The Review of Financial Studies

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

Christoph Albert
CEMFI

Andrea Caggese
UPF, Barcelona GSE, and CREI

We analyze a multiyear, multicountry entrepreneurship survey with more than one million observations to identify startups with low and high growth potential. 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. (*JEL* E32, D22, M13)

Received June 15, 2019; editorial decision July 6, 2020 by Editor Holger Mueller. Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web site next to the link to the final published paper online.

A well-established literature documents the importance of financial frictions for entrepreneurial entry and for the survival and growth of new firms (see Holtz-Eakin, Joulfaian, and Rosen 1994; Blanchflower and Oswald 1998; Corradin and Popov 2015; Schmalz, Sraer, and Thesmar 2017; Adelino et al. 2015, among others). However, less is known about the relation between financial factors, the

For their useful comments and suggestions, we thank Manuel Arellano, Davide Furceri, Peter Karadi, Geert Mesters, and Kirill Shakhnov (discussant); the participants at the 2018 Annual Conference of the Society for Economic Dynamics in Mexico City, the 6th ECB-CBRT Conference in Izmir, the 2019 Workshop in Macroeconomics in Marrakech, the 2019 Catalan Economic Society Conference in Barcelona, the 7th Macro Banking & Finance Workshop at Collegio Carlo Alberto; and the attendees at the seminars at Manchester University, Danmarks Nationalbank, and CREI. We thank Cristiano Mantovani and Marta Morazzoni for their excellent work as research assistants. All errors are our own. A previous version of this paper was entitled "Financial Frictions, Cyclical Fluctuations, and the Growth Potential of New Firms." Financial support from La Caixa Foundation [grant no. LCF/PR/RC14/10080001] is gratefully acknowledged. Andrea Caggese acknowledges financial support from the Spanish Ministry of Economy and Competitiveness [AEI-FEDER project UE-ECO2017-82596-P] and through the Severo Ochoa Programme for Centres of Excellence in R&D [SEV-2015-0563]. Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to Andrea Caggese, andrea.caggese@upf.edu.

The Review of Financial Studies 34 (2021) 2508–2548

doi:10.1093/rfs/hhaa112 Advance Access publication September 30, 2020

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 grows very rapidly, driving a higher mean net employment growth for younger firms than for older firms. Pugsley, Sedlacek 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? And 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 data set is drawn from the Global Entrepreneurship Monitor (GEM), a multicountry 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 data set 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 (gross domestic product [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 data set particularly suited for our purpose. First, it includes an individual's personal 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 survey responses of entrepreneurs in the GEM matches remarkably well with that obtained from administrative data sources. Third, it includes survey questions to ascertain the expected employment growth of new startups and the innovative nature of the products and services that will be offered; we use the survey questions to identify startups with high growth potential.

To formalize the intuition behind 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 the number of type 2 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 than for type 1 startups.

To test these predictions, we identify type 2 startups in the GEM data set as those businesses for which the entrepreneur is expecting high future employment (relative to the average employment of established firms in the same country and 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 two-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. Despite being limited to Spain, the matched firm-level data set is sufficiently representative for our purposes. Spain is the country with the most extensive coverage in GEM, with more than 200,000 observations. Indeed, all the main results we later obtain from the entire data set are also confirmed when considering only the Spanish GEM surveys. The matched sample includes 46 two-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 6 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.

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 by fluctuations in the excess cost of external finance. Our preferred indicator is the Gilchrist and Zakrajsek (2012) bond spread for financial institutions. Using additional data on European countries from Gilchrist and Mojon (2016), we compute the indicator for the United States, 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 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.

As an additional test of our hypothesis, we consider two indicators often used in the literature to determine the sectors that are most 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, Kadyrzhanova, and Sim 2013; Caggese and Perez 2017). The model predicts that startups in sectors with higher indicators should be more negatively affected by financial shocks, and we confirm these predictions in the data. Furthermore, our results are also confirmed when we control for the risk-free interest rate or variations in the term premium, which could differently affect the expected value of the different startup types, and when we control for additional individual-level characteristics, such as expectations about future business opportunities, income category, and previous entrepreneurial experience.

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 a recessionary period accompanied by a one-percentage-point increase in the bond spread changes the nature of newly created firms such that after 10 years, the employment level in these firms is on average 4.3% lower.

1. Related Literature

This paper is related to the large literature documenting the importance of financial constraints as a key factor influencing entrepreneurial entry. Holtz-Eakin, Joulfaian, and Rosen (1994) and Blanchflower and Oswald (1998), among others, 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, Schoar, and Severino (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, Sraer, and Thesmar (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 (2012) 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 (see Buera, Kaboski, and Shin 2011; Caggese and Cunat 2013; Midrigan and Xu 2014; Cole et al. 2016, among others).

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 entrepreneurial entry 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 (2019) 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 Sedlacek and Sterk (2017), who show that not only did firm entry strongly decline during the 2007–2009 financial crisis but also 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 entrepreneurial decisions to create different types of businesses while controlling for the quality of the entrepreneurial pool.[1]

Finally, our empirical analysis is related to studies, particularly Braun and Larrain (2005), Kroszner, Laeven, and Klingebiel (2007), and Dell'Ariccia, Detragiache, and Rajan (2008), that use multicountry and multisector 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.

2. 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 startups, 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 Benoit (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 data set, where entrepreneurs finance on average around 50% of their startup costs with external financing sources (see Figure F1 in the Internet 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 housing price channel could be introduced in the model by assuming that a higher cost of external finance increases the costs of mortgages, which reduces housing prices and thus 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 4.1, we confirm that our data set is able to identify these different startup types.

[1] Entrepreneurial choices among heterogeneous individuals also have been extensively analyzed in the occupational choice and innovation literatures (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. 2020). Other authors focus on the mobility of inventors and disruptive innovators and on the reallocation of highly skilled labor (see Acemoglu, Akcigit, and Celik 2014; Akcigit and Kerr 2018, among others).

2.1 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. Every period, firms exit with a certain probability. A Type j firm that does not exit in period t generates profits:

$$\pi_{j,t}=\theta_{j,t}^{\beta}L_{j,t}^{\alpha}-wL_{j,t}, \tag{1}$$

where $\theta_{j,t}$ is productivity, $L_{j,t}$ is labor input, w is the exogenously given wage, and $0<\alpha<1$, $0<\beta\leq 1$.[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 $L_{j,t}=\left(\frac{\alpha}{w}\right)^{\frac{1}{1-\alpha}}\theta_{j,t}$. Substituting this equation in Equation (1), we express profits as a linear function of $\theta_{j,t}$:

$$\pi(\theta_{j,t})=\Psi\theta_{j,t} \tag{2}$$

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

Startup types differ in their expected productivity growth:

Type $j=1$ indicates a startup with low growth potential, for which productivity $\theta_{1,t}$ grows at an exogenous gross rate m in all periods, so that $\theta_{1,t+1}=m\theta_{1,t}$. Starting a type 1 business represents 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 $j=2$ indicates a startup with high growth potential. Its productivity grows at the gross rate $l\leq m$ initially, but every year, with probability γ, the growth rate permanently increases from l to $h>m$. 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 than starting a type 1 business, and more time is required for the business to start generating revenues; however, the business has higher long-run growth potential.[3]

We introduce heterogeneity across entrepreneurs by assuming that their productivity is a function of their skills:

$$\theta_{i,j,0}=\phi_{i,j}E_i, \tag{3}$$

where $\theta_{i,j,0}$ is the initial productivity of Type j for entrepreneur i. E_i is the entrepreneur's generic skills, and $\phi_{i,j}$ the skills specific to type j projects. We

[2] $\theta_{j,t}$ 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.

[3] The growth potential of type 2 projects might also depend on different managerial and organizational strategies. For example, a restaurant owner might choose between managing a small traditional family-style restaurant or developing a new restaurant chain.

assume that E_i is uniformly distributed across entrepreneurs, $E_i \in [1-e, 1+e]$, with $0<e<1$. The skills required to operate type 2 firms, $\phi_{i,2}$ 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 $\phi_{i,2}$ determines the probability of starting a type 2 firm rather than a type 1 firm.[4]

2.2 Financing

Entrepreneurs have an initial endowment of $a \leq \kappa$ and need to borrow $b=\kappa-a$. In subsequent periods, debt can be repaid by using the flow of profits $\pi\left(\theta_{i,j,t}\right)$. One unit of debt implies a repayment of $1+r+r^b$ in the following period, where r is the borrowing rate in the absence of financial imperfections, and r^b can be interpreted as the financial spread or excess cost of debt caused by financial frictions.

2.3 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 finance a and the excess cost of external finance r^b. To better illustrate the economic intuition, and provide an analytical proof for all the propositions, we initially make the simplifying assumption that firms live deterministically for two periods. We then consider a more realistic version of the model with firms that live many periods and face a per-period stochastic death probability d. The general model also satisfies all the propositions for the calibrated realistic parameter values. To ease notation, we henceforth drop the i subscript.

