2018.

Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin, "Finance and Development: A Tale of Two Sectors," *American Economic Review*, August 2011, *101* (5), 1964–2002.

Caggese, Andrea, "Financing Constraints, Radical versus Incremental Innovation, and Aggregate Productivity," *American Economic Journal: Macroeconomics*, forthcoming.

_ **and Ander Perez**, "Capital Misallocation and Secular Stagnation," Finance and Economics Discussion Series 2017-009, Board of Governors of the Federal Reserve System (U.S.) January 2017.

_ **and Vicente Cunat**, "Financing Constraints, Firm Dynamics, Export Decisions, and Aggregate Productivity," *Review of Economic Dynamics*, January 2013, *16* (1), 177–193.

Campbell, Jeffrey R., Charles L. Evans, Jonas D.M. Fisher, and Alejandro Justiniano, "Macroeconomic Effects of Federal Reserve Forward Guidance," *Brookings Papers on Economic Activity*, 2012, *43* (1 (Spring), 1–80.

Carroll, Christopher D., "A Theory of the Consumption Function, with and without Liquidity Constraints," *Journal of Economic Perspectives*, Summer 2001, *15* (3), 23–45.

Cerra, Valerie and Sweta Chaman Saxena, "Growth Dynamics: The Myth of Economic Recovery," *American Economic Review*, March 2008, *98* (1), 439–457.

Clementi, Gian Luca and Berardino Palazzo, "Entry, Exit, Firm Dynamics, and Aggregate Fluctuations," *American Economic Journal: Macroeconomics*, July 2016, *8* (3), 1–41.

Cole, Harold L., Jeremy Greenwood, and Juan M. Sanchez, "Why Doesn't Technology Flow From Rich to Poor Countries?," *Econometrica*, July 2016, *84*, 1477–1521.

Corradin, Stefano and Alexander Popov, "House Prices, Home Equity Borrowing, and Entrepreneurship," *The Review of Financial Studies*, 03 2015, *28* (8), 2399–2428.

Corsetti, Giancarlo, Joao B. Duarte, and Samuel Mann, "One money, many markets: a factor model approach to monetary policy in the Euro Area with high-frequency identification," LSE Research Online Documents on Economics 87182, London School of Economics and Political Science, LSE Library February 2018.

Dell'Ariccia, Giovanni, Enrica Detragiache, and Raghuram Rajan, "The real effect of banking crises," *Journal of Financial Intermediation*, January 2008, *17* (1), 89–112.

Falato, Antonio, Dalida Kadyrzhanova, and Jae W. Sim, "Rising intangible capital, shrinking debt capacity, and the US corporate savings glut," Finance and Economics Discussion Series 2013-67, Board of Governors of the Federal Reserve System (U.S.) 2013.

Gertler, Mark and Peter Karadi, "Monetary Policy Surprises, Credit Costs, and Economic Activity," *American Economic Journal: Macroeconomics*, January 2015, *7* (1), 44–76.

Gilchrist, Simon and Egon Zakrajsek, "Credit Spreads and Business Cycle Fluctuations," *American Economic Review*, June 2012, *102* (4), 1692–1720.

— **and Mojon Benoit**, "Credit Risk in the Euro Area," *The Economic Journal*, 2016, *128* (608), 118–158.

Haltiwanger, John, "Job Creation, Job Destruction, and Productivity Growth: The Role of Young Businesses," *Annual Review of Economics*, August 2015, *7* (1), 341–358.

Haltiwanger, John C., Ron S. Jarmin, and Javier Miranda, "Who Creates Jobs? Small vs. Large vs. Young," NBER Working Papers 16300, National Bureau of Economic Research, Inc August 2010.

Haltiwanger, John, Ron S. Jarmin, Robert Kulick, and Javier Miranda, "High Growth Young Firms: Contribution to Job, Output, and Productivity Growth," in "Measuring Entrepreneurial Businesses: Current Knowledge and Challenges" NBER Chapters, National Bureau of Economic Research, Inc, 2016, pp. 11–62.

Holtz-Eakin, Douglas, David Joulfaian, and Harvey S Rosen, "Sticking It Out: Entrepreneurial Survival and Liquidity Constraints," *Journal of Political Economy*, February 1994, *102* (1), 53–75.

Hombert, Johan and Adrien Matray, "The Real Effects of Lending Relationships on Innovative Firms and Inventor Mobility," *The Review of Financial Studies*, 10 2016, *30* (7), 2413–2445.

Hsieh, Chang-Tai and Peter J. Klenow, "The Life Cycle of Plants in India and Mexico," *The Quarterly Journal of Economics*, 2014, *129* (3), 1035–1084.

Jarocinski, Marek and Peter Karadi, "Deconstructing monetary policy surprises: the role of information shocks," Working Paper Series 2133, European Central Bank February 2018.

Koellinger, Philipp D. and A. Roy Thurik, "Entrepreneurship and the Business Cycle," *The Review of Economics and Statistics*, November 2012, *94* (4), 1143–1156.

Krishnan, Karthik, Debarshi K. Nandy, and Manju Puri, "Does Financing Spur Small Business Productivity? Evidence from a Natural Experiment," *The Review of Financial Studies*, 11 2014, *28* (6), 1768–1809.

Kroszner, Randall S., Luc Laeven, and Daniela Klingebiel, "Banking crises, financial dependence, and growth," *Journal of Financial Economics*, April 2007, *84* (1), 187–228.

Laeven, Luc and Fabian Valencia, "The Real Effects of Financial Sector Interventions during Crises," *Journal of Money, Credit and Banking*, February 2013, *45* (1), 147–177.

Midrigan, Virgiliu and Daniel Yi Xu, "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, February 2014, *104* (2), 422–58.

Nakamura, Emi and Jón Steinsson, "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect," *The Quarterly Journal of Economics*, 2018, *133* (3), 1283–1330.

Paul, Pascal, "The Time-Varying Effect of Monetary Policy on Asset Prices," Working Paper Series 2017-9, Federal Reserve Bank of San Francisco May 2017.

Poschke, Markus, "Who becomes an entrepreneur? Labor market prospects and occupational choice," *Journal of Economic Dynamics and Control*, 2013, *37* (3), 693–710.

_ , "The Firm Size Distribution across Countries and Skill-Biased Change in Entrepreneurial Technology," *American Economic Journal: Macroeconomics*, July 2018, *10* (3), 1–41.

Pugsley, Benjamin, Petr Sedlacek, and Vincent Sterk, "The Nature of Firm Growth," CEPR Discussion Papers 12670, C.E.P.R. Discussion Papers January 2018.

Pugsley, Benjamin Wild and Aysegül Sahin, "Grown-up Business Cycles," *The Review of Financial Studies*, 06 2018, *32* (3), 1102–1147.

Rajan, Raghuram G and Luigi Zingales, "Financial Dependence and Growth," *American Economic Review*, June 1998, *88* (3), 559–586.

