Financial constraints and economic development: the role of innovative investment

Galina Vereshchagina*

Arizona State University

October 2018

Abstract

This paper argues that accounting for firms' endogenous productivity growth plays an important role in understanding the link between financial and economic development. First, using a simple analytically tractable model, it shows that incorporating endogenous investment in firm productivity into the model amplifies the negative impact of firm financing constraints on economic development, as long as the models with endogenous and exogenous productivity growth are calibrated to match the same data on firm size dynamics and firm owners' income. Second, the paper embeds productivity investment into an otherwise standard variation of the Bewley-Aiyagary-Hugget model used in the existing literature to evaluate the impact of borrowing constraints on economic development. It compares the effects of firm financing constraints in the two models, with endogenous and exogenous firm productivity growth, calibrated in such a way that they are observationally equivalent in the benchmark unconstrained environment. The main result is that the impact on financing constraints on measured TFP and GDP is significantly bigger in the model in which the evolution of firm productivity is endogenous. While measured TFP and GDP fall by 5% and 28% in the model with exogenous productivity growth, they fall by 13% and 37%, respectively, in the model in which firm productivity grows endogenously.

Keywords: borrowing constraints, innovative investment, endogenous productivity growth, economic development, misallocation

JEL codes: E22, L11, L26, D14, D24, O16, O3

*email: Galina.Vereshchagina@asu.edu, phone: 480-965-6860, fax: 480-965-0748.

1 Introduction

The question "to what extent limited access to external financing distorts the allocation of resources in the economy" has been asked in many studies.[1] The channel that has been thoroughly researched is the impact of credit constraints on the allocation of physical capital across firms: Because limited access to external financing leads to misallocation of resources between the poor and the rich firm owners, it adversely affects GDP and measured TFP. This paper draws attention to the fact that, in addition to distorting the allocation of *physical capital* across the firms, borrowing constraints also distort *innovative investment*. Firm owners need financial resources to invest in growth and development of their ideas. Typical examples of such investment are the resources spent on market research, product promotion, supply chain development, adopting best management practices, investment in own and managerial human capital, and many other productivity enhancing investment commonly referred to as investment in intangible or organizational capital. Naturally, limited excess to external financing may not only slow down the process of physical capital accumulation but also restrain the firm productivity growth. Indeed, recent empirical studies[2] document the evidence of such adverse effects. The main objective of this paper is to *study to what extent endogenizing the link between financial constraints and the evolution of firm productivity over time may magnify the effect of financial constraints on economic development*.

For this purpose, the paper develops two models, with exogenous (XG) and endogenous (NG) firm productivity growth, and compares the effects of imposing firm financing constraints in these two environments. In the benchmark scenario, in which firms have unlimited access to external financing, the two models are parameterized in such a way that they produce identical predictions regarding the firm dynamics, income and wealth distribution, as well as the aggregate quantities and prices. Then I introduce a firm financing constraint in the two models, and compare its effects on GDP, measured TFP, and other aggregate variables.

I start by developing a simple analytically tractable example which demonstrates that accounting for innovative investment is not simply a matter of recalibrating the firm technology by accounting for all the production costs, including investment in intangible capital. I show

[1] See, for example, Buera, Kaboski and Shin (2011), Greenwood, Sanchez and Wang (2013), Moll (2014), and Midrigan and Xu (2014), just to name a few.

[2] See, for example, Ferrando and Ruggieri (2015) Gorodnichenko and Schnitzer (2013), Levine and Warusawitharana (2016) and Manaresi and Pierri (2017).

that when the (XG) and (NG) models are calibrated to match the same data on the evolution of firm sizes and firm owners' income, the financing constraints restricting firms' capital input lead to a bigger losses in the (NG) model in which innovative investment is modeled explicitly. The reason for this amplification is that the benefits from productivity investment are persistent in nature: such investments increase not only the current, but also future firm productivity. Thus, having limited access to financing early in life has negative effects on future productivity in the (NG), but not in the (XG), model.

Then I develop full dynamic versions of the (XG) and (NG) models, calibrate them to match the U.S. data in the unconstrained scenario, and quantitatively evaluate the steady state effects of imposing the same firm financing constraint in the two models. The benchmark (XG) model is a standard Bewley-Aiyagari-Hugget model with production risk, different variations of which have been used in existing studies to analyze the impact of borrowing constraints on economic development. In this model, some agents draw productive ideas, and the quality of these ideas evolves over time according to an exogenously given random process. In contrast, in the (NG) model, the evolution of firm productivity is impacted by innovative investment made by the firm owners. The two models are identical in all other respects and, by design, the calibration procedure ensures that they generate the same firm dynamics, income dynamics, and the wealth distribution in the unconstrained benchmark scenario. In all counterfactual experiments, limited access to external financing reduces GDP and TFP in both models, but these effects are significantly bigger in the (NG) model in which innovative investment is modeled explicitly. For example, if the model is calibrated to generate the average share of innovative investment to GDP of 7.7% in the benchmark (NG) scenario, imposing the extreme firm financing constraint (self-financing) reduces the GDP by 28% in the (XG) model, and by 37% in the (NG) model. The differences in the measured TFP effects are even more pronounced: it falls by 5% in the (XG) model, and by 13% in the (NG) model. These findings suggest that accounting for innovative investment plays an important role for understanding the link between financial and economic development.

This paper is closely related to two strands in the literature. First, as mentioned in the opening paragraph, it complements to numerous papers studying the effects of borrowing constraints on economic development; Buera, Kaboski and Shin (2015) provide an excellent overview of this literature.[3] Perhaps, the most closely related to this paper is Midrigan and Xu (2014), who convincingly argue that the intensive margin alone (namely, the misallocation of capital across existing firms) generates a very modest impact of borrowing constraints on measured TFP (typically about 5%). The (XG) model studied in the paper generates the predictions consistent with the findings of Midrigan and Xu (2014). However, as the computations in the (NG) model show, the intensive margin becomes a lot more important in the model in which the firm productivity grows endogenously. While the extent of misallocation of capital across firms is still quite modest, borrowing constraints impede firm productivity growth, and the TFP falls by 2.6 times as much as in the (XG) model.

Second, this paper contributes to the literature which emphasizes the role of endogenous productivity-enhancing investment for understanding the effects of various distortions. Much of this work (e.g., Bhattacharyaa, Guner and Ventura (2013) using the occupational choice model or Bento and Restuccia (2017) using the firm dynamics model) has focused on the impacts of tax distortions, arguing that relatively high marginal tax rates imposed on the high income owners or large firms reduce the incentives to invest in productivity growth, thereby magnifying the distortionary impact of taxation. This paper studies a different form of distortion, the one that affects relatively poor firm owners, and outlines a different, and, to me knowledge, novel channel through which the distortionary effects are magnified.