2.3.1 No financial frictions. Access to finance is not a problem, if either $a<\kappa$ but $r^b=0$, meaning that the entrepreneur can borrow at the frictionless rate r, 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 2-period version of the model, a type 1 firm grows at the gross rate m both periods. Normalizing r to zero, its value for given initial productivity $\theta_{1,0}$ is

$$V^1=m\Psi\theta_{1,0}+m^2\Psi\theta_{1,0}. \tag{4}$$

Conversely, the productivity of a type 2 firm grows at the rate l, which we normalize to 1, in the first period, and with probability γ switches to the higher growth rate h in period 2. Therefore, its value, given initial productivity $\theta_{2,0}$, is

$$V^2=\Psi\theta_{2,0}+\gamma h\Psi\theta_{2,0}+(1-\gamma)\Psi\theta_{2,0}. \tag{5}$$

We make the following assumption.

---

[4] In Internet Appendix C, we show that the results are robust to replacing the uniform distribution of skills $\phi_{i,2}$ and E_i with a normal distribution. Furthermore, an alternative assumption to consider heterogeneous growth rates of productivity across entrepreneurs, rather than heterogeneous initial levels, would have similar implications.

Assumption 1. $2+\gamma(h-1)>m+m^2$

Assumption 1 implies that h is sufficiently high such that, for an entrepreneur who has the option to start the two types with the same initial productivity ($\theta_{1,0}=\theta_{2,0}$), a type 2 startup is more valuable than a type 1 startup in the absence of financial frictions. It then follows that:

Proposition 1. In the absence of financial frictions, there exists a threshold value $\overline{\phi}<1$ such that, for any generic skill value E, *entrepreneurs with skills* ϕ_2 above the threshold prefer a type 2 startup and entrepreneurs below the threshold prefer a type 1 startup.

Proof. For an entrepreneur with generic skills E and type 2 skills ϕ_2, initial productivities are $\theta_{1,0}=E$ and $\theta_{2,0}=E\phi_2$. Equations (4) and (5) imply that $V^2\geq V^1$ if $\phi_2\geq\overline{\phi}\equiv\frac{m+m^2}{2+\gamma(h-1)}$. Assumption 1 ensures that $\overline{\phi}<1$. Therefore, all entrepreneurs with $\phi_2\geq\overline{\phi}$ prefer type 2 to type 1 projects, regardless of their generic skills E. Given ϕ_2, it is possible to use (4) to obtain $E_{\min,2}(\phi_2)$, the minimum value of generic skills such that $V^2-\kappa\geq 0$ and the entrepreneur starts a type 2 firm:

$$E_{\min,2}(\phi_2)=\frac{\kappa}{\Psi\phi_2\left[2+\gamma(h-1).\right]} \tag{6}$$

Likewise, for all entrepreneurs with $\phi_2<\overline{\phi}$ there exists a minimum generic skill $E_{\min,1}$ such that $V^1-\kappa\geq 0$ and the entrepreneur starts a type 1 firm:

$$E_{\min,1}=\frac{\kappa}{\Psi\left(m+m^2\right)}. \tag{7}$$

For simplicity, we assume that the lower bound of the E_i distribution, $1-e$, is lower than the minimum thresholds of both types. ■

2.3.2 Financial frictions. Financial frictions matter if the entrepreneur needs to borrow $b=\kappa-a>0$ to start the firm, and external financing is costly, so that r^b>0. The larger b is, the longer it takes to repay the debt. To ensure a nontrivial case, we make the following assumption: b is sufficiently large so that at least for some type 1 and type 2 entrepreneurs full repayment takes two periods (see Internet Appendix B.1 for details):

Assumption 2. $b>\frac{\kappa}{(1+m)(1+r^b)}$

$C^1(\theta_{1,0},b)$ and $C^2(\theta_{2,0},b)$ denote the net present value of the expected excess financing costs given initial productivity and debt b, for a new business of type 1 and type 2, respectively. These values are computed as the net present value of the debt repayments minus the net present value of the repayments that would be due in the absence of financial frictions (see Equations (11) and (12) in Internet Appendix B.1 for details).

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 proof of Proposition 2, shown in detail in Internet Appendix B.1, is straightforward, since the excess costs $C^1(\theta_{1,0},b)$ and $C^2(\theta_{2,0},b)$ are zero in the absence of financial frictions ($r^b=0$), and both are increasing in r^b. An increase in these costs raises the minimum skill thresholds $E_{\min,1}$ and $E_{\min,2}(\phi_2)$, and reduces the proportion of entrepreneurs that start a business.

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:

Proposition 3. 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.

For a formal proof, see Internet Appendix B.1. Intuitively, the minimum generic skill threshold is higher for type 1 firms than for type 2 firms, $E_{\min,1} > \phi_2 E_{\min,2}(\phi_2)$, because the latter have higher growth prospects. Therefore, type 1 firms are initially more profitable on average and are able to repay a larger part of their debt in the first period, which reduces the amount of debt that needs to be rolled over to the next period and thus reduces the overall excess cost of finance. It follows that on average, type 2 entrepreneurs are more penalized by an increase in excess financing costs. A larger fraction of them do not start a firm, or switch from a type 2 to a type 1 firm. Proposition 3 considers an increase in the cost of external finance for given external financial needs b. The next proposition shows that these two factors interact with each other:

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

For a formal proof, see Internet Appendix B.1. Intuitively, an increase in financing needs increases the fraction of entrepreneurs that take two periods to repay the debt, and amplifies the differential effect of an increase in excess financing costs between the two types.

2.3.3 Calibrated multiperiod model. We graphically illustrate propositions 3 and 4 for the multiperiod version of the model, in which all firms face a constant death probability every period (see Internet Appendix B.2 for details). All the effects described above also extend to this model, and are quantitatively amplified by the fact that, for realistic parameter values, it takes more than two periods on average to repay the debt.

In particular, 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. We set $m=1.03$; therefore, the employment of type 1 firms grows on average at 3% every year, consistent with the median employment growth rate of U.S. firms.[5] For type 2 firms, the initial net growth rate is normalized to zero; hence, $l=1$. After switching, they grow at a rate of 4%; thus, $h=1.04$ until they die, and the switching probability γ is 20%, so that the resultant expected employment growth of type 2 relative to type 1 firms roughly matches the relative employment growth of the *high-growth* startups we identify from matching the GEM and SABI data sets (see Section 3.1 for details). The value of α is set to 0.6, 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 3.1 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.

Proposition 3 is illustrated in Figure 1, 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 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 $\kappa - a$.

Proposition 4 is illustrated in the left panel of Figure 2, which depicts the choice between startup types as a function of the cost of external finance r^b.

---

[5] The value comes from the authors' own calculations using Compustat data.

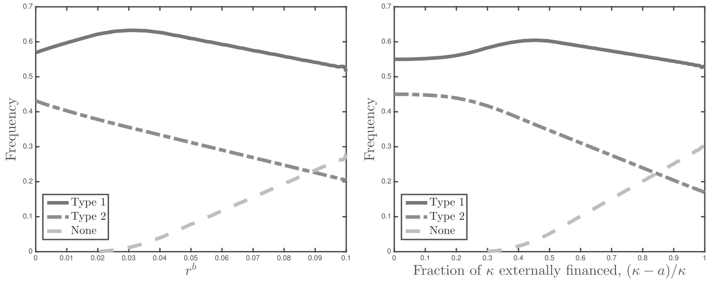

Figure 1
Predicted frequencies of startup types
$(\kappa - a)/\kappa$ is set to 0.75, and r^b is set to 0.075 to generate the frequencies in the left and right panels, respectively.

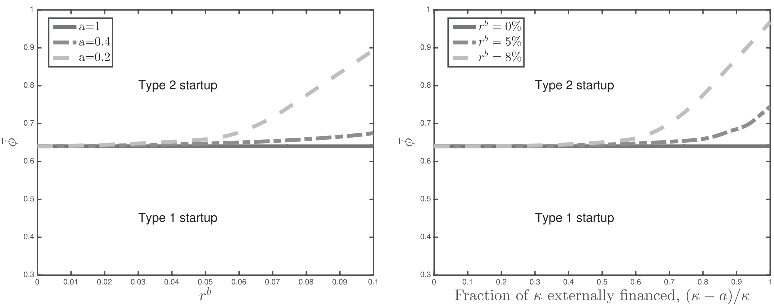

Figure 2
Interaction between financing needs and the cost of external finance
Each line represents the threshold value $\overline{\phi}$ given r^b and a. Entrepreneurs with $\phi_2 > \overline{\phi}$ prefer type 2 projects to type 1 projects. Both panels are generated for entrepreneurs with the average level of skills $E = 1$.

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 of finance is irrelevant to the choice of type of project, as stated in Proposition 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 2 considers the symmetric case of varying $\kappa - a$ for given levels of r^b.

2.4 Predictions

In the empirical section, we proxy r^b by indicators of the cost of external finance. 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 with GDP growth. We interpret a as funds 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.[6] Therefore, we can also test the following prediction.

Prediction 3. A decline in GDP growth increases the negative effects of r^b relatively more for type 2 startups than for 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 will be 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 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 affect high-growth startups more than low-growth ones.[7]

---

6 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 with income. Nonetheless, we believe that this assumption is without loss of generality. On the one hand, empirical models of households' precautionary savings show that households exhibit buffer stock behavior whereby their net financial wealth is highly sensitive to their 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. Moreover, Internet Appendix Figure F2 shows that the average share of external finance needed to start a business across all entrepreneurs in the GEM was considerably higher during the financial crisis and Table F2 in the Internet Appendix suggests that this share negatively correlates with GDP growth.