Reynolds, Paul Davidson and Diana M. Hechavarria, "Global Entrepreneurship Monitor [GEM]: Adult Population Survey Data Set, 1998-2013," 2016.

Robb, Alicia M. and David T. Robinson, "The Capital Structure Decisions of New Firms," *The Review of Financial Studies*, 07 2012, *27* (1), 153–179.

Romer, Christina D. and David H. Romer, "New Evidence on the Aftermath of Financial Crises in Advanced Countries," *American Economic Review*, October 2017, *107* (10), 3072–3118.

Schmalz, Martin C., David A. Sraer, and David Thesmar, "Housing Collateral and Entrepreneurship," *Journal of Finance*, February 2017, *72* (1), 99–132.

Schoar, Antoinette, "The Divide between Subsistence and Transformational Entrepreneurship," *Innovation Policy and the Economy*, 2010, *10* (1), 57–81.

Sedlacek, Petr and Vincent Sterk, "The Growth Potential of Startups over the Business Cycle," *American Economic Review*, October 2017, *107* (10), 3182–3210.

Siemer, Michael, "Employment Effects of Financial Constraints During the Great Recession," *The Review of Economics and Statistics*, forthcoming.

Appendix

A Derivations of value functions

Using the interest rate on lending (equal to zero) as the discount factor, the value of a newly created Type 1 firm (gross of the start-up costs C^1 and κ) is equal to:

$$V^1(\theta_t) = (1-d)[\pi(\theta_t) + V^1\left((1+g^{med})\theta_t\right)] \tag{14}$$

where $\theta_{t+1} = (1+g^{med})\theta_t$. Using Equation 4 and substituting recursively yields:

$$\begin{aligned} V^1(\theta_0) &= (1-d)\Psi\left[\theta_0 + \frac{1}{1+\rho}\theta_0(1+g^{med}) + \frac{1}{(1+\rho)^2}\theta_0(1+g^{med})^2 + ...\right] \\ &= \Psi\frac{\theta_0}{\frac{d}{1-d} - g^{med}} = (1-d)\Psi\frac{\theta_0}{d-(1-d)\,g^{med}} \end{aligned}$$

The value of a Type 2 firm that switched permanently to high growth is:

$$V^{high}(\theta_t) = (1-d)\Psi\frac{\theta_t}{d-(1-d)\,g^{high}} \tag{15}$$

To compute its initial value, assume that with probability $1-\gamma$, the firm continues to grow at rate g^{low} such that $\theta_{t+1} = (1+g^{low})\theta_t$. However, with probability γ, it switches permanently to high growth, and its value becomes that as determined in Equation 15. Therefore, the initial value is:

$$V^2(\theta_0) = (1-d)\Psi\left[(1-\gamma)\theta_0 + \gamma\frac{\theta_0}{d-(1-d)\,g^{high}} + ...\right] \tag{16}$$

Rearranging yields:

$$V^2(\theta_0) = (1-d)\Psi\Phi\left\{\begin{array}{c} \theta_0 + (1-\gamma)(1-d)(1+g^{low})\theta_0 \\ +\left[(1-\gamma)(1-d)(1+g^{low})\right]^2\theta_0 + ... \end{array}\right\} \tag{17}$$

$$\Phi \equiv (1-\gamma) + \frac{\gamma}{d-(1-d)\,g^{high}} \tag{18}$$

Solving recursively yields:

$$\begin{aligned} V^2(\theta_0) &= (1-d)\Psi\Phi(1-\gamma)(1-d)\frac{\theta_0}{\frac{1-(1-\gamma)(1-d)}{(1-\gamma)(1-d)} - g^{low}} \\ &= (1-d)\Psi\Phi\frac{\theta_0}{1-(1-\gamma)(1-d)\left(1+g^{low}\right)} \end{aligned}$$

A.1 Calculation of C^1 and C^2

Substituting Equation 7 recursively and given the n periods necessary to repay the debt, for a Type 1 firm, its initial debt can be written as:

$$b = \Psi\theta_0 \left[\frac{1-\left(\left(1+g^{med}\right)\frac{1-d}{1+r^b}\right)^n}{\frac{r^b+d}{1-d} - g}\right] \tag{19}$$

Solving for n yields:

$$n^*(b, g^{med}, \Psi\theta_0) = \frac{\log\left\{1 - \frac{b}{\Psi\theta_0}\left(\frac{r^b+d}{1-d} - g\right)\right\}}{\log\left(\left(1+g^{med}\right)\frac{1-d}{1+r^b}\right)} \tag{20}$$

$n^*(b, g^{med}, \Psi\theta_0)$ is the number of periods necessary to repay debt b with growth g^{med} and initial profits $\Psi\theta_0$. Once we find n^*, we compute Equation 19 discounting the flows using $r = 0$ instead of $r = r^b$:

$$b^* = \Psi\theta_0 \left[\frac{1-\left(\left(1+g^{med}\right)(1-d)\right)^{n^*}}{\frac{d}{1-d} - g}\right] \tag{21}$$

b^* represents the net present value of the stream of revenues generated during the n^* periods. The difference between b^* and b is by construction equal to C^1. Note that in general, the procedure above can be used to compute $C(b, g, \theta_0, r^b)$, the excess cost of finance conditional on debt b, productivity growth g, initial productivity θ_0, and the interest rate premium r^b. It is then straightforward to show that $C(b, g, \theta_0, 0) = 0$ and that $C(b, g, \theta_0, r^b)$ increases in r^b.

Consider now a Type 2 firm. In the first period, the firm pays an excess return $r^b b_0$. The residual debt is $b_1 = \left(1 + r^b\right) b_0 - \Psi\theta_0$. In the second period, with probability γ, the firm switches to high growth so that $\pi_1 = \Psi\theta_0 \left(1 + g^{high}\right)$ and the residual cost is $C(b_1, g^{high}, \pi_1)$. With probability $(1 - \gamma)$, the firm remains a low-growth firm and pays an excess return $r^b b_1$, so that $b_2 = \left(1 + r^b\right) b_1 - \pi_1^{low}$. In this case, $\pi_1^{low} = \Psi\theta_0 \left(1 + g^{low}\right)$. Substituting recursively, this yields Equation 9.

B Data and variable definitions

B.1 Business types identified from GEM questions

High-growth startups

To identify a startup with high growth potential, we refer to the following two questions:

1. "Currently, how many people, not counting the owners but including exclusive sub-contractors, are working for this business?"