Finally, two recent papers focus on the role of endogenous productivity in the presence of financial constraints. Mestieri, Schauer and Townsend (2017) argue that poor financial development distorts schooling decision and, via it, impacts productivity of workers and entrepreneurs. Since they focus on schooling, their model does not allow for endogenous firm productivity growth and assumes that all productivity-enhancing investment are made before agents become a workers or entrepreneurs. In contrast, the possibility of gradual productivity investment, as in this paper, allows the firm to gradually save out of the constrained region and then catch up on productivity growth, thereby reducing the potential negative impact of the financing constraints. Despite this, I find that accounting for endogenous gradual productivity investment considerably magnifies the impact of the financing constraints on measured TFP and GDP.

[3] Some papers in this literature, for example Buera, Kaboski and Shin (2011), assume that the firm exogenous productivity remains constant over lifecycle, while the calibration procedure in this paper targets the firm size-age profile. This, however, is not a critical distinction. The main point in this paper is that explicitly recognizing in the model that the path of firm-level productivity over time is endoegenous and depends on firms' innovative investment is an important step in quantifying the impact of financial frictions on economic development.

Caggese (2018) analyzes the role that firm innovative investment play in understanding the imact of financing constraints on the firm growth profile. He argues that incremental investment alone, like the ones considered in this paper, cannot explain the differences in the evolution of firm sizes over time in rich and poor countries, and shows that incorporating radical innovations into the model can address this shortcoming. In addition to focusing on a different question, my paper uses a different modeling environment from Caggese (2018) – production incorporated into a Bewley-Ayiagari-Hugget model, as opposed to a Hopenhayn firm dynamics model – which allows to link the financing constraints to firm owners' endogenous assets, in contrast to an exogenously binding financing constraint assumed in the aforementioned paper. In addition, in an extension considered in this paper (inclusion of traditional firms representing 'push entrepreneurs'), I offer an alternative mechanism that can contribute to explaining some of the patterns in cross-country differences in firm dynamics.

The paper is organized as follows. Section 2 develops a simple analytical example to demonstrate the key mechanism and outline the calibration strategy. Section 3 sets up the full quantitative (XG) and (NG) model. Sections 4 and 5 describe the calibration procedure and the results of the main counterfactual experiments. Finally, Section 6 considers a two-sector extension of the main model, in which innovative investments have a bigger quantitative impact and, in addition, which generates interesting predictions regarding the variations of firm sizes across rich and poor countries.

2 An example

The purpose of this example is to illustrate why accounting for endogenous evolution of firm productivity is important for quantifying the effects of firm financing constraints. One may suspect that explicitly endogenizing firm productivity in the model is redundant if the model is calibrated to account for expenses on innovation while choosing the parameters of the production technology. The simple example below shows that this indeed would be true in a static environment with homogeneous production technology, but this intuition fails in a dynamic setting where innovative investment have persistent effects on future productivity.

Formally, I demonstrate that if the two models, with exogenous and endogenous firm productivity, are calibrated to match the same data on firm sizes (measured by the capital input) and the income of the firm owners (measured by firm profit), financing constraints have a bigger impact in the model with endogenous firm productivity, but only if innovative investments have persistent effects over time. Matching the two aforementioned targets is important if the objective is to assess the impact of firm financing constraints: the extent of misallocation crucially depends on the firm size distribution, and the ability of firm owners to accumulate assets and overcome financing constraints is impacted by their disposable income.[4]

Note that matching the aforementioned calibration targets in the models with endogenous and exogenous firm productivity growth necessitates that the production technology (mapping capital and labor inputs into the revenue) differs across the two models, because in the model in which firm productivity is endogenous, firm owners incur additional expenses on innovative investment. If, in contrast, one were to use the same production technology in the two models, the net income of firm owners would be lower in the model in which firm owners make innovative investment. Naturally, this would magnify the adverse effects of the borrowing constraints because the additional expenses on innovative investment would slow down the process of capital accumulation by firm owners. Such a naive magnification channel, however, is absent if the production technology is calibrated in such a way that the firms of the same size generate identical income for the owners (net of all costs, including expenses on innovation) in the two models. This imposes the discipline in the argument developed below, as well as in the calibration of the full model in Section 4.1.

2.1 Static model

Suppose we observe that an unconstrained firm uses k^* units of physical capital generates income π^* for its owner, and we want to assess the effects of a financing constraint $k \leq \widehat{k} < k^*$. The original unconstrained allocation can be formalized using a model (XG) in which the firm productivity is exogenously given, or a model (NG) in which the firm productivity is endogenous. Assume that in the (XG) model, a firm produces output Ak^γ by using capital input k, while in the (NG) model, a firm produces output $z^\varphi k^\nu$ by using capital input k and productivity z.[5] Suppose that capital is rented on a competitive market at rate R, and the cost of acquiring

[4]Indeed, previous studies – see, for example, Buera, Kaboski and Shin (2015) or Bhattacharyaa, Guner and Ventura (2013) – calibrate the models to match these targets.

[5]For brevity, labor input is omitted from this model. All the results naturally extend to the environment with production technology using a composite input $k^\alpha n^{1-\alpha}$, as long as there are no separate restrictions imposed on

the productivity level z is just z. Then the unconstrained firms in the (XG) environment solves

$$\pi^{XG} = \max_{k} Ak^{\gamma} - Rk \tag{1}$$

implying that the optimal capital and the ratio of profit to capital cost are, respectively:

$$k^{XG} = \left(\frac{\gamma A}{R}\right)^{\frac{1}{1-\gamma}} \quad \text{and} \quad \frac{\pi^{XG}}{Rk^{XG}} = \frac{1}{\gamma} - 1. \tag{2}$$

In the (NG) environment, the unconstrained firm solves

$$\pi^{NG} = \max_{k,z} z^{\varphi} k^{\nu} - Rk - z, \tag{3}$$

implying that the optimal capital and the ratio of profit to capital cost are, respectively:

$$k^{NG} = \left[\left(\frac{\nu}{R}\right)^{1-\varphi} \varphi^{\varphi}\right]^{\frac{1}{1-\varphi-\nu}} \quad \text{and} \quad \frac{\pi^{NG}}{Rk^{NG}} = \frac{1-\varphi}{\nu} - 1. \tag{4}$$

In order for the two models to produce the same predictions regarding the firm sizes and the income of firm owners, it must be that $k^{XG} = k^{XG} = k^{*}$ and $\pi^{XG} = \pi^{NG} = \pi^{*}$, implying that

$$\gamma = \frac{\nu}{1-\varphi} \quad \text{and} \quad A = (1-\varphi)\varphi^{\frac{\varphi}{1-\varphi}} \tag{5}$$

Now assume that the firm faces a borrowing constraint $k \leq \widehat{k} \leq k^{*}$.[6] Then, trivially, the firm's profit in the (XG) model is given by

$$\widehat{\pi}^{XG} = A\widehat{k}^{\gamma} - R\widehat{k}.$$

In the (NG) model, the constrained firm chooses the optimal investment in productivity $\widehat{z}^{NG}$ given the amount of available capital $\widehat{k}$:

$$\widehat{z}^{NG} = \left(\varphi\widehat{k}^{\nu}\right)^{\frac{1}{1-\varphi}},$$

the use of labor input n.