7 Another simplification of the model is that financial frictions affect only the initial startup cost, not variable production inputs. If capital is also used to produce output, and optimal capital investment is increasing in

Other simplifying assumptions of the model are that the riskless interest rate is kept constant, and that the law of motion of productivity $\theta_{j,t}$ is unaffected by aggregate shocks. Given their different intertemporal profiles, the two startup types would be affected differently by temporary fluctuations in the interest rate. In our empirical analysis, we show that all the results are robust to controlling for country-specific riskless interest rates, and for their term structure.

With respect to productivity, in Internet Appendix D, we introduce unexpected and permanent aggregate productivity shocks. We show that these shocks have a small effect on the relation between financial frictions and startup decisions and that this effect is symmetric for type 1 and type 2 projects, and therefore it does not affect the model's predictions.

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 long-term or uncertain projects and are therefore more likely to be subject to asymmetric information or other financial frictions. 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 5.1, 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.

3. Data

3.1 GEM data set

Our main data source is the GEM, the most comprehensive cross-country survey on entrepreneurial activity currently available (Reynolds and Hechavarria 2016). The GEM includes random samples of adult individuals from over 100 countries, with sample sizes ranging from approximately 1,000 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 the 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

productivity and potentially subject to financial frictions, then type 1 startups would be initially more affected, because their productivity initially grows faster. However, type 1 startups also repay their debt faster, and so, at some point, type 2 startups would become more affected, because they are still financially constrained, while type 1 startups are not. Intuitively, the longer it takes type 2 firms to accumulate enough savings to become unconstrained, relative to type 1 firms, the more likely it is that this additional constraint on capital reinforces the predictions of the model.

[8] The survey began in 1999, but the first 3 years have fewer observations and variables; therefore, we include only the years 2002–2013.

countries with observations in fewer than 9 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, of which around 100,000 are not included in the estimations because of zero or missing weights. We use the following two survey questions to identify individuals who start a business (referred to as "nascent entrepreneurs" in the GEM):

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 have been active in establishing a new business during the last year and own at least part of the 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 3 months. However, we believe that this might lead to the exclusion of too many 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 5 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 two-digit sector and country. A total of 34% of all startups are classified in this category.[11] All remaining startups are classified as

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

[10] Approximately 27% of nascent entrepreneurs in our sample report having paid salaries or wages for more than 3 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.

having low growth potential (see Figure A2 and Table F1 in Internet Appendix F for details).[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 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 4.4 we show that 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 4.1, we verify that the correlation between the probability to be a high-growth startup and the ex post growth of firms is positive and significant, even after controlling for sector-specific year fixed effects.

Finally, as an additional robustness check, in Internet Appendix G.2, we use the survey questions from the GEM that identify entrepreneurs who plan to offer a product or service that is considered new by the potential customers and/or that embodies new technologies. 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.

3.2 Business cycle and financial crisis data

In our empirical analysis, we use data on GDP per capita from the Penn World Tables and calculate real GDP growth rates (details are in Appendix A.2). Furthermore, we use data on short-term and long-term risk-free interest rates from the OECD. Short-term rates are 3-month nominal interest rates. Long-term rates are 10-year government bond yields. We subtract the inflation rates from both series to obtain the real rates.

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 crisis dummy, which is based on systemic banking crises data from Laeven and Valencia (2013). According to their measure, 14 countries

[12] The survey provides information on the size of the initial startup investment. Figure F1 in the Internet Appendix reports this information. 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: funds are often borrowed by the entrepreneur rather than by the firm (see the literature review in Section 1). Therefore, this information is not very useful for distinguishing between the entrepreneur's own savings and external financing.

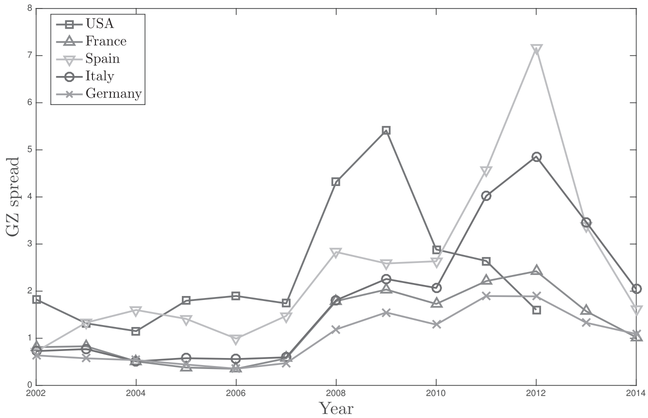

Figure 3
GZ spread by country
The figure plots the interest spread between the bonds of financial institutions and the risk-free rate based on Gilchrist and Zakrajsek (2012).

in our sample suffered a financial crisis, lasting from 2007 to 2013 in the United States and the United Kingdom 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 indicator for the United States, Italy, France, and Germany (details are in Appendix A.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 3 depicts the evolution of the measure by country over the sample period. The measure spikes in 2009 in the United States and in 2012 in Spain and Italy, whereas it is only moderately elevated between 2008 and 2013 in France and Germany. Third, in Internet Appendix G.1, we provide additional results considering the financial distress indicator of Romer and Romer (2017) as an alternative measure.

3.3 Firm-level data set from SABI

The GEM data set 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 database, which contains the number of employees for nearly the entire universe of firms that were established in 2003 or later until the year 2018.

4. Empirical Analysis

4.1 Firm-level analysis

In this section, we analyze the firm dynamics in the Spanish SABI data set to verify that 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 for this purpose. 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 data set are also confirmed when considering only the Spanish surveys.

We cannot link the GEM and SABI data sets at the firm level, but we can do so at the industry level. Using the GEM data, we compute the variable $Share\_growth_{s,t}$, that is, the share of *high-growth* startups in a two-digit sector s in year t in Spain. This is computed for a total of 46 two-digit sectors listed in Table F1 in the Internet Appendix. 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.[13] 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 started 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). To reduce the noise in the variable $Share\_growth_{s,t}$, we also drop firms created in years and sectors in which the share of high-growth firms is based on fewer than five observations. This reduces the number of sectors from 46 to 37. The variable $Share\_growth_{s,t}$ is available for 185 sector-year observations and we perform the analysis on 898,797 firm-year observations in SABI. We estimate the following model:

13 Because the Spain sample of the GEM has very few observations in 2002 and 2003 (approximately 150 startups), we drop these years and only consider firms created from 2004 onward.

$$Employment_{i,s,t} = \beta_0 + \sum_{k=0}^{10} \beta_1^k age_{i,s,t}^k + \sum_{k=0}^{10} \beta_2^k age_{i,s,t}^k \cdot Share\_growth_{i,s} + \sum_{k=0}^{N} \gamma_k X_{i,s,t}^k + \varepsilon_{i,s,t}. \quad (8)$$

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 started interacted with the age dummies. We estimate the model with OLS, and we cluster standard errors at the sector-year level. 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 effect of the probability of being a high-growth startup is identified using both cross-sectional variation (in a given year, two firms of the same age but in different sectors will have different high-growth shares) and time-series variation (in a given sector, two firms of the same age but observed in different years will have different high-growth shares).

Table 1 reports the regression results (β_2 coefficients only; the full set of β_1 coefficients is shown in Table E2 in Internet Appendix E.2). In columns 1–3, we estimate the age profile of firm employment growth. In column 1, we show our baseline specification with year and sector fixed effects. In column 2, we saturate the model with year-sector effects, which absorb all factors that are common to all firms in each sector and year. In column 3, we include an additional control variable, namely, the growth rate of GDP in the year the firm was started interacted with age. As explained above, this controls for the possibility that the share of high-growth startups does not measure the nature of the new businesses but rather the expectations related to the economy when the firm was started. The results are similar across the 3 specifications. Consistent with the model, employment growth is initially lower for likely *high-growth* firms than for other firms, but it becomes progressively higher as firms age, and the difference is positive and statistically significant from age 4 in column 1 and from age 6 in columns 2 and 3.