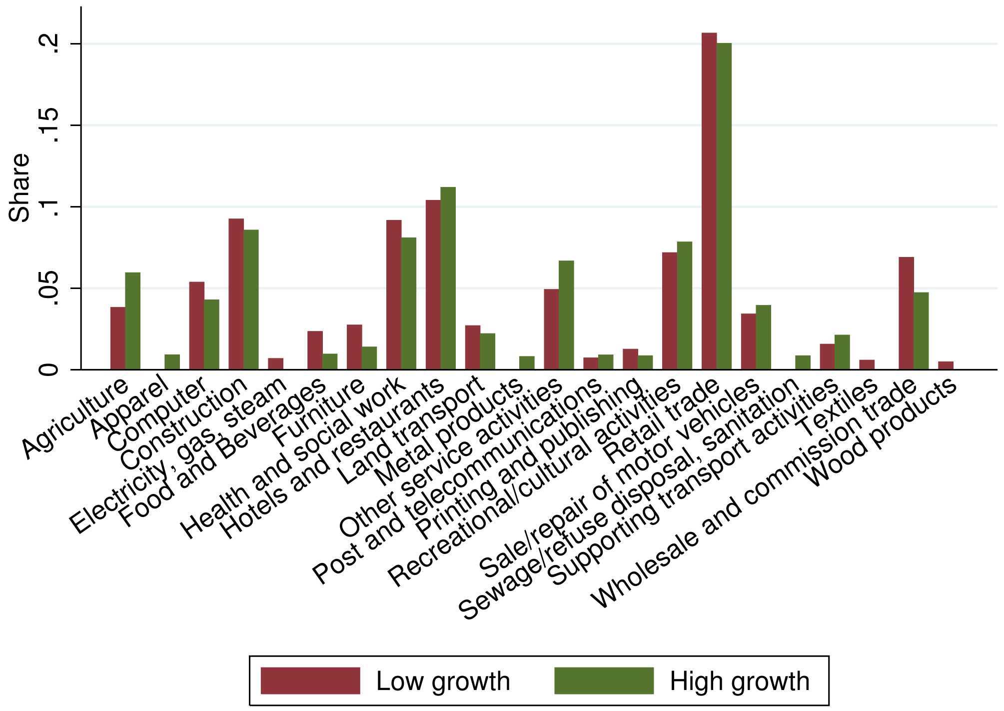
2. "Not counting the owners but including all exclusive sub-contractors, how many people will be working for this business when it is five years old?"

We compute the average size of the established firms by sector (at the 2-digit level) and country by using the answer to the first question given by respondents that are currently owners of firms that are 5 or more years old.[27] We then classify a startup as having high growth potential if the answer to the second question, i.e., the expected size in five years, exceeds the average size of the established firms at the sector-country level. Ideally, we would use only firms that are exactly 5 years old as the comparison benchmark. However, this results in very few observations in many country-sectors; therefore, we choose to consider all firms at least 5 years old. We have confirmed that the main results are not sensitive to using different ranges of the firm age, e.g., 5 to 10 years,

[27] As there is no information on the date of firm creation in the GEM data, we use the first year a firm paid wages or profits to the owners as a proxy.

to compute the average size of established firms. Figure 8 shows that the distribution of low-growth and high-growth startups for each 2-digit sector.

Figure 8: Distribution of low-growth and high-growth startups in 2-digit sectors



Notes: The figure shows the sector shares of startups in the 21 most common sectors, which account for approximately 94% of all startups, separately for the low-growth and high-growth categories.

B.2 Business cycle data

We take yearly GDP per capita data from the Penn World Tables. We compute yearly GDP growth as the percentage change in expenditure-side real GDP in chained PPP values.

B.3 Financial crisis data

We identify years in which a particular country is in a financial crisis by using data on systemic banking crises from Laeven and Valencia (2013). The following table shows the countries in our sample, the corresponding crisis period and the number of observations.

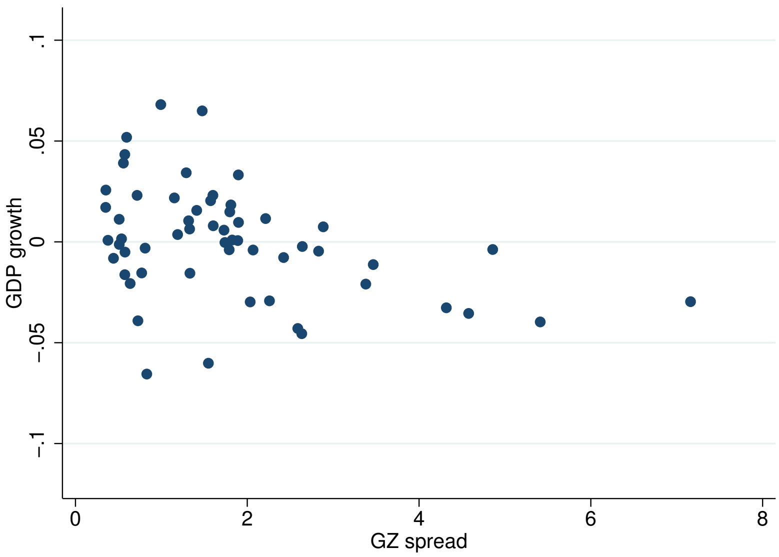
Table 16: Countries and financial crisis years

Country	Start year	End year	Obs.
Belgium	2008	2013	29995
Chile	-	-	36306
Croatia	-	-	22377
Denmark	2008	2013	28183
Finland	-	-	22231
France	2008	2013	23089
Germany	2008	2013	67619
Greece	2008	2013	20430
Hungary	2008	2013	22029
Iceland	2008	2013	16477
Ireland	2008	2013	20601
Italy	2008	2013	24572
Japan	-	-	22042
Netherlands	2008	2013	39500
Norway	-	-	22016
Slovenia	2008	2013	28865
Spain	2008	2013	233625
Sweden	2008	2013	45298
Switzerland	2008	2013	21079
United Kingdom	2007	2013	187967
United States	2007	2013	50589

Notes: The periods for systemic banking crises are taken from Laeven and Valencia (2013)

B.4 GZ bond spread

Figure 9: Correlation between GDP growth (deviation from country average) and bond spread



As a proxy for the financing costs of firms r^b at the country-year level, we rely on the excess bond premium for financial firms from Gilchrist and Zakrajsek (2012), who measure the bond premium with respect to the yields of 10-year US government bonds. We make our index comparable across countries by measuring the premiums of all countries with respect to the German bund. For the US, we take the domestic spread directly from Gilchrist and Zakrajsek (2012)[28], and we add the spread between US and German government bonds.[29] For France, Spain, Italy and Germany, we take the data from Gilchrist and Mojon (2016), who calculate the spread at the individual bond level and aggregate it.[30] We finally compute the yearly means of the monthly data.