[6]Note that the financing constraint restricts only the amount of capital input used in the firm, but not the amount of productivity investment. Had one imposed the constraint $k + z \leq \widehat{k}$, the effects in the (NG) model would naturally be amplified.

which implies that

$$\widehat{\pi}^{NG} = (1-\varphi)\varphi^{\frac{\varphi}{1-\varphi}}\widehat{k}^{\frac{\nu}{1-\varphi}} - R\widehat{k}. \tag{6}$$

Obviously, calibration strategy (5) results in $\widehat{\pi}^{XG} = \widehat{\pi}^{NG}$, implying that if the two models are calibrated to match the same observations on firm size (measured by the amount of capital used) and the income of firm owners (measure by the firm profit), they would also make similar predictions regarding the effects of firm financing constraints on owners' income. This finding is not surprising; it simply reinforces the idea that, under homogeneous production technology, the variety of inputs can be compressed into one common input.

Looking forward, the firm profit maximization problem will be embedded into the dynamic consumption / savings problem of the firm owner as in Quadrini (2000). The argument above suggests that if the firm productivity were solely determined by the firms' innovative expenses incurred within the same period, there would be no need to explicitly model such innovative expenses in order to evaluate the effects of financial constraints, as long as the model is calibrated to match firm sizes and income of firm owners. This is because a financial constraint would have the same impact on firm owners' income regardless of whether the firm productivity is endogenous, and thus would affect in the same way the owner's ability to accumulate assets over time. The dynamic example developed below demonstrates that these predictions do not hold if the firm's innovative investment have persistent effects.

2.2 A simple two-period model

Next, let us extend the simple static models above to a two-period settings. Suppose that the firm lives for two periods and, in an unconstrained environment, the models are targeted to match the same observations on the evolution of firm sizes (k_1^*, k_2^*) and profits (π_1^*, π_2^*). The objective is to analyze whether the effects of a financial constraint in period 1 depend on whether the evolution of firm productivity is exogenously assumed or endogenously determined within the model.

Assume that in the (XG) model the firm has technology $A_1k_1^{\gamma_1}$ and $A_2k_2^{\gamma_2}$ in periods 1 and 2, respectively. In the absence of financing constraints, the firm solves

$$\Pi^{XG} = \max_{k_1,k_2} A_1k_1^{\gamma_1} - Rk_1 + \beta\cdot\left(A_2k_2^{\gamma_2} - Rk_2\right), \tag{7}$$

where β is the discount factor (if $\beta = 0$, the model collapses to the static setting analyzed above). The fact that the production technology in period 2 is different from the production technology in period 1 is not important, it simply accounts for the fact that profit in period 2, in a reduced form, may represent the stream of profits generated from period 2 onwards. Obviously, (7) is just a sequence of two static problems, and the unconstrained solution for each is given by (2).

Suppose that, in the (NG) model, the firm's innovative investments have persistent effect. Namely, assume that, having invested z_1 in period 1, the firm has access to technology $z_1^{\varphi_1} k^{\nu_1}$ and $z_1^{1-\nu_2} k^{\nu_2}$ in periods 1 and 2, respectively.[7] Then the firm's decision problem is

$$\Pi^{NG} = \max_{k_1, z_1, k_2} z_1^{\varphi_1} k_1^{\nu_1} - Rk_1 - z_1 + \beta \cdot \left(z_1^{1-\nu_2} k_2^{\nu_2} - Rk_2 \right). \tag{8}$$

The choice of k_2 in the second period is static, yielding

$$\Pi^{NG} = \max_{k_1, z_1} z_1^{\varphi_1} k_1^{\nu_1} - Rk_1 - z_1 + \beta \Omega z_1, \tag{9}$$

where $\Omega = \frac{1-\nu_2}{\nu_2} \left(\frac{\nu_2}{R} \right)^{\frac{1}{1-\nu_2}} \cdot R$. The solution to (9) is

$$k_1^{NG} = \left[\left(\frac{\nu_1}{R} \right)^{1-\varphi_1} \cdot \left(\frac{\varphi_1}{1-\beta\Omega} \right)^{\varphi_1} \right]^{\frac{1}{1-\varphi_1-\nu_1}} \quad \text{and} \quad z_1^{NG} = \left(\frac{\varphi_1}{1-\beta\Omega} (k_1^{NG})^{\nu_1} \right)^{\frac{1}{1-\varphi_1}}, \tag{10}$$

and the share of firm's profit to its capital cost in period 1 is

$$\frac{\pi_1^{NG}}{Rk_1^{NG}} = \frac{1-\varphi_1-\beta\Omega}{(1-\beta\Omega)\nu_1}. \tag{11}$$

Notice that if $\beta = 0$, (10) and (11) collapse to (4).

Thus, if the (XG) and (NG) models are calibrated to match the same data on firm size dynamics and firm owners' profits, it must be that

$$\gamma_1 = \frac{(1-\beta\Omega)\nu_1}{1-\phi_1-\beta\Omega} \quad \text{and} \quad \left(\frac{\gamma_1 A_1}{R} \right)^{\frac{1}{1-\gamma_1}} = \left[\left(\frac{\nu_1}{R} \right)^{1-\varphi_1} \cdot \left(\frac{\varphi_1}{1-\beta\Omega} \right)^{\varphi_1} \right]^{\frac{1}{1-\varphi_1-\nu_1}} \tag{12}$$

[7] The Cobb-Douglas technology in period 2 allows for analytical tractability, but the results can be easily generalized beyond this assumption. What's important for the argument is that the profit in period 2 increases with the productivity investment made in period 1.

to equalize the first period's outcomes in the two models, and $A_2 = (z_1^{NG})^{1-\nu_2}$ and $\gamma_2 = \nu_2$ to equalize the second period's outcomes. Note that such a calibration strategy also results in $\Pi^{XG} = \Pi^{NG}$, since firms' profit in both models is equalized in each period.