In columns 4–6, we repeat the analysis of columns 1–3 with the employment level as dependent variable. In the baseline specification in column 4, the log employment level shows dynamics consistent with the findings in columns 1–3. For new firms, a larger share of high-growth startups is related to a smaller initial size, but also to faster growth and to a significantly larger size from age six onward. In the more saturated models in columns 5 and 6, the

Table 1
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				−1.547***	−0.459	−0.509
				(0.3256)	(0.4450)	(0.4607)
Age 1 x share	−0.535***	−0.332**	−0.344**	−2.450***	−1.044***	−0.978***
	(0.1350)	(0.1364)	(0.1378)	(0.5731)	(0.2772)	(0.2953)
Age 2 x share	−0.070**	0.073	0.087	−2.429***	−0.829**	−0.773**
	(0.0341)	(0.0687)	(0.0590)	(0.6134)	(0.3230)	(0.3286)
Age 3 x share	0.034	0.077	0.083	−1.556***	−0.465	−0.433
	(0.0393)	(0.0585)	(0.0587)	(0.5523)	(0.3373)	(0.3303)
Age 4 x share	0.079**	0.020	0.021	−0.585	−0.285	−0.279
	(0.0366)	(0.0289)	(0.0321)	(0.4068)	(0.3365)	(0.3630)
Age 5 x share	0.056	0.014	0.022	−0.047	−0.058	−0.097
	(0.0366)	(0.0295)	(0.0273)	(0.3411)	(0.3116)	(0.3376)
Age 6 x share	0.060**	0.047**	0.044*	0.686**	0.222	0.189
	(0.0271)	(0.0214)	(0.0227)	(0.3425)	(0.2867)	(0.3009)
Age 7 x share	0.073**	0.041*	0.041	1.174***	0.420	0.442
	(0.0344)	(0.0247)	(0.0269)	(0.3693)	(0.2878)	(0.2816)
Age 8 x share	0.068**	0.057***	0.056***	1.748***	0.827**	0.802**
	(0.0311)	(0.0193)	(0.0190)	(0.4331)	(0.3342)	(0.3411)
Age 9 x share	0.009	0.078***	0.077***	1.819***	0.943**	0.960***
	(0.0349)	(0.0182)	(0.0183)	(0.5074)	(0.3666)	(0.3503)
Age 10 x share	−0.054	0.055*	0.058**	1.742***	0.946*	0.841*
	(0.0329)	(0.0296)	(0.0289)	(0.6060)	(0.5077)	(0.4998)
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	671,652	671,652	671,652	898,797	898,797	898,797
R-squared	.112	.115	.115	.149	.150	.150

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 log(employment level). *share* is the share of high-growth startups, as measured from the GEM data, that are in the two-digit sector to which the firm belongs in the year it was started. Firms created in sector-years with fewer than five startup observations in the GEM are dropped. Standard errors are clustered at the sector-year level. *p <.1; **p <.05; ***p <.01.

estimated coefficients decrease slightly in magnitude. Nonetheless, the results are confirmed, particularly that likely high-growth firms are initially smaller but become significantly larger over time.

Based on the specification in column 4, Figure 4 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 started. We show the paths for the 10th and 90th percentiles of the share, which are 18% and 66%, respectively. Firms started when the share of high-growth startups was lower are predicted to be larger initially. However, after 2 years, firms started when the share was higher are predicted to grow faster and to eventually overtake the other firms after 6 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, this selection effect would imply a different interpretation of 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, in Internet Appendix E.1 we examine this possibility and find that the likelihood of being a high-growth

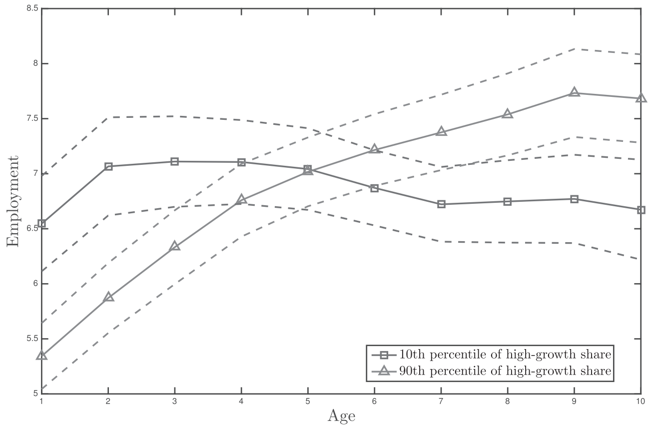

Figure 4
Predicted employment by age from SABI
The figure shows the predicted firm employment based on column 4 of Table 1 for different values of *share*. The 10th percentile is 18%, and the 90th percentile is 66%. The dashed lines represent 90% confidence intervals.

firm slightly increases the exit risk only in the first year of existence, while it reduces this risk for firms between 2 and 8 years old. Thus, high-growth firms seem, if anything, less risky than low-growth firms, and selection effects should not play an important role in the previous results.

Overall, a clear and statistically significant pattern is obvious in the data: firms that are 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/long term than do firms that are more likely to have been derived from low-growth startups.[14] This pattern does not seem to be driven by selection effects, by sector-year effects, and by market conditions prevailing at the time of their birth. These findings are therefore consistent with the behavior of type 2 firms in the model, as well as with our claim that the high-growth startup indicator constructed in the GEM data set is a valid measure of the intrinsic growth potential of these new firms and does not just capture market-related factors.

4.2 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

14 These regressions do not control for differences in sector-specific age profiles of employment. In Internet Appendix E.3, we show that most sectors have similar age profiles of employment at the firm level, and we verify that the results are confirmed after excluding the few sectors with markedly different age profiles. Furthermore, we also perform other robustness checks: we exclude firms started during the Great Recession (2008 and 2009); we aggregate similar sectors to increase the number of startup observations per sector in the GEM, and we restrict the sample to firms created in sector-years with at least 10 startup observations. In each case, we obtain similar results to those presented here. These additional results are available on request.

the following probit model:

$$Pr(start_{i,j,t}=1|X_{i,j,t})=\Phi\left(\beta_0+\beta_1 bus_{j,t}+\beta_2 fin_{j,t}+\sum_{k=0}^{N}\gamma_k X^k_{i,j,t}+\varepsilon_{i,j,t}\right), \tag{9}$$

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 section 3.2. $X^k_{i,j,t}$ is a vector of N control variables including country dummies, gender, age and educational level.[15] We weight observations by using the weight variable for the 18–64 labor force included in the GEM.[16]

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, to test Prediction 3, we estimate a model that includes the interaction $bus_{j,t}\cdot fin_{j,t}$:

$$Pr(start_{i,j,t}=1|X_{i,j,t})=$$
$$\Phi\left(\beta_0+\beta_1 bus_{j,t}+\beta_2 fin_{j,t}+\beta_3 bus_{j,t}\cdot fin_{j,t}+\sum_{k=0}^{N}\gamma_k X^k_{i,j,t}+\varepsilon_{i,j,t}\right). \tag{10}$$

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.

This estimation strategy requires that cyclical fluctuations and financing conditions are not perfectly correlated in the data, and we find that this is the

15 In Section 4.4, we present the results including dummies for the household income level (three categories). The information on the actual income level of respondents' households is not available in the GEM data. Instead, the GEM contains a variable that indicates whether a person's household in a specific year and country is in the lowest, the middle or the upper tercile of the income distribution of the households in the respondent's country. 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.

16 In accordance with the description of the GEM, the weights are "developed such that proportions of different subgroups (gender and age, e.g.) match the most recent official data descriptions of the population of a country." Our results are robust to not weighting the observations.

Table 2
Financial crisis, GZ spread, and an entrepreneur's 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	894,126	894,126	894,126	370,280	370,280	370,280
R-squared	.060	.045	.075	.037	.035	.034
p-value for $\beta_2^{low}=\beta_2^{high}$		.018			.002	

The dependent variable is a dummy equal to one if an individual is a nascent entrepreneur in the respective category. Controls include dummies for three education levels, sex, age, and country fixed effects. Standard errors are clustered at the country level. $*p<.1$; $**p<.05$; $***p<.01$.

case in our sample. The correlation between the GZ spread and GDP growth is −0.39, thus low enough that their effects can be separately identified. This is shown in detail in Appendix A.4, where we report the scatterplots between GDP growth (deviations from country averages) and the value of the GZ spread. The plot shows 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.

4.3 Individual-level analysis: Baseline results

Table 2 shows the results of the baseline probit model (9). 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, whereas the financial crisis indicator has a significantly 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 the 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 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 United States, 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. For both indicators, the hypothesis that their coefficients are the same can be rejected as shown by the *p*-values reported at the bottom of the table.[17] Table G3 in the Internet Appendix shows the full set of coefficients of the demographic controls, which indicate that startup creation is positively related to income and negatively related to being female or being older.

Overall, the results in Table 2 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% (vs. 4.5% in the complementary group).[18] The result that for low-growth startups both the Financial Crisis coefficient (column 2) and the GZ spread coefficient (column 4) are not significantly different from zero is consistent with the ambiguous relationship between r^b and type 1 startups predicted by the model. As shown in the left panel of Figure 1, an increase in the spread r^b has both a positive effect (some entrepreneurs switch from type 2 to type 1 startups) and a negative effect (some entrepreneurs switch from starting a type 1 business to not starting a business) on type 1 startups. While the negative effect dominates overall in Figure 1, in Internet Appendix C we show that alternative assumptions on the distribution of entrepreneurial skills imply that the positive effect might dominate instead (see Figure C3).

Table 3 presents the results of estimating model (10) 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 significant than in Table 2. Finally, the interaction between *GDP growth* and *Financial crisis* is positive and statistically significant. In general, a positive interaction coefficient indicates 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 also can be interpreted as showing a

[17] We cannot include an interaction between the financial frictions indicators and a high-growth indicator in the regression model to test for a differential effect by startup type as a value of zero for this indicator would perfectly predict an outcome of no startup creation. Thus, to test the equality of the coefficients, we estimate the simultaneous covariance matrix of the two equations by generating a joint sample consisting of two appended copies of the data set, in which the low-growth startup indicator is the dependent variable in the first copy and the high-growth startup indicator is the dependent variable in the second copy. We then run a regression including an interaction of all regressors with a dummy D equal to one if the observation is from the second copy and equal to zero otherwise. The noninteracted coefficients (both their value and their standard error) are identical to the coefficients estimated in the original regression with the low-growth indicator as the dependent variable. Moreover, the interacted coefficients represent the deviations between coefficients in the low-growth and high-growth regressions, allowing us to perform the test. This approach is identical to the Hausman-type test implemented by the *suest* command in Stata (for further details, see Weesie 2000).