B.5 Romer & Romer indicator

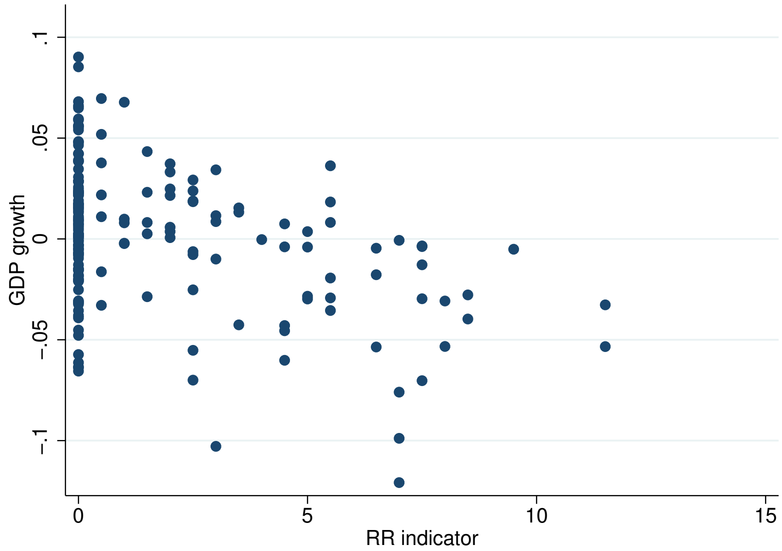
Romer and Romer (2017), based on qualitative information from the OECD Economic Outlook reports, which have been published by the OECD for individual countries since 1967, develop a measure of financial distress for 24 advanced countries. The indicator ranges from 0 to 14 and covers all countries in our sample until 2012, except Hungary, Chile, Croatia and Slovenia. The aim of this measure is to capture the "cost of credit intermediation", i.e., the costs of obtaining funds for financial institutions (relative to the riskless rate) and the costs of screening, monitoring and administering loans to borrowers. This makes it a suitable indicator for the spread between the lending rate and the riskless rate, represented by r^b in our model.

[28]Data are available at `http://people.bu.edu/sgilchri/Data/data.htm`

[29]Retrieved from `https://fred.stlouisfed.org/series/IRLTLT01USM156N`

[30]Data are available at `https://publications.banque-france.fr/en/economic-and-financial-publications-working-papers/credit-risk-euro-area`

Figure 10: Correlation between GDP growth (deviation from country average) and the RR indicator



B.6 Financial dependence and intangibility data

The GEM dataset contains information on the industrial sector in which a business is started. The sectors are classified following the ISIC Rev.3 classification until 2008 and the ISIC Rev.4 classification from 2009 onwards. We complement the analysis with two sector-level indicators that are related to the financing needs of firms and to the collateralizability of their assets.

First, Kroszner et al. (2007) provide a version of the Rajan and Zingales indicator of external financial dependence (EFD) for manufacturing sectors under the ISIC Rev.2 classification. EFD is defined as the fraction of capital expenditures not financed with cash flows from operations. It is computed based on US data and constant for each sector across time, as it is intended to capture differences in external financing needs caused by technological differences across sectors, such as the length of the gestation periods of the projects. We match these data to the sector variable of the GEM, obtaining information on EFD for approximately 2,000 manufacturing startups (5.4% of all business started). We use this information to classify startups into sectors with low or high EFD, where the latter proxy for sectors with higher external financing needs (a high value of $\kappa - a$ in the model).

Second, Caggese and Perez (2017) use Compustat data to compute an indicator of the share of intangible over total assets for US industrial sectors. We match their sectors to the sector variable of the GEM, obtaining information on the sector-level share of intangible assets for approximately 17,000 startups (54% of all businesses started). We use this information to classify startups into sectors with a high or low share of intangible assets. Several authors argue that intangible assets have low collateral value, and therefore, we consider our category of *high intangibility* as a proxy for sectors with higher average costs of external finance (high r^b in the model). In other words, both high EFD and high intangibility might proxy for factors that increase the financial frictions of entrepreneurs and could be used as an additional test of the model. Note that the high EFD and high intangibility categories are quite uncorrelated (the correlation coefficient is 0.14). This is reasonable because they are conceptually different; this is also a desirable property since it implies that they provide independent sets of information.

We match the values for external dependence (1980-1999) from Table 12 of Kroszner et al. (2007) to the 22 manufacturing sectors identified in the GEM dataset. For the sectors that we can match across the Compustat SIC classification and the 2-digit sectors in the GEM dataset, we take the intangible capital share from Caggese and Perez (2017). We then calculate the median values for both measures and classify a sector as having high (low) external dependence or intangible share if its value is above (below) the median.

Table 17: External financial dependence, intangible asset share and startups by sector

Sector	Name	EFD	Intangible	# start-ups	% high growth
1	Agriculture and hunting	-	low	972	44.3
2	Forestry, logging and related service activities	-	-	79	49.2
5	Fishing	-	-	68	34
14	Other mining and quarrying	-	-	48	50.1
15	Food and Beverages	high	low	441	17.4
17	Textiles	high	high	102	23.9
18	Apparel	-	-	112	60.2
19	Leather	low	low	25	56.5
20	Wood products	high	low	122	41.4
21	Paper products	low	low	12	53.3
22	Printing and publishing	low	high	244	25.9
23	Petroleum and coal	high	low	10	9.3
24	Other chemical products	low	high	85	28.5
25	Rubber and plastic products	high	low	17	32
26	Non-metal products	low	low	67	50.5
27	Iron and steel	high	low	55	30.3
28	Metal products	low	high	87	59.8
29	Machinery	high	high	76	48
30	Office and computing	high	high	16	29.9
31	Electrical machinery	high	high	42	71.4
32	Radio	high	high	16	31.1
33	Professional equipment	high	high	33	30.9
34	Motover vehicles, trailers	low	low	46	11.4
35	Other transport equipment	low	high	22	51.5
36	Furniture	low	high	503	20.7
37	Recycling	-	high	25	13.9
40	Electricity, gas, steam	-	-	167	37.1
41	Collection, purification and distribution of water	-	-	12	44.6
45	Construction	-	high	1774	32.2
50	Sale, maintenance, repair of motor vehicles	-	low	769	37.1
51	Wholesale and commission trade	-	high	1280	26
52	Retail trade	-	low	4297	33.2
55	Hotels and restaurants	-	low	2156	35.5
60	Land transport; transport via pipelines	-	-	523	29.6
61	Water transport	-	-	15	23.1
63	Supporting and auxiliary transport activities	-	-	381	40.9
64	Post and telecommunications	-	-	178	39
71	Renting of machinery and equipment	-	high	85	30
72	Computer and related activities	-	high	1066	29
73	Research and development	-	high	87	55
85	Health and social work	-	low	1839	31.1
90	Sewage and refuse disposal, sanitation	-	-	125	47.8
91	Activities of membership organizations n.e.c.	-	-	60	25
92	Recreational, cultural and sporting activities	-	low	1454	35.9
93	Other service activities	-	-	1169	41
95	Activities of private households as employers of domestic staff	-	-	31	43.5
Total				20793	33.9

Notes: External financial dependence based on Kroszner et al. (2007) and intangible share based on Caggese and Perez (2017).