Now suppose that the capital input in the first period is limited by $k_1 \leq \widehat{k}_1 \leq k_1^*$. First, observe that, in the (NG) model, this would induce the firm to make smaller innovative investment z_1 in period 1 and, hence, result in less capital input and profit in period 2. In contrast, imposing a limit on k_1 in period 1 in the (XG) model has no impact on the firm size in the second period. This simple observation underscores an obvious mechanism through which persistent effects of innovative investment may alter the link between the financial constraints and economic activity. Second, one can argue that limiting the capital input has a bigger impact on the total life-time profit in the (NG) model than in the (XG) model.[8] To see this, denote by $\widehat{\Pi}^{XG}(\widehat{k}_1)$ and $\widehat{\Pi}^{NG}(\widehat{k}_1)$ the constrained profit functions computed as (7) and (8), respectively, but with the additional constraint $k_1 \leq \widehat{k}_1$. The calibration strategy described above ensures that

$$\widehat{\Pi}^{XG}(k_1^{XG}) = \widehat{\Pi}^{NG}(k_1^{NG}). \tag{13}$$

Both profits decline as $\widehat{k}_1$ falls below $k_1^* = k_1^{XG} = k_1^{NG}$. However, because reducing $\widehat{k}_1$ also lowers z_1 in the (NG) model, life-time profit $\widehat{\Pi}^{NG}(\widehat{k}_1)$ is affected by more than $\widehat{\Pi}^{XG}(\widehat{k}_1)$. At the limit, as $\widehat{k}_1$ converges to 0, $\widehat{\Pi}^{NG}(\widehat{k}_1)$ also approaches 0 while $\widehat{\Pi}^{XG}(\widehat{k}_1)$ remains positive because the firm's second period profit in the (XG) model is not affected by the constraint on capital input imposed in the first period. Formally, one can verify (see section 8.1 in the Appendix) that, for the calibrated parameters (12),

$$\frac{\partial \widehat{\Pi}^{XG}(\widehat{k}_1)}{\partial \widehat{k}_1} < \frac{\partial \widehat{\Pi}^{NG}(\widehat{k}_1)}{\partial \widehat{k}_1}, \quad \text{as long as} \quad \widehat{k}_1 \leq k_1^{XG} = k_1^{NG}, \tag{14}$$

which, in conjunction with (13), implies that $\widehat{\Pi}^{XG}(\widehat{k}_1) > \widehat{\Pi}^{NG}(\widehat{k}_1)$ for all $\widehat{k}_1 < k_1^*$.[9] The following Proposition summarizes these results:

[8] In other words, reduced innovative investment in the (NG) model does not simply shift the profit from period 2 to period 1, but leads to a loss in life-time profit.

[9] Mechanically, the equivalence argument used in the static version of the model fails here because the life-time profit in the (XG) model has a 'constant' term – the second period profit – which is not affected by the change in capital. Had the second-period profit been proportional to either k_1 or $k_1^{\gamma_1}$, and the model been calibrated to match the total profit and k_1, the two-period model would be isomorphic to the one-period example studied earlier.

Proposition 1 *Suppose that the unconstrained two-period (XG) and (NG) models are calibrated to match the same values of capital input and profits in both periods. Then, imposing the constraint on capital input in period 1 results in*

(i) a reduction in period 2's productivity, capital, and profit in the (NG) model, but not in the (XG) model;

(ii) a bigger decline in the life-time profit in the (NG) model than in the (XG) model.

Proposition 1 has two important implications. First, it illustrates that, due to persistency of firm's innovative investment, a financing constraint in one period reduces the firm's future productivity and size. Second, because the financing constraint leads to a bigger reduction in firm profit in the (NG) model than in the (XG) model, it has a bigger negative effect on the firm owner's ability to accumulate assets and overcome future financial constraints. The full model developed in the next section incorporates these channels into a dynamic consumption / savings framework with uninsured risk in order to quantify their importance.

3 The Full Model

3.1 Brief description of the environment

The modeling environment is a variation of the Bewley-Aiyagari-Hugget[10] model with production risk. The economy is populated by a continuum of agents of mass one who receive stochastic labor productivity shocks, supply labor to the competitive labor market, and can save, subject to a borrowing constraint, in a risk-free asset traded on a competitive market. In addition, with some probability, agents receive productive ideas and can operate them, thereby becoming firm owners.

Productive ideas generate output using labor and capital inputs, and their productivity evolves over time. I consider two settings, with endogenous and exogenous modern firm growth, labeled (NG) and (XG), respectively. In the (NG) model, the firm owners can invest in building up intangible capital to raise firm productivity. In contrast, in the (XG) model, the productivity of a firm grows stochastically over time, as it is commonly assumed in existing models studying

[10]Hugget (1993)

the effects of financial frictions on economic development.[11] In the benchmark scenario, firms have unlimited access to external financing and rent as much capital as they need at the risk-free interest rate. In the counterfactual experiments, the firm owners face financing constraints and thus may use inefficiently little capital.

The objective of the analysis is to build and calibrate (NG) and (XG) models in such a way that they are observationally equivalent in the benchmark scenario without financial frictions, and then compare the effects of firms' limited excess to external financing across the two environments. This allows to identify the relative importance of endogenous firm productivity growth for understanding the link between financial and economic development.

3.2 Formal setup

There is a continuum of consumers of mass one, with preferences given by

$$E\sum_{t=0}^{\infty}\beta^t u(c_t), \quad 0<\beta<1.$$

In period t, a consumer has total assets a_t, and may have a productive idea of quality z_t which generates proprietary income $\pi(a_t, z_t)$ specified below. Additionally, in every period, each consumer receives labor productivity shock $\xi_t \in F_W(\xi)$, i.i.d. across agents and over time, and earns income $w\xi_t$ by supplying labor to the competitive labor market.[12] After the income in the current period is realized, a consumer chooses how much to consume and how much to save in a risk-free asset which generates the rate of return r, subject to the borrowing constraint $a_{t+1} \geq \underline{a}$.

A consumer operating an idea z_t has access to production technology

$$y(z_t, k_t, n_t) = z_t^{1-\gamma}\left(k_t^{\alpha} n_t^{1-\alpha}\right)^{\gamma}, \quad \gamma, \alpha \in (0,1) \tag{15}$$

where k_t and n_t are capital and labor inputs, respectively. The capital used in the firm may be subject to a borrowing constraint, $k_t \leq \overline{k}(a_t, z_t)$. At the most extreme scenario, all firm owners have to finance the operations of their firms from own assets, i.e. $k_t \leq a_t - \underline{a}$ is imposed. Thus,

[11] See, for example, Buera, Kaboski and Shin (2011), Greenwood, Sanchez and Wang (2013) or Moll (2014).

[12] The analysis focuses on a stationary equilibrium, so prices do not depend on t.

the proprietary income generated by a firm of type z_t is

$$\begin{aligned}\pi(z_t, a_t) = &\max_{k_t, n_t} z_t^{1-\gamma}\left(k_t^{\alpha} n_t^{1-\alpha}\right)^{\gamma} - w n_t - (r+\delta) k_t \\ &\text{s.t.} \quad k_t \leq \overline{k}(a_t, z_t)\end{aligned} \tag{16}$$

Importantly, as in the simple model in the previous section, the financing constraint only applies to the amount of physical capital used in the firm. The cost of labor and productivity investment (specified below) are not directly restricted by the constraint.