[18] We compute these semielasticities 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. The inherent nonlinearity of the probit model has no considerable impact on the marginal effects as we find them to be not sensitive to the particular values at which the controls or GDP growth are evaluated. For example, the marginal effect of the GZ spread in column 6 in Table 2 is -11.03% with GDP growth evaluated at -0.03 and -11.6% with GDP growth evaluated at +0.03.

Table 3
Financial crisis, GZ spread, and an entrepreneur's 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	894,126	894,126	894,126	370,280	370,280	370,280
R-squared	.062	.046	.077	.039	.035	.039
p-value for $\beta_2^{low}=\beta_2^{high}$		.129			.030	
p-value for $\beta_3^{low}=\beta_3^{high}$		.484			.000	

The dependent variable is a dummy 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. Standard errors are clustered at the country level. *p <.1; **p <.05; ***p <.01.

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. However, these differences are not statistically significant.

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 strong 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* × *GDP growth* is large and statistically significant for the startups with high growth potential, and both coefficients are significantly different from the coefficients for low-growth startups. 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, Figure 5 compares the frequencies of high-growth and low-growth startups implied by the model with those estimated in the data. For the left-hand-side model-based graphs, we

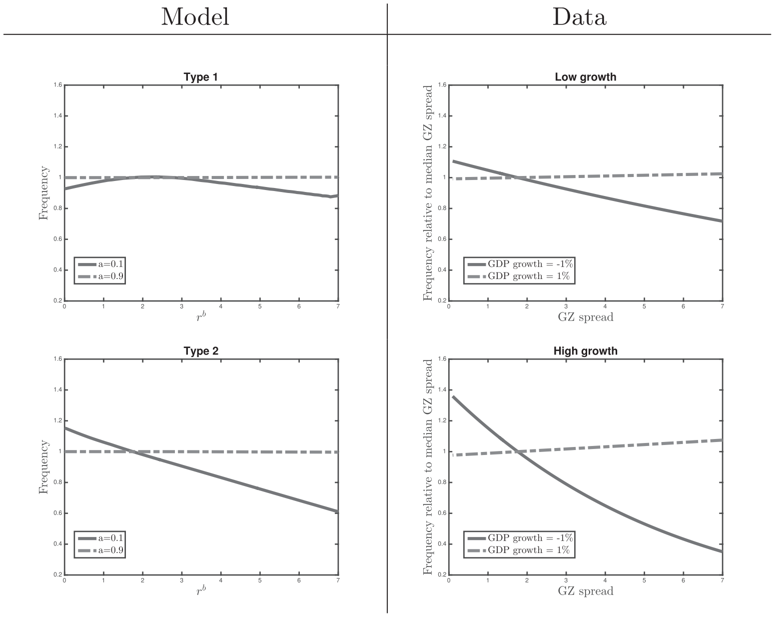

Figure 5
Comparison of model and empirical predictions

choose two values of the entrepreneur's initial endowment to represent the case in which 90% of the entire initial sunk cost κ has to be financed externally ($a = 0.1$) and the case in which only 10% has to be financed externally ($a = 0.9$). The other parameters are those defined for the benchmark calibration described in Section 2.3. 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.

The right-hand-side graphs use the estimated coefficients in Table 3 to compute the marginal effects of the *GZ spread*, conditional on a given value of GDP growth. Here the solid and dashed lines depict the cases of GDP growth rates equal to -1% and 1%, respectively. Therefore, consistent with our model, we match low/high GDP growth periods in the data with low/high values of wealth a in the model. As mentioned before, this mapping is supported by the fact that the average share of external finance needed to start a business across all entrepreneurs in the GEM negatively correlates with GDP growth (see Table F2 in the Internet Appendix).

Qualitatively, Figure 5 shows that 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. Both startup types are more negatively affected by financial frictions during periods of declining

GDP, and the effect is stronger for high-growth than for low-growth startups. Quantitatively, the effect of the interest rate spread on the relative frequencies is stronger in the data than in the model. This might be because in the model we assume a highly dispersed distribution of entrepreneurial skills. In Internet Appendix C, we show that assuming a normal distribution of skills magnifies the effect of the interest rate spread on these relative frequencies.

4.4 Additional control variables

In this section, we confirm that our results are robust to the inclusion of additional control variables. In particular, we add the share of firm exits for each sector/country/year observation, household income categories (for the definition, see footnote 15), and a dummy variable indicating expertise in running a business. The first of these controls captures the possibility that new startups are driven by the presence of serial entrepreneurs who seek to start a new business.

A potential threat to the consistency of the effect of the GZ spread on the high-growth startups relative to low-growth startups is that positive economic shocks might affect financial conditions but also change agents' expectations about the future state of the economy. To control for this factor, we include two further controls. First, we add a variable indicating that a respondent in the GEM expects good business opportunities in the future.[19] Second, we include the riskless short-term interest rate.

Table 4 presents the results based on this specification. We find that a higher household income tends to increase the probability to start a new firm. The effects are larger and more significant for the high-growth category. This finding is consistent with the notion that financial frictions matter more for high-growth startups, given that a higher household income potentially enables an entrepreneur to self-finance a larger part of the initial investment in the new business. The share of exits is not statistically significant, whereas having business expertise has a significantly positive effect. As anticipated, the expectation of good business opportunities has a positive and very strong effect on startup creation. Moreover, the riskless rate is positively related with the probability of starting a new firm, most likely because it is a leading empirical indicator of the business cycle.

In the results shown earlier, we make the GZ spread comparable across countries by using the German bund as the risk-free benchmark rate (details in Appendix Section A.4). As a spread between the rate of the bund and that of U.S. Treasuries is likely to not be entirely due to financial frictions,

[19] 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 who answer *Don't know* and include in the analysis the variable *Opportunity expectations*, which is equal to one for *Yes* and zero otherwise. Although the time horizon of this expectations variable is relatively short, we 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 4
Including additional control variables

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	0.406	0.744	−0.416	4.700***	3.639***	5.549***
	(0.5279)	(0.4823)	(0.5646)	(1.6664)	(1.3401)	(1.8138)
Fin. crisis	−0.179***	−0.134***	−0.220***			
	(0.0633)	(0.0514)	(0.0748)			
Fin. crisis x GDP growth	4.360*	3.461*	5.030			
	(2.4393)	(1.8463)	(3.3071)			
GZ spread				−0.016	−0.007	−0.040
				(0.0216)	(0.0175)	(0.0267)
GZ spread x GDP growth				2.175	1.147	3.808**
				(1.7186)	(1.3518)	(1.9054)
Middle income	0.060	0.041	0.082	0.111***	0.075**	0.151***
	(0.0426)	(0.0322)	(0.0516)	(0.0372)	(0.0300)	(0.0401)
High income	0.013	−0.035	0.100***	0.038**	0.001	0.106***
	(0.0228)	(0.0294)	(0.0152)	(0.0152)	(0.0168)	(0.0118)
Share of exits	−1.411	−1.352	−0.850	−0.408	0.025	−0.317
	(5.4840)	(4.9684)	(5.0735)	(4.9201)	(5.7236)	(2.6569)
Business expertise	0.837***	0.775***	0.754***	0.886***	0.813***	0.815***
	(0.0344)	(0.0284)	(0.0445)	(0.0105)	(0.0140)	(0.0285)
Opportunity expectations	0.385***	0.347***	0.341***	0.386***	0.349***	0.334***
	(0.0342)	(0.0314)	(0.0353)	(0.0588)	(0.0540)	(0.0522)
Riskless interest rate	0.038***	0.034***	0.036***	0.056	0.055*	0.036
	(0.0084)	(0.0086)	(0.0091)	(0.0360)	(0.0323)	(0.0317)
Observations	639,248	639,248	639,248	316,450	316,450	316,450
R-squared	.135	.121	.119	.142	.127	.125
p-value for $\beta_2^{low}=\beta_2^{high}$		.022			.022	
p-value for $\beta_3^{low}=\beta_3^{high}$		.447			.000	

The dependent variable is a dummy equal to one if an individual is a nascent entrepreneur in the respective category. Controls include dummies for three education levels, sex, age, and country fixed effects. Standard errors are clustered at the country level. *p <.1; **p <.05; ***p <.01.

pooling the U.S. and European countries together might introduce some noise in our measurement of financial frictions. Therefore, in Table 5, we repeat the regressions in Table 4 after excluding the United States in order to reduce this noise. We find that, if anything, the results are stronger and more significant in Table 5, with the coefficient of the interaction between the GZ spread and GDP growth increasing from 3.8 to 5.5.

Altogether, the inclusion of the additional control variables does not significantly change the results obtained previously, except that the estimates of the coefficients of the interaction between the financial crisis dummy and GDP growth are somewhat more noisy.