C Additional figures and tables

Figure 11: Predicted corporate GZ spread by country

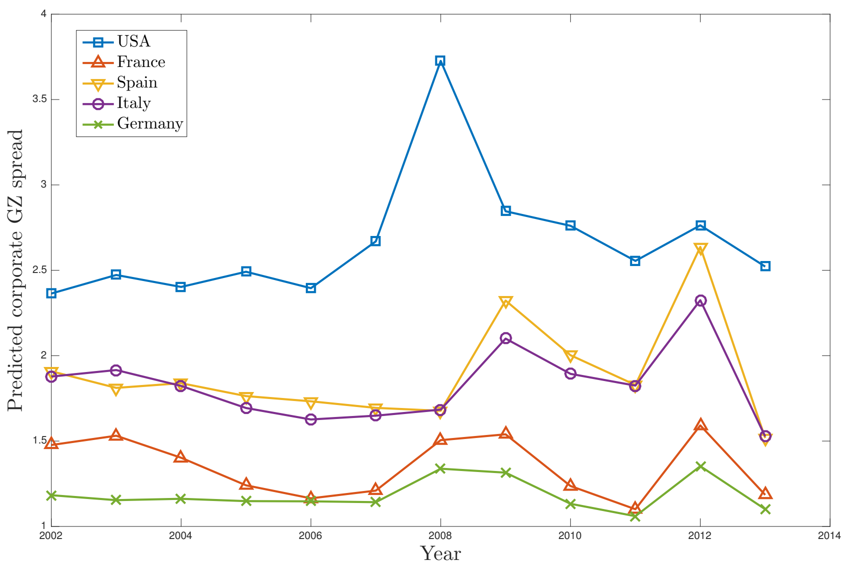
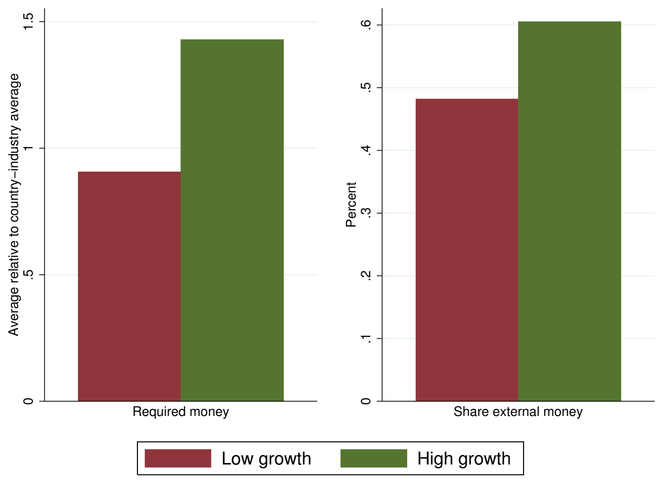


Figure 12: Financing needs of low-growth and high-growth startups



Notes: The bars on the left indicate, relative to the country-industry average, the average required money to start a business. The bars on the right indicate the average shares of external money needed (one minus the entrepreneur's own money divided by the required money). The observations in which own money > required money are dropped.

Table 18: Share of high-growth startups at firm creation and employment from SABI

	(1) Empl. growth	(2) Empl. growth	(3) Empl. growth	(4) Employment	(5) Employment	(6) Employment
Age 0				5.235***	5.591***	5.633***
				(0.1180)	(0.6739)	(0.6740)
Age 1	1.251***	1.227***	1.224***	6.848***	7.139***	7.246***
	(0.0301)	(0.2122)	(0.2123)	(0.1215)	(0.6740)	(0.6742)
Age 2	0.512***	0.493**	0.503**	7.325***	7.590***	7.687***
	(0.0274)	(0.2118)	(0.2118)	(0.1263)	(0.6747)	(0.6752)
Age 3	0.384***	0.396*	0.401*	7.389***	7.763***	7.835***
	(0.0271)	(0.2117)	(0.2117)	(0.1289)	(0.6751)	(0.6751)
Age 4	0.321***	0.356*	0.360*	7.307***	7.874***	7.923***
	(0.0271)	(0.2117)	(0.2117)	(0.1347)	(0.6766)	(0.6763)
Age 5	0.309***	0.339	0.343	7.224***	7.943***	8.007***
	(0.0272)	(0.2117)	(0.2117)	(0.1438)	(0.6787)	(0.6776)
Age 6	0.276***	0.301	0.302	6.944***	7.768***	7.817***
	(0.0274)	(0.2117)	(0.2117)	(0.1457)	(0.6792)	(0.6794)
Age 7	0.255***	0.282	0.289	6.691***	7.609***	7.632***
	(0.0275)	(0.2117)	(0.2118)	(0.1544)	(0.6812)	(0.6823)
Age 8	0.251***	0.284	0.297	6.665***	7.693***	7.697***
	(0.0277)	(0.2118)	(0.2120)	(0.1790)	(0.6883)	(0.6983)
Age 9	0.258***	0.253	0.274	6.668***	7.694***	7.984***
	(0.0281)	(0.2118)	(0.2122)	(0.1866)	(0.6904)	(0.7282)
Age 10	0.267***	0.246	0.226	6.596***	7.591***	8.130***
	(0.0289)	(0.2119)	(0.2124)	(0.1965)	(0.6931)	(0.7383)
Age 0 x share				-1.284***	-0.588***	-0.681***
				(0.1081)	(0.1363)	(0.1384)
Age 1 x share	-0.315***	-0.206***	-0.225***	-1.812***	-0.959***	-0.910***
	(0.0355)	(0.0399)	(0.0406)	(0.1081)	(0.1238)	(0.1271)
Age 2 x share	-0.040**	0.060***	0.071***	-1.708***	-0.821***	-0.766***
	(0.0164)	(0.0201)	(0.0205)	(0.1230)	(0.1346)	(0.1353)
Age 3 x share	0.010	0.033**	0.037**	-1.227***	-0.639***	-0.589***
	(0.0132)	(0.0153)	(0.0155)	(0.1296)	(0.1412)	(0.1451)
Age 4 x share	0.057***	0.020	0.022	-0.540***	-0.418***	-0.407**
	(0.0129)	(0.0145)	(0.0147)	(0.1408)	(0.1539)	(0.1606)
Age 5 x share	0.026**	-0.001	0.001	-0.133	-0.379**	-0.369**
	(0.0132)	(0.0148)	(0.0149)	(0.1561)	(0.1709)	(0.1794)
Age 6 x share	0.052***	0.037**	0.034**	0.553***	0.084	0.117
	(0.0144)	(0.0158)	(0.0160)	(0.1744)	(0.1911)	(0.1951)
Age 7 x share	0.070***	0.047***	0.047***	1.163***	0.511**	0.533***
	(0.0146)	(0.0158)	(0.0158)	(0.1879)	(0.2049)	(0.2055)
Age 8 x share	0.065***	0.027	0.026	1.469***	0.618**	0.616**
	(0.0154)	(0.0179)	(0.0182)	(0.2373)	(0.2587)	(0.2521)
Age 9 x share	0.019	0.061***	0.059***	1.756***	0.904***	0.861***
	(0.0180)	(0.0207)	(0.0209)	(0.2648)	(0.2865)	(0.2861)
Age 10 x share	-0.020	0.057**	0.063**	1.819***	1.064***	0.979***
	(0.0227)	(0.0242)	(0.0246)	(0.2946)	(0.3155)	(0.3190)
Year FE	Yes	No	No	Yes	No	No
Sector FE	Yes	No	No	Yes	No	No
Year-sector FE	No	Yes	Yes	No	Yes	Yes
Age-growth interactions	No	No	Yes	No	No	Yes
Observations	706578	706578	706578	947696	947696	947696
R-squared	0.110	0.113	0.113	0.149	0.150	0.150