The productivity of the idea evolves over time. In the (XG) model, the firm productivity changes according to the exogenous transition process $\Pi(z_{t+1}|z_t)$. In contrast, in the (NG) model, the productivity in the next period depends on the productivity z_t in the current period, and on how much x_t the firm owner invests in intangible capital:

$$z_{t+1} = (1-\delta_z) z_t + B z_t^{\theta} x_t^{1-\theta} = \left(1 - \delta_z + B\left(\frac{x_t}{z_t}\right)^{1-\theta}\right) \cdot z_t. \tag{17}$$

The functional form of this law of motion is borrowed from Bhattacharyaa, Guner and Ventura (2013). If no investment in intangible capital is made, the quality of the idea depreciates at rate δ_z. As seen from the second term in parenthesis, the rate of growth of firm productivity increases with x_t at a decreasing rate, but is smaller for higher z_t because more productive firms are harder to grow. Solving (17) for x_t yields the cost of achieving productivity z_{t+1} in period $t+1$ given that the previous period's productivity was z_t:

$$x(z_t, z_{t+1}) = \left(\frac{z_{t+1} - (1-\delta_z) z_t}{B z_t^{\theta}}\right)^{\frac{1}{1-\theta}}, \tag{18}$$

which will be used in the decision problems below.

The total disposable income of a firm owner who operates the idea of quality z_t, invests x_t in intangible assets, and experiences a labor productivity shock ξ_t is

$$\pi(z_t, a_t) - x_t + w\xi_t. \tag{19}$$

Note that it is assumed that the consumer does not need to forego wages in order to manage a productive idea, which distinguishes this modeling environment from the span-of-control model in Lucas (1978), as well as from many occupational choice models studying the link between the financial and economic development. In this case, all the productive ideas are implemented regardless of the extent of financial development, and the presence of financial constraint impacts TFP only via two *intensive* margins: misallocation of labor and capital across firms, and reduced firm productivity due to lower innovative investment. Previous studies have found that misallocation alone (the former margin) plays little role in the link between financial and economic development because productive firms accumulate assets quickly and borrowing constraints have relatively small effect on them.[13] This paper focuses on the effect of financial constraints on innovative investment, which is only present in the (NG) model and can magnify the impact of financial frictions on TFP via the intensive margin. Thus, to isolate this effect, I consider an environment in which, by design, all the effects of financial constraints manifest only through the intensive margin.[14]

The transition function $\Pi(z_{t+1}|z_t)$ and the law of motion (17) guide the evolution of productivity for surviving incumbent firms. The arrival and survival of productive ideas follows an exogenously given process. If an agent starts the period without any idea, he draws one with probability η_M and chooses whether or not to pursue it. If the idea is pursued, its initial quality $z_t \sim F_M(z)$ is realized, i.i.d. across agents. Productive ideas disappear at the end of the period with exogenous probability χ_M. The agent operating an idea learns if the idea survives in the next period before all the investment decisions, including innovative investment, are made.[15]

Now let us set up the agents' decision problems recursively. The analysis focuses on a stationary equilibrium, so the prices (w, r) are omitted from the set of the state variables. Denote by $V_W(a)$ the expected value of the agent pursuing no productive ideas in the current period and by $V_M(a, z)$ the expected value of the agent pursuing an idea of quality z, with a being the asset level in the beginning of the period. It is also convenient to introduce an auxiliary value function, $V_0(a)$, denoting the value of the agent who has no idea in the beginning of the period

[13]See, for example, Midrigan and Xu (2014).

[14]Adding extensive margin creates an additional channel via which endogenous innovative investment multiply the impact of borrowing constraints on economic activity. Thus, focusing on intensive margin only, does not only allow to keep analysis simple and transparent, but also provides a *lower bound* on the quantitative effects of endogenous innovative investment.

[15]This assumption helps to generate higher growth rates among smaller firms, consistently with the data. Had the intangible investment decision been made before the realization of survival shock χ_M, smaller firms, operated by owners with less assets, and thus more risk averse, would choose to invest less in intangible capital because the possibility of idea death poses a lot of risk and, therefore, would exhibit smaller growth rates than larger firms.

but still has a chance to draw one. Such value is determined as

$$V_0(a) = (1 - \eta_M)V_W(a) + \eta_M \max\left\{V_W(a), \int V_M(a, z)dF_M(z)\right\}. \tag{20}$$

Since operating a productive idea does not require to forego labor earnings, the discrete choice between the value $V_W(a)$ and the expected value of $V_M(a, z)$ can be omitted. The expected value of the worker $V_W(a)$ is given by

$$V_W(a) = \int \max_{a' \geq \underline{a}} \{u(w\xi + a - \frac{a'}{1+r}) + \beta V_0(a')\} \, dF_W(\xi). \tag{21}$$

The value of operating an idea of quality z depends on whether the firm productivity grows exogenously or endogenously. In an exogenous firm growth model (XG), $V_M(a, z)$ is given by

$$\begin{aligned} V_M(a, z) = \chi_M \cdot \int \max_{a' \geq \underline{a}} \left\{ u(w\xi + \pi(a, z) - \frac{a'}{1+r}) + \beta V_0(a') \right\} dF_w(\xi) \\ +(1 - \chi_M) \cdot \int \max_{a' \geq \underline{a}} \left\{ u(w\xi + \pi(a, z) - \frac{a'}{1+r}) \right. \\ \left. + \beta \int \max\{V_M(a', z'), V_0(a')\} \, d\Pi(z'|z) \right\} dF_w(\xi). \end{aligned} \tag{22}$$

In contrast, in an endogenous firm growth model (NG), the value of operating an idea of quality z is

$$\begin{aligned} V_M(a, z) = \chi_M \cdot \int \max_{a' \geq \underline{a}} \left\{ u(w\xi + \pi(a, z) - \frac{a'}{1+r}) + \beta V_0(a') \right\} dF_w(\xi) \\ +(1 - \chi_M) \cdot \int \max_{a' \geq \underline{a}, z' \geq (1-\delta_z)z} \left\{ u(w\xi + \pi(a, z) - x(z, z') - \frac{a'}{1+r}) \right. \\ \left. + \beta \max\{V_M(a', z'), V_0(a')\} \right\} dF_w(\xi), \end{aligned} \tag{23}$$

where $x(z, z')$ is defined in (18). In both settings, if the firm owner learns that the firm does not survive in the next period, the continuation value is $V_0(a')$, but if the firm does not die exogenously, there is a choice of whether to abandon it. The difference between (22) and (23) is that in the latter case, in addition to choosing how much to save in a risk-free asset a', the firm owner also chooses how much $x(z, z')$ to invest in intangible capital. As in (26), the discrete choice between continuing with an idea and foregoing it to take $V_0(a)$ can be omitted because the former option is always chosen.

The optimal policies in (21)-(23), together with the evolution of idea and labor productivity shocks, induce a stationary distribution over agents' asset levels, idea ownership and idea quality. Denote by $\mu_W(a)$ and $\mu_M(a,z)$ the stationary distributions across assets for the workers, and stationary distribution across asset and idea productivitites the firm owners. Then, in a stationary competitive equilibrium, prices w and r clear the labor and capital market, respectively:

$$\int n(a,z)\, d\mu_M(a,z) = 1 \tag{24}$$

$$\int k(a,z)\, d\mu_M(a,z) = \int a\, d\mu_W(a) + \int a\, d\mu_M(a,z), \tag{25}$$

where $n(a,z)$ and $k(a,z)$ represent the labor and capital demand of the modern firm of quality z whose owner has asset level a. In the benchmark unconstrained economy, these quantities depend only on firm productivity z, but in the constrained economy they vary with the assets of the firm owner.