4.5 Quantifying employment losses

In the following, we combine the results obtained in the previous section with the analysis of SABI from Section 4.1 to assess the impact of a severe recession on overall employment (abstracting from the general equilibrium effects on wages and interest rates) through the *composition of entry* channel. To do so,

Table 5
Including additional control variables (without the United States)

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	0.129	0.467	−0.617	5.764***	4.273***	6.912***
	(0.5676)	(0.5006)	(0.6236)	(1.6970)	(1.5718)	(1.2499)
Fin. crisis	−0.207***	−0.161***	−0.241***			
	(0.0665)	(0.0531)	(0.0800)			
Fin. crisis x GDP growth	4.990**	4.107**	5.455			
	(2.4255)	(1.7789)	(3.3869)			
GZ spread				−0.040**	−0.025*	−0.068***
				(0.0157)	(0.0139)	(0.0158)
GZ spread x GDP growth				3.618***	2.159*	5.463***
				(1.2286)	(1.1884)	(0.8323)
Middle income	0.055	0.032	0.085	0.115***	0.070*	0.171***
	(0.0462)	(0.0355)	(0.0527)	(0.0437)	(0.0387)	(0.0316)
High income	0.014	−0.035	0.100***	0.051***	0.009	0.120***
	(0.0247)	(0.0328)	(0.0166)	(0.0114)	(0.0163)	(0.0096)
Share of exits	−3.935	−4.089	−2.259	−0.530	−2.011	2.607
	(5.4943)	(4.8545)	(5.3240)	(4.5500)	(4.7942)	(3.0230)
Business expertise	0.835***	0.773***	0.753***	0.891***	0.818***	0.820***
	(0.0384)	(0.0318)	(0.0480)	(0.0108)	(0.0145)	(0.0302)
Opportunity expectations	0.368***	0.329***	0.333***	0.344***	0.308***	0.307***
	(0.0321)	(0.0276)	(0.0368)	(0.0290)	(0.0236)	(0.0373)
Riskless interest rate	0.033***	0.029***	0.033***	0.033	0.036	0.013
	(0.0078)	(0.0076)	(0.0099)	(0.0359)	(0.0360)	(0.0240)
Observations	609,605	609,605	609,605	289,749	289,749	289,749
R-squared	.129	.112	.118	.134	.115	.126
p-value for $\beta_2^{low}=\beta_2^{high}$		.05			.000	
p-value for $\beta_3^{low}=\beta_3^{high}$		.535			.000	

The dependent variable is a dummy equal to one if an individual is a nascent entrepreneur in the respective category. Controls include dummies for three education levels, sex, age, and country fixed effects. Standard errors are clustered at the country level. *p <.1; **p <.05; ***p <.01.

we perform a simple back-of-the-envelope calculation to quantify the effects of an increase in financial frictions and a worsening of business cycle conditions on the size of new firms emerging via the decline in the share of high-growth startups created.

First, we compute the predicted changes in the probabilities of creating a startup using the regression models in columns 5 and 6 in Table 5.[20] We then use column 4 of Table 1 and, in particular, the coefficient of the *Age 10 x share* interaction to calculate the implied effect on employment per firm 10 years after firm creation. We focus on this long-run effect because, consistently with our model, our estimates in Table 1 indicate that the *composition of entry* channel reduces employment growth from around 3 to 4 years of age onward. We also

[20] Specifically, we predict the probabilities of creating a low-growth or a high-growth startup for different values of GDP growth and the GZ spread, evaluating the control variables at their means, and calculate the high-growth share as Pr(high-growth=1)/(Pr(low-growth=1)+Pr(high-growth=1)). The results in this section are roughly similar, if we instead use the estimates from Table 3, the estimates including the United States from Table 4 or the estimates obtained when using the GEM sample of Spain only.

compare this channel with an estimate of the employment losses caused by the drop in the number of entering firms, which we call the *extensive margin*. For a more accurate comparison, we should compute how many fewer jobs are created in the long term because of the decrease in firm entry. This step is practically unfeasible with our data, but it is plausible to assume that the long-term losses from the extensive margin are smaller than the short-term ones, because some entrepreneurs might simply delay their entry, meaning that the reduction in firm creation due to unfavorable financial conditions does not one-to-one translate into a lower number of firms many years ahead. Therefore, we compute the short-term losses from this margin, and we assume they are an upper bound of the long-term ones.

We consider the hypothetical scenario of a recession that reduces GDP growth from its long-run average of 1.8% to -3%.[21] Using the estimates from Table 5 and the SABI data, we obtain that the GDP contraction alone, with the GZ spread constant at its average value, decreases the probability of creating any startup from 1.5% to 0.74%. We then compute the *composition of entry* margin. Before the recession, the high-growth share is equal to 31.7%. When GDP growth drops to -3%, this share falls to 21.7%. Combining these changes in the composition of startups with the estimates from the SABI data in Table 1, we obtain that firms created in the recession have a 2.4% lower average employment level after 10 years. In Internet Appendix E.4, we estimate that these job losses correspond to around 8% of our upper bound estimate of the extensive-margin job losses.

Consider now that in addition to the drop in GDP, there is also a one-percentage-point increase in the GZ spread. Such tightening of financial conditions has a small negative effect on the extensive margin, further reducing the probability of creating a startup from 0.74% to 0.51%. Moreover, it reduces the high-growth share by an additional 7.5 percentage points, from 21.7% to 14.2%.[22] This decreases the future employment level after 10 years by an additional 1.9%, which corresponds to approximately 22% of the additional job losses from our upper bound estimate of the extensive margin (for details, see Internet Appendix E.4). This value is quantitatively significant, especially considering that during financial crises the GZ spread typically increases by several percentage points. We recognize that there might be other channels operating in the opposite direction, and that a more detailed analysis would be necessary to precisely quantify the employment effects of the *composition of entry* channel. Nevertheless, our estimates suggest that these effects are an important determinant of the medium- and long-term job losses caused by financial shocks.

---

21 This negative growth rate roughly corresponds to that of Spain's during the Great Recession.

22 Note that this relatively large additional effect of financial frictions is caused by the interaction effect between GDP growth and the GZ spread. If the same tightening of the spread happened when GDP growth was at its long-run average, the high-growth share would remain approximately unchanged.

5. Robustness Checks

In this section, we complement our analysis with a number of robustness checks. We use industry-level measures of financial frictions; we control for the term premium of the interest rate, and we present the results of a two-stage Heckman selection model. Finally, we summarize the results of additional robustness checks included in the Internet Appendix.

5.1 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 more external financing on average 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 be easily 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, Laeven, and Klingebiel (2007), and we identify the manufacturing startups with low- and high-external financial dependence (*low EFD* and *high EFD*). Appendix A.5 offers the details.

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 Compustat SIC codes with two-digit sectors in the GEM data set, 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.[23]

[23] 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 results shown in Table 3 also hold in this subsample.

Table 6
Baseline results by intangibility and external financial dependence of sectors

	By intangibility				By external financial dependence			
	(1) Low	(2) High	(3) Low	(4) High	(5) Low	(6) High	(7) Low	(8) High
GDP growth	0.535	0.617	4.082***	7.337**	1.650**	−0.973	3.856*	7.181***
	(0.5267)	(0.6626)	(1.5757)	(3.2281)	(0.8264)	(1.7349)	(2.1633)	(1.7670)
Fin. crisis	−0.108***	−0.252***			−0.089**	−0.240**		
	(0.0403)	(0.0858)			(0.0385)	(0.0941)		
Fin. crisis x GDP growth	3.155**	6.663**			2.672*	7.688**		
	(1.5244)	(2.7952)			(1.5425)	(3.3112)		
GZ spread			−0.017	−0.026			−0.016	0.025
			(0.0155)	(0.0362)			(0.0194)	(0.0275)
GZ spread x GDP growth			2.067*	3.415			0.406	2.693***
			(1.1846)	(2.6825)			(1.8754)	(0.7913)
Observations	894,126	894,126	370,280	370,280	894,126	894,126	370,280	370,280
R-squared	.057	.063	.028	.053	.047	.067	.032	.032
p-value for $\beta_2^{low}=\beta_2^{high}$	.010		.674		.070		.047	
p-value for $\beta_3^{low}=\beta_3^{high}$	.036		.375		.064		.100	

The dependent variable is a dummy equal to one if an individual is a nascent entrepreneur in the respective category. Controls include dummies for three education levels, sex, age, and country fixed effects. Standard errors are clustered at the country level. *p <.1; **p <.05; ***p <.01.

Columns 1–4 of Table 6 report the regression results when we select startups in sectors with low or high intangibility. Columns 5–8 repeat the same exercise using the low and high EFD indicators. When using the financial crisis as an indicator of a high cost of external finance, we find that startups in more financially constrained sectors (high intangibility or high EFD) decline significantly more on average than in the other sectors, and their interaction term is also significantly higher, confirming our predictions. When using the GZ spread, we find results that go in the same direction, although they are generally more noisy and in some cases not significant. Importantly, we find that high-growth startups are always more negatively affected by financial frictions than low-growth startups, with differences that are statistically significant in all specifications, except those in columns 3 and 4.

5.2 Controlling for the term premium in interest rates

In the model, we normalize the real interest rate to zero in both short and long runs. However, in reality, changes in the term structure of interest rates should affect high- and low-growth startups differently, given their different intertemporal profiles. In particular, a higher term premium should more negatively affect high-growth startups because their current value depends more on future than current profits.