Notes: In columns 1-3, the dependent variable is the yearly employment growth of firms established in 2003 or later; 0.1% of the tails are winsorized. In columns 4-6, the dependent variable is the log employment level. *share* is the share of high-growth startups in the 2-digit sector to which the firm belongs in the year it was born. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 19: Share of high-growth and innovative startups at firm creation and exit rate

	(1) Share high-growth	(2) Share high-growth	(3) Share innovative	(4) Share innovative
Age 0	0.076***	0.062***	0.087***	0.087***
	(0.0045)	(0.0223)	(0.0039)	(0.0219)
Age 1	0.114***	0.104***	0.112***	0.114***
	(0.0041)	(0.0221)	(0.0035)	(0.0220)
Age 2	0.124***	0.110***	0.108***	0.110***
	(0.0040)	(0.0221)	(0.0035)	(0.0220)
Age 3	0.122***	0.103***	0.107***	0.107***
	(0.0043)	(0.0221)	(0.0037)	(0.0220)
Age 4	0.113***	0.101***	0.101***	0.099***
	(0.0045)	(0.0222)	(0.0038)	(0.0220)
Age 5	0.101***	0.093***	0.098***	0.093***
	(0.0048)	(0.0222)	(0.0040)	(0.0220)
Age 6	0.107***	0.102***	0.092***	0.085***
	(0.0052)	(0.0223)	(0.0041)	(0.0220)
Age 7	0.100***	0.099***	0.090***	0.086***
	(0.0058)	(0.0225)	(0.0045)	(0.0221)
Age 8	0.082***	0.093***	0.087***	0.085***
	(0.0060)	(0.0226)	(0.0046)	(0.0222)
Age 9	0.064***	0.083***	0.072***	0.075***
	(0.0066)	(0.0227)	(0.0049)	(0.0222)
Age 10	0.045***	0.071***	0.061***	0.064***
	(0.0077)	(0.0231)	(0.0054)	(0.0223)
Age 0 x share	0.022***	0.041***	-0.004	-0.007
	(0.0072)	(0.0080)	(0.0074)	(0.0079)
Age 1 x share	0.001	0.010*	0.015**	0.002
	(0.0054)	(0.0061)	(0.0060)	(0.0064)
Age 2 x share	-0.022***	-0.005	0.036***	0.021***
	(0.0052)	(0.0057)	(0.0058)	(0.0063)
Age 3 x share	-0.031***	-0.004	0.015**	0.008
	(0.0056)	(0.0061)	(0.0061)	(0.0066)
Age 4 x share	-0.021***	-0.012*	0.023***	0.018**
	(0.0061)	(0.0066)	(0.0067)	(0.0071)
Age 5 x share	-0.013*	-0.013*	-0.001	0.008
	(0.0069)	(0.0075)	(0.0074)	(0.0078)
Age 6 x share	-0.040***	-0.047***	-0.002	0.016*
	(0.0077)	(0.0083)	(0.0083)	(0.0089)
Age 7 x share	-0.031***	-0.044***	-0.007	0.003
	(0.0090)	(0.0099)	(0.0118)	(0.0131)
Age 8 x share	-0.009	-0.048***	-0.038***	-0.031**
	(0.0096)	(0.0105)	(0.0131)	(0.0146)
Age 9 x share	0.027**	-0.031***	0.033**	0.013
	(0.0109)	(0.0118)	(0.0163)	(0.0180)
Age 10 x share	0.049***	-0.020	0.050***	0.028
	(0.0131)	(0.0140)	(0.0179)	(0.0199)
Year FE	Yes	No	Yes	No
Sector FE	Yes	No	Yes	No
Year-sector FE	No	Yes	No	Yes
Observations	847250	847250	847250	847250
R-squared	0.175	0.177	0.175	0.177

Notes: The dependent variable is a dummy indicating the exit of the firm in the SABI data. In columns 1-2, *share* is the share of high-growth startups, and in columns 3-4, *share* is the share of innovative startups in the 2-digit sector to which the firm belongs in the year it was born. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Baseline results with Spain only

Table 20: Financial crisis, GZ spread and probability of starting a firm

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	5.458***	4.827***	5.143***	3.720***	3.519***	3.022***
	(0.4226)	(0.4616)	(0.7075)	(0.2301)	(0.2675)	(0.3377)
Fin. crisis	0.120***	0.102**	0.121*			
	(0.0384)	(0.0416)	(0.0655)			
GZ spread				-0.022***	-0.012	-0.039***
				(0.0069)	(0.0078)	(0.0111)
Observations	232749	232749	232749	232749	232749	232749
R-squared	0.029	0.024	0.026	0.029	0.024	0.027

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 21: Financial crisis, GZ spread and probability of starting a firm

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	2.434***	2.699***	1.312*	9.803***	8.045***	10.496***
	(0.4462)	(0.4963)	(0.6924)	(0.5536)	(0.6324)	(0.8548)
Fin. crisis	-0.106***	-0.054	-0.187***			
	(0.0374)	(0.0416)	(0.0587)			
Fin. crisis x GDP growth	11.280***	7.881***	15.164***			
	(0.8849)	(0.9938)	(1.4373)			
GZ spread				-0.012	-0.004	-0.032**
				(0.0077)	(0.0084)	(0.0136)
GZ spread x GDP growth				4.339***	3.224***	5.351***
				(0.3518)	(0.3967)	(0.5646)
Observations	232749	232749	232749	232749	232749	232749
R-squared	0.035	0.028	0.038	0.034	0.027	0.035

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Estimating predicted bond spreads