At this stage, a few remarks would help to clarify the role of some of the modeling features. I have on purpose tried to keep the model as simple as possible, in order to isolate the channel via which endogenizing firm productivity impacts the effects of financing constraints. To simplify the model even further, one could omit the labor productivity shocks ξ_t; the uninsured risk associated with arrival and death of production ideas is sufficient to produce a non-degenerate asset distribution as long as $\beta(1+r) < 1$. However, such a specification of the model leads to too much concentration at the bottom of the wealth distribution. The presence of uninsured labor risk helps to create more wealth dispersion among relatively poor agents, and generating the realistic wealth distribution is obviously important for evaluating the effects of financial frictions. The model makes a simplistic assumption about the evolution of ξ_t; this is because the focus of the model is not to reproduce the dynamics of earnings and assets of workers but to focus on the dynamics of firms. Adding persistence in the labor productivity shocks would add an additional state variable into the decision problem of the firm owner (23), which would make computations a lot harder because even under the current formulation the problem already has two continuous state and two continuous choice variables.

Table 1: Parameters set outside of the model or calibrated individually

Parameter	Value	Target and/or Source
σ , risk aversion	2	standard value
$\underline{a}$, borrowing limit	1	mean workers's annual income
δ, capital depreciation rate	0.06	Hugget (1996)
χ_M, ideas' death rate	0.0738	firm exit rate, BDS data, 1998-2006
α, capital vs. labor returns 15	0.3264	share of capital to labor cost

4 Parameter values

The model is parameterized to match the relevant features of the U.S. firm dynamics, earnings, and wealth distribution in the unconstrained economy. Sections 4.1 and 4.2 describe the key steps of the calibration in the (NG) and (XG) models, while sections 8.2 and 8.3 in the Appendix briefly describe the numerical algorithm solving the model and offer some technical details. The main numerical results are presented in section 5, which compares the effects of a borrowing constraint in (NG) and (XG) models.

4.1 Calibrating the (NG) model

The model's parameters are chosen in such a way that the labor and capital markets clear at the normalized wage $w = 1$ and the annual interest rate $r = 0.03$. There are fours sets of parameters to be calibrated – the parameters affecting preferences, production technology, endogenous modern firm growth in (NG) model and exogenous evolution of firm productivity in (XG) model, as well as the death / arrival of productive ideas. Tables 1 and 2 list all the parameters, their values in the benchmark specification, as well as the calibration targets.

Table 1 summarizes the parameter values that are set outside the model. It is assumed that the consumers have CRRA utility function, $u(c) = \frac{c^{1-\sigma}-1}{1-\sigma}$ with the risk aversion parameter $\sigma = 2$, as is common in macroeconomic literature. The borrowing limit $\underline{a}$ is set to be equal to the mean annual income of the workers, which is normalized to 1. Turning to the firm evolution and technology, the capital depreciation rate is set at $\delta = 0.06$ following Hugget (1996). The exit rate of the modern firms χ_M, together with other relevant moments of firm dynamics, are directly computed from the Business Dynamics Statistics (BDS) data, see section 8.3.1 in the Appendix for details.[16] Finally, the parameter α in the production technology (15) is set to match the share of capital to labor cost ratio.

The remaining parameters ($F_W(\xi)$, η_M, $F_M(z)$, γ, B, δ_z, θ, and β) are calibrated jointly to match the following moments: the size distribution of age-0 firms, the average firm size, the share of the the top 5% income earners, the equilibrium wage, the share of total revenues spent on intangible investment, the average growth of firms over the first 15 years of life, the equilibrium interest rate and the percentage of population with negative asset holdings. Table 2 summarizes these parameters, along with the targeted moments.

The labor productivity shock is assumed to be binary,[17] taking values ξ_H with probability p and ξ_L with probability $1-p$, normalized such that $p\xi_H + (1-p)\xi_L = 1$. The values of ξ_H and p affect the fraction of the population with assets below 0 and the volatility of the idiosyncratic component of the earnings process, and are pinned down to match their empirical counterparts.

Given the value of χ_M (calibrated above), the arrival rate η_M of productive ideas determines the average firm size: since the steady state mass of productive ideas is η_M/χ_M and every agent supplies labor to the labor market, the average firm size is $\frac{1}{\eta_M/\chi_M}$. Its counterpart in the BDS is 22.7, implying that $\eta_M = 0.0034$.

The returns to scale parameter γ affects the profit to output ratio of the firms, and thus controls the income share of the firm owners who, given the aforementioned entry and exit rates constitute about 4.4% of the population. Idea ownership is also highly correlated with wealth holding and, therefore, with capital income. Thus, the choice of γ affects the income share of the top 5% of population which, according to Piketty and Saez (2003) amounts to 26%.[18] Once the technology parameters α and γ are calibrated, the distribution of the productivity shocks for the entering firms $F_M(z)$ is pinned down to match the distribution of age 1 firms in the BDS

[16] BDS is publicly available at https://www.census.gov/ces/dataproducts/bds/data.html. I use the averages across 1998-2006 to construct all the relevant statistics.

[17] Since the sole purpose of the uncertainty in labor productivity is to generate less concentrated wealth distribution, a more dense grid for labor productivity offers little advantage, but makes the already intense computational procedure more time consuming.

[18] The calibrated value $\gamma = 0.77$ results in the capital-output ratio of 2.8, consistent with the values used in the macro literature.

Table 2: Parameters calibrated jointly, (NG) model

Parameter	Value	Target / Source	Data	Model
ξ_H, earnings shock	1.209	fract. of pop. with negative assets Daz-Gimnez, Glover and Ros-Rull (2011)	0.10	0.11
p, prob. of ξ_H	0.5	var. of log earnings Hugget (1996)	0.045	0.045
η_M , arrival of ideas	0.0034	average firm size, BDS data	22.7	22.8
γ, returns to scale in (15)	0.77	top 5% income share Piketty and Saez (2003)	0.26	0.26
$F_M(z)$, new ideas distr.		BDS data, see Table 6		
B, productivity in (17)	1.4	normalized equilibrium wage	1	1
δ_z, depreciation in (17)	0.05	share of intangible investment in output McGrattan and Prescott (2010), Corrado et al. (2018)	0.07 - 0.15	0.077
θ, returns to scale in (17)	0.10	average size of 10-15 y.o. firms, BDS	15.9	15.3
β, time discount factor	0.9425	equilibrium interest rate	0.03	0.03

data (reported in Table 6 in the Appendix), assuming that the equilibrium wage is $w = 1$ and the interest rate is $r = 0.03$.