Therefore, in Table 7, we again estimate the models shown in Table 3 while also controlling for the term premium, that is, the difference between long-term and short-term interest rates. One caveat is that long-term rates are measured by 10-year government bonds and might partly capture the excess cost of finance measured by the GZ spread in countries affected by sovereign crises. Nonetheless, we find that the main results remain qualitatively unchanged, in the sense that the GZ spread penalizes significantly more high-growth startups than

Table 7
Including the term premium

	(1) All	(2) Low growth	(3) High growth	(4) All	(5) Low growth	(6) High growth
GDP growth	0.626	0.920**	−0.189	2.282***	1.994***	2.208***
	(0.4129)	(0.4347)	(0.3993)	(0.6356)	(0.7458)	(0.3302)
Fin. crisis	−0.112***	−0.089***	−0.134***			
	(0.0307)	(0.0301)	(0.0353)			
Fin. crisis x GDP growth	3.154***	2.695**	3.442***			
	(1.1167)	(1.1238)	(1.3019)			
GZ spread				0.014	0.014	0.003
				(0.0232)	(0.0223)	(0.0169)
GZ spread x GDP growth				0.748***	0.315	1.392***
				(0.2003)	(0.2888)	(0.2625)
Term premium	−0.044*	−0.034**	−0.054*	−0.100***	−0.081***	−0.108***
	(0.0225)	(0.0170)	(0.0281)	(0.0101)	(0.0060)	(0.0152)
Observations	816,895	816,895	816,895	370,280	370,280	370,280
R-squared	.045	.040	.045	.046	.040	.047
p-value for $\beta_2^{low}=\beta_2^{high}$		.142			.056	
p-value for $\beta_3^{low}=\beta_3^{high}$		.560			.017	

The dependent variable is a dummy equal to one if an individual is a nascent entrepreneur in the respective category. Controls include dummies for three education levels, sex, age, and country fixed effects. Standard errors are clustered at the country level. *p <.1; **p <.05; ***p <.01.

Table 8
Heckman selection model with sector value-added growth in the second stage

	(1)	(2)	(3)	(4)
GDP growth	−0.363	−0.810*	0.041	2.778***
	(0.3801)	(0.4475)	(0.7043)	(1.0426)
Fin. crisis	−0.039	−0.073**		
	(0.0278)	(0.0336)		
Fin. crisis x GDP growth		1.618*		
		(0.8838)		
GZ spread			−0.047**	−0.037*
			(0.0203)	(0.0205)
GZ spread x GDP growth				2.751***
				(0.7496)
Sector output growth rate	−0.246	−0.267*	−0.686**	−0.714**
	(0.1531)	(0.1534)	(0.3328)	(0.3326)
Observations	889,080	889,080	369,874	369,874

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. *p <.1; **p <.05; ***p <.01.

low-growth startups. Moreover, we find that the term premium has a negative effect, which is stronger for high-growth startups, consistent with our model.

5.3 Heckman selection model

In Table 8, we estimate a two-step Heckman selection model. This has the advantage of allowing us to directly test whether financial conditions and the business cycle affect the choice of creating a high-growth startup *instead* of a low-growth one, while correcting for the potential bias due to the selection into being an entrepreneur.

In the first-stage selection equation, we estimate the probability of starting a business and in addition to GDP growth, we include the indicator for financial frictions and their interaction as well the controls for sex, education, age, and country fixed effects as explanatory variables. In the second-stage equation, we estimate the effects of GDP growth and financial frictions on the type of business created. This 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.

Another advantage of this approach is that it allows us to control for sector-level factors. Sector-level output growth is a good proxy for the demand conditions faced by new entrepreneurs, but we cannot include it as a control variable in our benchmark regressions because we only know the sectors of new or continuing entrepreneurs. However, we can include it in the second step of the Heckman selection model. As a proxy for sector-level output, we use the growth rate of gross output at the sector level obtained from the OECD STAN database. Table 8 shows the results for the second stage, which confirm that startups with high growth potential are less frequent during a financial crisis and are significantly more sensitive to financing conditions than low-growth startups. Interestingly, the effect of the output growth rate on high-growth startups is negative and significant, which implies that higher output growth increases the probability of low-growth relative to high-growth startups. This is consistent with the model because short-run demand shocks should be more beneficial to startups that generate higher profits in the short term.

5.4 Additional robustness checks

We perform additional robustness checks of our main results by including additional control variables and by adopting different selection criteria for the sample. These checks are briefly summarized here, while the detailed results can be found in the Internet Appendix.

In Table G4, we show that the negative effect of bond spreads on high-growth startups is confirmed when considering only the Spanish GEM surveys.

In Table G5, 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. In Table G6, we exclude startups that have already paid some wages and thus might have been established before. Both of these robustness checks confirm the previous results.

In Table G7, 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 data set, the financial crisis dummy becomes insignificant. Nonetheless, the main results for the interaction between financial frictions and GDP growth are confirmed.

In Table G8, we include the country-specific riskless interest rate as regressor, not only in isolation but also when interacted with GDP growth. Our main results are again confirmed in this case.

6. 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 data set to identify high-growth startups in the data. For the case of Spain, which has very extensive coverage in the GEM data set, we use firm-level data from SABI to confirm that high-growth startups are more likely to grow faster and employ more workers in the long term than other 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.

Appendix

A. Data and Variable Definitions

A.1 Business Types Identified from GEM Questions

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 subcontractors, 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 size of the established firms by sector (at the two-digit level) and country (averaged across all years) by using the answer to the first question given by respondents that are owners of firms that are 5 or more years old.[24]

[24] 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.

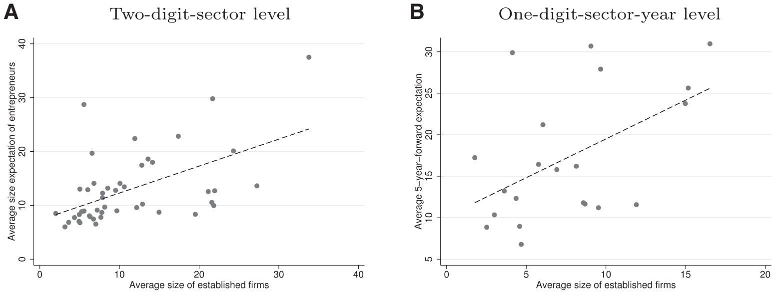

Figure A1
Comparison of actual and expected size of firms in the GEM data (sector averages)
The size of established firms is calculated as the number of employees of firms that are at least 5 years old (proxying for the year of creation by the first year in which the firm paid wages). The size expectation for startups is taken from the answer to question 2, which is described above. To control for outliers, we drop answers above the 99th percentile. In panel A, points indicate the averages of these measures for 46 two-digit sectors across all years. In panel B, points indicate the average size of established firms in year t and the average 5-year-forward firm size expectation of entrepreneurs in year $t-5$ in the same one-digit sector (keep only those averages that are based on more than 10 firm observations). The dashed lines represent the best linear fit.

We then classify a startup as having high growth potential if the answer to the second question, that is, the expected size in 5 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 process would result in very few observations in many country-sectors; therefore, we choose to consider all firms that are at least five years old.[25]

Panel A of Figure A1 plots the two-digit sector averages of the size of these established firms against the expectations of entrepreneurs. It suggests a strong relationship between actual sizes and expectations across sectors. The coefficient of correlation is 0.54. The linear fit is flatter than the 45-degree line, most likely because firms older than 5 years are also included to compute the size of established firms. Panel B shows the relationship between the actual size of established firms and the 5-year-forward expectations of entrepreneurs 5 years earlier at the one-digit-sector-year level. We opt for a lower level of sector aggregation in this case because the number of firm observations at the sector-year level is very low for many two-digit sectors, leading to very noisy averages. Again, we find a strong positive relationship with a similar correlation coefficient of 0.51. Figure A2 shows the distribution of low-growth and high-growth startups for each two-digit sector.

A.2 Business Cycle Data

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

25 We confirm that the main results are not sensitive to using different ranges of the firm age, for example, 5 to 10 years, to compute the average size of established firms.

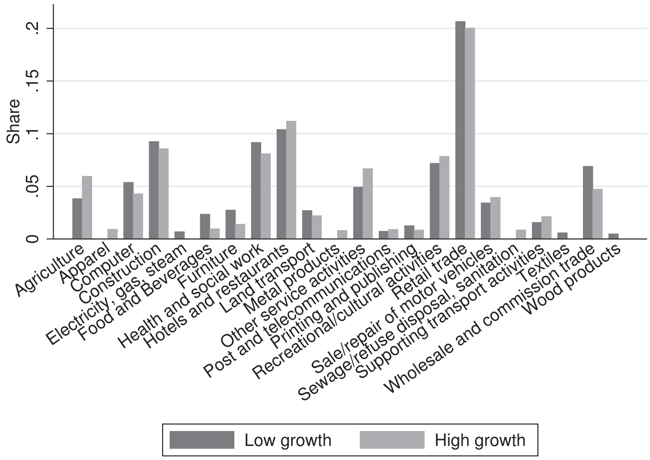

Figure A2
Distribution of low-growth and high-growth startups in two-digit sectors
The figure shows the sector shares of startups in the 21 most common sectors, which account for approximately 94% of all startups, separately for the low-growth and high-growth categories.