Table 22: Monetary policy shocks and bond spreads

	(1) Bank GZ spread		(2) Corporate GZ spread	
	Pre-2007	Post-2007	Pre-2007	Post-2007
USA x MP shock	0.256	-3.265	-0.284	0.419
	(1.1463)	(11.1314)	(1.5745)	(5.8500)
USA x MP shock (t-1)	0.510	-7.599	-0.151	-1.918
	(1.2386)	(10.1838)	(1.6382)	(5.7117)
USA x MP shock (t-2)	-0.012	-12.224	-1.437	-4.120
	(1.4395)	(8.7537)	(1.4265)	(4.7421)
USA x MP shock (t-3)	-0.083	-8.250	-1.184	0.251
	(1.7215)	(8.6156)	(1.8761)	(4.8331)
FRA	-0.676***		-1.254***	
	(0.1467)		(0.1062)	
FRA x MP shock	-0.239	8.762**	3.915**	8.754**
	(1.6338)	(4.3539)	(1.9888)	(3.4593)
FRA x MP shock (t-1)	-1.073	8.870**	3.132	9.957***
	(1.7194)	(4.4624)	(2.0734)	(3.2251)
FRA x MP shock (t-2)	-0.672	6.236	3.162*	6.413*
	(1.6659)	(5.3425)	(1.6858)	(3.7183)
FRA x MP shock (t-3)	-0.578	-1.068	3.526*	0.846
	(1.6898)	(7.4470)	(1.8081)	(4.6806)
SPA	0.423**		-0.740***	
	(0.1855)		(0.1229)	
SPA x MP shock	-2.858	20.066**	3.550	11.304**
	(3.7938)	(9.7093)	(2.1842)	(5.4775)
SPA x MP shock (t-1)	-4.356	19.663*	2.163	10.527*
	(3.9906)	(10.3961)	(2.3772)	(5.4396)
SPA x MP shock (t-2)	-4.393	18.982*	1.019	7.867
	(3.5555)	(11.4578)	(2.2461)	(5.8141)
SPA x MP shock (t-3)	-3.842	3.357	1.479	0.108
	(3.5470)	(14.1604)	(2.4530)	(7.6492)
ITA	-0.207		-0.809***	
	(0.1869)		(0.1249)	
ITA x MP shock	-1.693	12.690	2.819	10.587*
	(3.4673)	(9.3929)	(2.3463)	(6.3210)
ITA x MP shock (t-1)	-3.177	11.408	1.977	9.157
	(3.6267)	(9.0966)	(2.4965)	(6.1618)
ITA x MP shock (t-2)	-3.105	7.548	2.297	5.416
	(3.3502)	(10.7918)	(2.3343)	(6.8756)
ITA x MP shock (t-3)	-2.929	-8.030	3.070	-4.266
	(3.3866)	(13.8543)	(2.3956)	(9.2036)
GER	-0.900***		-1.366***	
	(0.1388)		(0.1016)	
GER x MP shock	-0.025	5.827*	1.479	6.780**
	(1.3233)	(3.1741)	(1.2631)	(3.0341)
GER x MP shock (t-1)	-0.755	5.908*	0.363	7.364***
	(1.4163)	(3.3913)	(1.2750)	(2.8148)
GER x MP shock (t-2)	-0.926	4.935	0.479	4.410
	(1.3521)	(4.1950)	(1.1461)	(3.4404)
GER x MP shock (t-3)	-0.913	-0.751	0.584	-0.953
	(1.3294)	(5.3414)	(1.0927)	(4.5946)
Observations	821		833	
R-squared	0.333		0.384	
F-statistic	5.8		10.8	

Notes: Columns Pre-2007 and Post-2007 show the coefficients interacted with a dummy indicating the respective period (2007 is included in Pre-2007). The country fixed effects are restricted to be the same across periods. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Estimating predicted bond spreads: additional results

Table 23: Spain only

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	3.742***	3.361***	3.439***	4.718***	4.312***	4.298***
	(0.1405)	(0.1603)	(0.2016)	(0.3003)	(0.3401)	(0.4626)
Predicted GZ spread	-0.222***	-0.159***	-0.311***	-0.184***	-0.122***	-0.278***
	(0.0232)	(0.0261)	(0.0347)	(0.0243)	(0.0279)	(0.0380)
Pr. GZ spread x GDP growth				6.693***	6.562***	5.778**
				(1.8572)	(2.1255)	(2.9008)
Observations	233625	233625	233625	233625	233625	233625
R-squared	0.031	0.025	0.032	0.032	0.026	0.032

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. The standard errors are obtained by using a bootstrap procedure with 5000 replications. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 24: Using bank spread instead of corporate spread

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	3.511***	3.135***	3.367***	3.348***	3.031***	3.125***
	(0.1254)	(0.1433)	(0.1902)	(0.1354)	(0.1569)	(0.2224)
Predicted GZ spread	-0.103***	-0.074***	-0.145***	-0.102***	-0.073***	-0.145***
	(0.0089)	(0.0099)	(0.0146)	(0.0088)	(0.0100)	(0.0149)
Pr. GZ spread x GDP growth				0.711***	0.474	0.935**
				(0.2669)	(0.2978)	(0.4505)
Observations	399494	399494	399494	399494	399494	399494
R-squared	0.041	0.037	0.040	0.041	0.038	0.040

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. The standard errors are obtained by using a bootstrap procedure with 5000 replications. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 25: Effects of MP shocks same pre- and post-2007

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	4.147***	3.658***	4.006***	4.550***	3.961***	4.531***
	(0.1379)	(0.1594)	(0.2083)	(0.1491)	(0.1724)	(0.2276)
Predicted GZ spread	0.080***	0.080**	0.054	0.107***	0.104***	0.066
	(0.0299)	(0.0333)	(0.0482)	(0.0309)	(0.0350)	(0.0521)
Pr. GZ spread x GDP growth				3.079***	2.257***	4.410***
				(0.4755)	(0.5133)	(0.8792)
Observations	399462	399462	399462	399462	399462	399462
R-squared	0.039	0.037	0.037	0.040	0.037	0.038

* p<0.1, ** p<0.05, *** p<0.01

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. The standard errors are obtained by using a bootstrap procedure with 5000 replications. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 26: Using bank spread and effects of MP shocks same pre- and post-2007

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	3.972***	3.541***	3.736***	3.532***	3.221***	3.082***
	(0.1328)	(0.1494)	(0.2011)	(0.1367)	(0.1547)	(0.2125)
Predicted GZ spread	-0.002	0.021	-0.067	-0.015	0.013	-0.095**
	(0.0263)	(0.0284)	(0.0453)	(0.0273)	(0.0296)	(0.0481)
Pr. GZ spread x GDP growth				3.103***	2.407***	3.871***
				(0.2977)	(0.3269)	(0.4813)
Observations	399494	399494	399494	399494	399494	399494
R-squared	0.040	0.037	0.037	0.041	0.037	0.038

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. The standard errors are obtained by using a bootstrap procedure with 5000 replications. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Additional robustness checks