The rest of the parameters are tightly linked together. Intuitively, increasing B increases the firm growth and, via it, the average firm size.[19] Increasing δ_z implies that the productive ideas depreciate faster and, therefore, more innovative investment is required in order to match the targeted average firm size. Parameter θ controls how much harder it is to grow high- versus low-productivity firms, and, thereby, affects how much of the firm's life-time growth is achieved over the first 15 years of life. Finally, β affects consumers' incentives to save and its choice ensures

[19]Formally, changes in B impact the labor demand. The choice of η_M described above guarantees that the labor market would clear at $w = 1$ if the modern firm, on average hire 22.7 workers. Thus, adjusting B to match the average firm size also ensures that the labor market is in equilibrium at $w = 1$.

that the asset market clears at $r = 0.03$, while borrowing limit $\underline{a}$ affects the fraction of agents with negative assets.

As can be seen from Table 2, the model matches the data well. The only moment that warrants additional discussion is the share of intangible investment in output. Since this variable is hard to measure, existing literature provides a wide range of estimates for it, anywhere from 7% to 15%.[20] In the benchmark calibrated version of the model, the intangible investment share is a rather conservative 7.7%. Naturally, the bigger is the share of innovative investment in revenue, the bigger is their impact on the link between financial and economic development.

4.2 Calibrating the (XG) model

Since the purpose of the analysis is to study to what extent accounting for endogenous firm growth impacts the link between financial and economic development, the (XG) model is calibrated in such a way that in all the relevant dimensions its benchmark unconstrained allocation is identical to that of the (NG) model. In particular, to be comparable, the two models must generate identical firm dynamics, income dynamics, wealth distribution, as well as the aggregate variables in the unconstrained environment. In other words, the calibration procedure of the (XG) model treats the simulated outcome of the (NG) model as the data, and sets the parameters to match the same set of moments that were targeted in the calibration of the (NG) model (and which are the typical targets in the previous studies).

To ensure that this is the case, the parameters reported in Table 1 remain unchanged, implying that the workers' earnings, as well as the ideas' ownership, exhibit the same dynamics in both models. To guarantee that the earnings of the firm owners are also comparable across the two models, the parameters guiding the production technology must be recalibrated because in the (XG) model there are no expenses associated with innovative investment, and the share of firms' profit relative to their output is just $1-\gamma$. In contrast, in the (NG) model, the firm owners spend additional $X^{NG}/Y^{NG} = 0.077$ of revenue on innovative investment, where X^{NG} and Y^{NG} are the aggregate amounts of innovative investment and output in the (NG) model. Thus, using the same value of γ in both settings would result in neglecting an additional expense the firm owners incur in the (NG) model, which would obviously magnify the effect of financial frictions in the model with innovative investment because these additional expenses would slow

[20]See, for example, McGrattan and Prescott (2010) or Corrado et al. (2018).

down the asset accumulation by firm owners.[21] Thus, to avoid such a short-sided approach, I follow the approach used in the simple model in Section 2 and choose the span-of-control parameter γ^{XG} in the (XG) model in such a way that the average business income of firm owners, net of the expenses on capital, labor, and innovative investment (the latter occurring only in the (NG) model), are equalized across the two calibrated environments. The firm owners' total disposable business income in the (NG) model is $\left(1-\gamma^{NG}\right)Y^{NG}-X^{NG}$. The calibration procedure must assure that the two models generate the same aggregate output net of all costs, implying that $Y^{XG}=Y^{NG}-X^{NG}$ must hold.[22] Thus, the firm owners' earnings are equalized in the two models if $\left(1-\gamma^{XG}\right)\left(Y^{NG}-X^{NG}\right)=\left(1-\gamma^{NG}\right)Y^{NG}-X^{NG}$, implying that[23]

$$\gamma^{XG}=\frac{\gamma^{NG}}{1-X^{NG}/Y^{NG}}.$$

Plugging $\gamma^{NG}=0.77$ and $X^{NG}/Y^{NG}=0.077$, results in $\gamma^{XG}=0.837$.

To ensure that the firms in both models use the same amounts of labor and capital inputs at the same prices, the exogenous productivity shocks in the (XG) model are then rescaled relative to their values in the (NG) model. After that, the process of evolution of productivity shocks $\Pi(z'|z)$ is directly constructed using the simulations of the calibrated (NG) model to ensure that the evolution of the firm sizes across the two models is similar.[24] Finally, the grid for entrant's productivity levels is recalibrated to match the same size distribution of entrants as in the (NG) model. This procedure ensures that the evolution of the firms (in terms of employment and physical capital dynamics), as well as the income of their owners (measured by firm profits net of all the expenses, including innovative investment in the (NG) model), is indeed nearly identical in the two environments. As a result, the values of the time discount factor β and the borrowing limit $\underline{a}$ remain remain the same as in the (NG) model because the income profiles in the two models are similar by construction.

[21] Additionally, using the same γ in both settings would imply that the firm owners accumulate less assets in the (NG) model than in the (XG) model. As a result, the time discount factor β would have to be adjusted to clear the credit market at the same interest rate, implying that the wealth distributions in the two models would be different.

[22] By the choice of productivity levels described below, the two models produce identical firm dynamics, implying the same aggregate costs of labor and capital.

[23] Note that this condition is equivalent to (5) assuring that (XG) and (NG) models are observationally equivalent in the simple example studied in Section (2).

[24] Even though the law of motion for firm productivity in (17) is deterministic, different firms with the same current productivity z may have different productivity levels z' in the next period because their owners differ in how much assets they have and may decide to invest different amounts in firm productivity growth. Thus the calibrated $\Pi(z'|z)$ is non-degenerate.

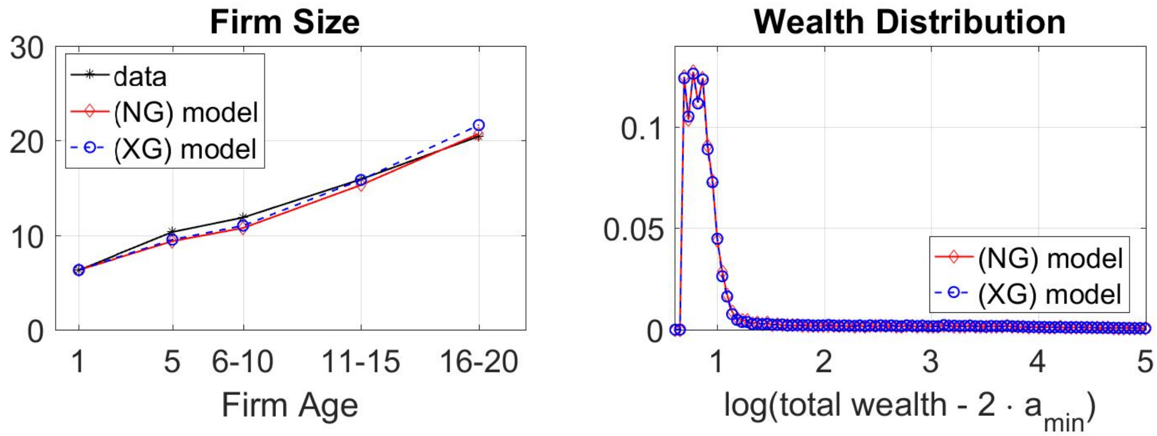

Figure 1: Firms size-age profile (left) and wealth distribution (right) in the calibrated (NG) and (XG) models.