A.3 Financial Crisis Data

We identify years in which a particular country is in financial crisis by using data on systemic banking crises from Laeven and Valencia (2013). Table A1 shows the countries in our sample, the corresponding crisis period and the number of observations.

A.4 GZ 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 U.S. government bonds. We make our index comparable across countries by measuring the premiums of all countries with respect to the German bund. For the United States, we take the domestic spread directly from Gilchrist and Zakrajsek (2012)[26], and we add the spread between U.S. and German government bonds.[27] 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.[28] We finally compute the yearly means of the monthly data.

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

[27] Retrieved from *https://fred.stlouisfed.org/series/IRLTLT01USM156N*

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

Table A1
Countries and financial crisis years

Country	Start year	End year	Obs.
Belgium	2008	2013	28304
Chile	-	-	32911
Croatia	-	-	18972
Denmark	2008	2013	27954
Finland	-	-	21049
France	2008	2013	18687
Germany	2008	2013	60618
Greece	2008	2013	20432
Hungary	2008	2013	21979
Iceland	2008	2013	15547
Ireland	2008	2013	19163
Italy	2008	2013	23210
Japan	-	-	21176
Netherlands	2008	2013	30315
Norway	-	-	18506
Slovenia	2008	2013	27879
Spain	2008	2013	232751
Sweden	2008	2013	39648
Switzerland	2008	2013	18510
United Kingdom	2007	2013	157880
United States	2007	2013	38594

The periods for systemic banking crises come from Laeven and Valencia (2013).

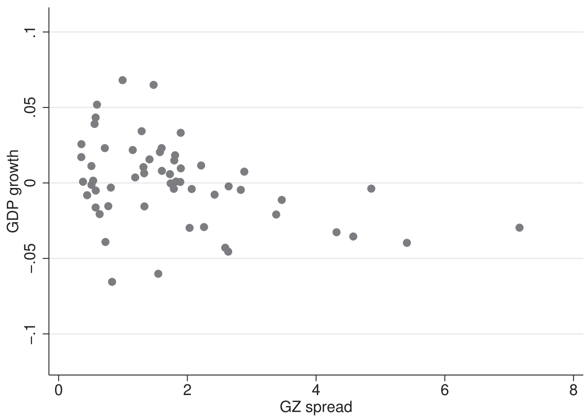

Figure A3
Correlation between GDP growth (deviation from the country average) and bond spread

A.5 External Financial Dependence and Intangibility Data

The GEM data set contains information on the industrial sector in which a business is started. Sectors are classified following ISIC Rev.3 until 2008 and ISIC Rev.4 from 2009 onward. 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, Laeven, and Klingebiel (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 U.S. data and is 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 project's gestation period. 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 U.S. 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, Laeven, and Klingebiel (2007) to the 22 manufacturing sectors identified in the GEM data set. For sectors that we can match across Compustat SIC codes and the two-digit sectors in the GEM data set, 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.

References

Acemoglu, D., U. Akcigit, and M. A. Celik. 2014. Young, restless and creative: Openness to disruption and creative innovations. Working Paper, MIT.

Adelino, M., A. Schoar, and F. Severino. 2015. House prices, collateral, and self-employment. *Journal of Financial Economics* 117:288–306.

Akcigit, U., and W. R. Kerr. 2018. Growth through heterogeneous innovations. *Journal of Political Economy* 126:1374–443.

Almeida, H., and M. Campello. 2007. Financial constraints, asset tangibility, and corporate investment. *Review of Financial Studies* 20:1429–60.

Azoulay, P., B. Jones, J. D. Kim, and J. Miranda. 2020. Age and high-growth entrepreneurship. *American Economic Review: Insights* 2:65–82.

Blanchflower, D. G., and A. J. Oswald. 1998. What makes an entrepreneur? *Journal of Labor Economics* 16:26–60.

Braun, M., and B. Larrain. 2005. Finance and the business cycle: International, inter-industry evidence. *Journal of Finance* 60:1097–128.

Brown, J. D., Earle, J. S., M. J. Kim, and K. M. Lee. 2018. High-growth entrepreneurship. IZA Discussion Papers 11662 Working Paper, Institute of Labor Economics (IZA).

Buera, F. J., J. P. Kaboski, and Y. Shin. 2011. Finance and development: A tale of two sectors. *American Economic Review* 101:1964–2002.

Caggese, A., and V. Cunat. 2013. Financing constraints, firm dynamics, export decisions, and aggregate productivity. *Review of Economic Dynamics* 16:177–93.

Caggese, A., and A. Perez. 2017. Capital misallocation and secular stagnation. Finance and Economics Discussion Series 2017-009, Board of Governors of the Federal Reserve System (U.S.).

Carroll, C. D. 2001. A theory of the consumption function, with and without liquidity constraints. *Journal of Economic Perspectives* 15:23–45.

Cerra, V., and S. C. Saxena. 2008. Growth dynamics: the myth of economic recovery. *American Economic Review* 98:439–57.

Clementi, G. L., and B. Palazzo. 2016. Entry, exit, firm dynamics, and aggregate fluctuations. *American Economic Journal: Macroeconomics* 8:1–41.

Cole, H. L., J. Greenwood, and J. M. Sanchez. 2016. Why doesn't technology flow from rich to poor countries? *Econometrica* 84:1477–521.

Corradin, S., and A. Popov. 2015. House prices, home equity borrowing, and entrepreneurship. *Review of Financial Studies* 28:2399–428.

Dell'Ariccia, G., E. Detragiache, and R. Rajan. 2008. The real effect of banking crises. *Journal of Financial Intermediation* 17:89–112.

Falato, A., D. Kadyrzhanova, and J. W. Sim. 2013. 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.).

Gilchrist, S., and M. Benoit. 2016. Credit risk in the euro area. *Economic Journal* 128:118–58.

Gilchrist, S., and E. Zakrajsek. 2012. Credit spreads and business cycle fluctuations. *American Economic Review* 102:1692–720.

Haltiwanger, J., R. S. Jarmin, R. Kulick, and J. Miranda. 2016. High growth young firms: Contribution to job, output, and productivity growth. In *Measuring entrepreneurial businesses: current knowledge and challenges*, pp. 11–62. Washington, DC: National Bureau of Economic Research.

Holtz-Eakin, D., D. Joulfaian, and H. S. Rosen. 1994. Sticking it out: Entrepreneurial survival and liquidity constraints. *Journal of Political Economy* 102:53–75.

Hombert, J., and A. Matray. 2016. The real effects of lending relationships on innovative firms and inventor mobility. *Review of Financial Studies* 30:2413–45.

Koellinger, P. D., and A. R. Thurik. 2012. Entrepreneurship and the business cycle. *Review of Economics and Statistics* 94:1143–56.

Krishnan, K., D. K. Nandy, and M. Puri. 2014. Does financing spur small business productivity? Evidence from a natural experiment. *Review of Financial Studies* 28:1768–809.

Kroszner, R. S., L. Laeven, and D. Klingebiel. 2007. Banking crises, financial dependence, and growth. *Journal of Financial Economics* 84:187–228.

Laeven, L., and F. Valencia. 2013. The real effects of financial sector interventions during crises. *Journal of Money, Credit and Banking* 45:147–77.

Midrigan, V., and D. Y. Xu. 2014. Finance and misallocation: Evidence from plant-level data. *American Economic Review* 104:422–58.

Poschke, M. 2013. Who becomes an entrepreneur? Labor market prospects and occupational choice. *Journal of Economic Dynamics and Control* 37:693–710.

Poschke, M. 2018. The firm size distribution across countries and skill-biased change in entrepreneurial technology. *American Economic Journal: Macroeconomics* 10:1–41.

Pugsley, B., P. Sedlacek, and V. Sterk. 2018. The nature of firm growth. Working Paper, University of Notre Dame.

Pugsley, B. W., and A. Sahin. 2018. Grown-up business cycles. *Review of Financial Studies* 32:1102–47.

Rajan, R. G., and L. Zingales. 1998. Financial dependence and growth. *American Economic Review* 88:559–86.

Reynolds, P. D., and D. M. Hechavarria. 2016. Global entrepreneurship monitor [gem]: Adult population survey data set, 1998–2013. Report.

Robb, A. M., and D. T. Robinson. 2012. The capital structure decisions of new firms. *Review of Financial Studies* 27:153–79.

Romer, C. D., and D. H. Romer. 2017. New evidence on the aftermath of financial crises in advanced countries. *American Economic Review* 107:3072–118.

Schmalz, M. C., D. A. Sraer, and D. Thesmar. 2017. Housing collateral and entrepreneurship. *Journal of Finance* 72:99–132.

Schoar, A. 2010. The divide between subsistence and transformational entrepreneurship. *Innovation Policy and the Economy* 10:57–81.

Sedlacek, P., and V. Sterk. 2017. The growth potential of startups over the business cycle. *American Economic Review* 107:3182–210.

Siemer, M. 2019. Employment effects of financial constraints during the great recession. *Review of Economics and Statistics* 101:16–29.

Weesie, J. 2000. Seemlingly unrelated estimation and the cluster-adjusted sandwich estimator. *Stata Technical Bulletin* 9(52).