Table 27: Excluding countries without financial crisis

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	1.163**	1.329**	0.500	5.447***	4.418**	6.182***
	(0.5501)	(0.5604)	(0.4140)	(1.9928)	(1.7647)	(1.6892)
Fin. crisis	-0.138***	-0.112***	-0.160***			
	(0.0392)	(0.0333)	(0.0517)			
Fin. crisis x GDP growth	4.096***	3.455***	4.490*			
	(1.5200)	(1.2040)	(2.2918)			
GZ spread				-0.020	-0.013	-0.033*
				(0.0197)	(0.0189)	(0.0173)
GZ spread x GDP growth				2.450	1.532	3.829**
				(1.6126)	(1.3356)	(1.5513)
Observations	800019	800019	800019	370280	370280	370280
R-squared	0.042	0.038	0.041	0.039	0.035	0.039

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 28: Excluding construction sector

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	0.689	1.027	-0.061	4.979***	4.034**	5.694***
	(0.7050)	(0.6652)	(0.5285)	(1.8103)	(1.6127)	(1.5254)
Fin. crisis	-0.152***	-0.116***	-0.185***			
	(0.0490)	(0.0400)	(0.0592)			
Fin. crisis x GDP growth	4.203***	3.371***	4.806**			
	(1.6229)	(1.2837)	(2.2861)			
GZ spread				-0.018	-0.010	-0.033**
				(0.0184)	(0.0180)	(0.0156)
GZ spread x GDP growth				2.076	1.191	3.487**
				(1.3994)	(1.1551)	(1.3585)
Observations	891932	891932	891932	369436	369436	369436
R-squared	0.060	0.045	0.075	0.035	0.031	0.036

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 29: Excluding startups that have paid wages

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	0.458	0.778	-0.219	2.757**	2.430**	2.881***
	(0.4883)	(0.5092)	(0.2854)	(1.0722)	(1.0755)	(0.7090)
Fin. crisis	-0.094***	-0.071**	-0.117***			
	(0.0311)	(0.0294)	(0.0434)			
Fin. crisis x GDP growth	2.974***	2.556**	3.021***			
	(0.9589)	(1.0006)	(1.0331)			
GZ spread				-0.015	-0.007	-0.032**
				(0.0126)	(0.0134)	(0.0123)
GZ spread x GDP growth				1.088	0.650	1.922***
				(0.8799)	(0.8450)	(0.7090)
Observations	888862	888862	888862	367460	367460	367460
R-squared	0.054	0.040	0.072	0.029	0.028	0.026

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 30: Including year fixed effects

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	0.855	1.134*	0.261	2.757***	2.198***	3.014***
	(0.6419)	(0.6178)	(0.4701)	(0.7804)	(0.7086)	(0.8113)
Fin. crisis	-0.120	-0.072	-0.179*			
	(0.1110)	(0.0970)	(0.1053)			
Fin. crisis x GDP growth	3.005***	2.483***	3.195**			
	(1.0741)	(0.8860)	(1.4086)			
GZ spread				-0.022	-0.015	-0.041***
				(0.0157)	(0.0176)	(0.0142)
GZ spread x GDP growth				1.652**	0.983	2.691***
				(0.7490)	(0.6163)	(0.8906)
Observations	894126	894126	894126	370280	370280	370280
R-squared	0.065	0.049	0.081	0.049	0.043	0.050

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 31: Including dummy for Great Recession

	(1) All	(2) Low growth	(3) High growth
GDP growth	2.006	1.740*	1.828
	(1.4180)	(1.0438)	(1.5625)
Great Recession	0.039	0.024	0.045
	(0.0948)	(0.0724)	(0.0999)
GR x GDP growth	3.949***	3.211**	4.131***
	(1.3555)	(1.2721)	(1.3530)
Observations	894126	894126	894126
R-squared	0.061	0.041	0.072

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 32: Heckman selection model

	(1)	(2)	(3)	(4)	(5)	(6)
GDP growth	-0.582**	-0.793***	-0.516	2.201**	-0.290	-0.057
	(0.2685)	(0.2987)	(0.6565)	(0.9871)	(0.3910)	(0.4056)
Fin. crisis	-0.032	-0.052*				
	(0.0242)	(0.0275)				
Fin. crisis x GDP growth		1.249*				
		(0.7539)				
GZ spread			-0.042**	-0.031		
			(0.0196)	(0.0199)		
GZ spread x GDP growth				2.745***		
				(0.7375)		
RR indicator					0.003	0.001
					(0.0052)	(0.0054)
RR indicator x GDP growth						0.306**
						(0.1366)
Observations	894126	894126	370280	370280	731881	731881

Notes: The first-stage selection equation for starting a business includes sex, education, age and country dummies. The second-stage equation for starting a high-growth business includes country dummies in addition to the reported variables. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 33: Including riskless interest rate

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	2.455***	2.450***	1.792**	3.266***	3.079***	2.683***
	(0.7229)	(0.6058)	(0.7836)	(0.4625)	(0.4856)	(0.2721)
Fin. crisis	-0.035	-0.022	-0.056			
	(0.0438)	(0.0420)	(0.0399)			
GZ spread				-0.011	-0.003	-0.025***
				(0.0091)	(0.0093)	(0.0067)
Riskless interest rate	0.032***	0.027***	0.035***	0.052***	0.046***	0.048***
	(0.0080)	(0.0072)	(0.0092)	(0.0048)	(0.0065)	(0.0060)
Observations	816895	816895	816895	370280	370280	370280
R-squared	0.042	0.038	0.041	0.039	0.036	0.036

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 34: Including riskless interest rate interacted with GDP growth

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	1.303**	1.518***	0.510	6.221***	5.059***	7.069***
	(0.5113)	(0.4861)	(0.4908)	(1.8901)	(1.6675)	(1.7455)
Fin. crisis	-0.111**	-0.085**	-0.141**			
	(0.0442)	(0.0361)	(0.0597)			
Fin. crisis x GDP growth	3.882**	3.200***	4.400*			
	(1.6439)	(1.2104)	(2.5728)			
GZ spread				0.014	0.015	0.004
				(0.0178)	(0.0176)	(0.0159)
GZ spread x GDP growth				2.593**	1.691*	3.901***
				(1.2518)	(0.9988)	(1.3746)
Riskless interest rate	0.039***	0.034***	0.038**	0.075***	0.064***	0.075***
	(0.0126)	(0.0099)	(0.0165)	(0.0048)	(0.0060)	(0.0060)
RIR x GDP growth	-0.201	-0.242	-0.079	-0.861*	-0.650	-1.107***
	(0.2497)	(0.2095)	(0.2878)	(0.4673)	(0.4875)	(0.3169)
Observations	816895	816895	816895	370280	370280	370280
R-squared	0.043	0.039	0.043	0.041	0.037	0.041

Notes: The dependent variable is a dummy that is equal to one if an individual is a nascent entrepreneur in the respective category. The controls include dummies for three education levels, sex, age and country fixed effects. Significance levels: * $p<0.1$, ** $p<0.05$, *** $p<0.01$.