Figure 1 illustrates the firm size / age profile and the asset distributions in the two models. As can be seen, the benchmark allocations in the two models are very similar. The slight differences appear because in the (NG) model the rate of investment in intangible assets actually varies across the firms (it is higher for smaller and younger firms), while the calibration procedure only accounts for the average rate of innovative investment. Notably, while the (NG) model was calibrated to match only the average sizes of the entrants, the firms that are 11-15 years old, and all the firms, it reproduces very well the overall evolution of firm size over age in the BDS data.

5 Findings: the effects of financial frictions in (NG) and (XG) models

This section compares the steady state effects of borrowing constraints in the two models, with endogenous and exogenous firm productivity growth. I start by considering the most extreme case, when the firm owners have no access to external financing, and must finance firm's expenses on physical capital from own assets. For consistency, since consumers are allowed to borrow up to $-\underline{a}$, I also allow firm owners borrow up to $-\underline{a}$, implying that

$$k \leq a - \underline{a}.$$

To understand the role of the borrowing constraints, it is convenient to distinguish between the partial and general equilibrium effects. In the (NG) model, for given prices w and r, imposing the borrowing limits discourages innovative investment in two ways. First, the marginal product of investment in the firm's future productivity decreases because the firm is limited in how much the future capital can be adjusted in response to higher future productivity z'. This results in lower innovative investment in the current period and reduces firm's productivity in the following periods because innovative investment have persistent effects.[25] Second, the constrained firm owners receive less business income. As described in the simple example in section 2, this effect is more pronounced in the (NG) model than in the (XG) model, implying that the firm owners accumulate wealth slower and make further reductions in innovative investment in the (NG) model. Overall, financial frictions have an adverse effect on firm productivity growth in the (NG) model, but not in the (XG) model, where firm productivity evolves exogenously. Therefore, holding prices and a distribution across assets fixed, borrowing constraints lead to a larger GDP and TFP declines in the (NG) model than in the (XG) model. Firm owners' lower income, in turn, induces a more pronounced leftward shift in the wealth distribution, further exacerbating the impact of borrowing constraints in the (NG) model.[26] As a result, for given w and r, the firm productivity, output and the labor demand decrease by more in the (NG) model than in the (XG) model. In the calibrated (NG) model, for $w = 1$ and $r = 0.03$, the steady-state labor demand, measured TFP[27] and output fall by 90%, 34% and 90%, respectively, while in the (XG) model the same variables decline by only 65%, 3.5% and 65%, respectively.

To clear the labor market, equilibrium wages must fall. This decline mitigates the large partial equilibrium effects reported above, because it allows firm owners to generate more profit and accumulate assets at a faster speed to start operating and investing in productivity at the

[25] This reduction in returns to innovative investment due to binding capital constraint manifested in the simple example in section 2 via a reduction in z_1 in response to the financing constraint in that period.

[26] Note that this effect was absent in the simple example considered in Section 2 because the borrowing limit in that example was exogenous, while in the main model borrowing limits are endogenous, in the sense that they are determined by the endogenous wealth distribution.

[27] The TFP is measured as $\frac{Y}{(N^{1-\alpha}K^{\alpha})^{\gamma}}$, where Y, N and K are the aggregate quantities of output, labor and capital, and parameters α and γ are calibrated separately in each model as described in sections 4.1 and 4.2.

Table 3: The effects of borrowing constraints in (NG) and (XG) models

	Benchmark	Self-financing			
		open economy		closed economy	
		(NG)	(XG)	(NG)	(XG)
interest rate, r	0.03	0.03	0.03	-0.083	-0.085
wage, w	1	0.697	0.770	0.629	0.722
measured TFP	1	0.893	0.969	0.868	0.949
mean $z^{1-\gamma}$	1	0.914	1	0.898	1
capital input	1	0.359	0.432	0.277	0.364
Output	1	0.681	0.755	0.626	0.717
Output minus inn. inv.	1	0.682	0.755	0.628	0.717
access capital supply	0	384%	393%	0	0
mean firm size	22.8	22.8	22.8	22.8	22.8

efficient scale. The second and third columns in Table 3 report the steady state equilibrium effects for an open economy, in which the interest rate is fixed at $r = 0.03$. Because firms are affected by the borrowing constraints in the (NG) model more than in the (XG) model, the equilibrium wage in the former must fall further to clear the labor market, from 1 to 0.697, as opposed to 0.770 in the (XG) model. Because wages fall so much, the constrained firms are able to hire more labor, and accumulate assets (and grow productivity in the (NG) model) at a faster speed; as a result, the adverse effects of the borrowing constraints are considerably mitigated compared to the partial equilibrium outcomes, but are still significant.

In the steady state equilibrium of the open economy, the measured TFP falls by 11% in the (NG) model, and by only 3% in the (XG) model. The small effect in the (XG) model is consistent with the findings in preceding literature (e.g. Midrigan and Xu (2014)) that borrowing constraints lead to small losses in TFP due to the misallocation margin alone – and, by design, this is the only margin present in the (XG) model. In the (NG) model, the decline in measured TFP occurs for two reasons, due to misallocation of resources across firms with different productivity levels, and due to lower investment in productivity.[28] Had the resources been allocated efficiently across firms, the measured TFP would be equal to the average of $z^{1-\gamma}$, which falls by 8.6%. Thus, more than three quarters of a decline in TFP in the (NG) model is due to a reduction in innovative investment.[29]

The firm owners in (NG) model have lower productivity, receive smaller profits, and thus accumulate capital at a slower pace than in the (XG) model. As a result, self-financed firms in the (NG) model use less capital than the self-financed firms in the (XG) model. Together with the bigger TFP loss in the (NG) model, this leads to a bigger drop in steady state output, 32% in the (NG) model compared to 24% in the (XG) model. Recall that the models are calibrated to generate identical levels of output net of innovative investment in the benchmark scenario, thus one would also want to compare how this variable is affected by borrowing constraints. Table 3 reports that the output net of innovative investment falls by more in the (NG) model than in the (XG) model, 32% vs. 24%.[30]

In columns 2 and 3 of Table 3 (open economy), the interest rate does not adjust after the borrowing constraints are introduced, and thus the allocation is characterized by a large excess supply of capital.[31] In a closed economy, the interest rate would have to fall to clear the credit market. The constrained firm owners accumulate wealth, become relatively rich and, once the productive idea disappears, hold a lot of assets. Since, due to self-financing constraints, other firm owners cannot borrow from these agents to finance their firm operations, all this excess wealth is lent to the consumers who borrow to finance personal consumption (as in a standard Bewley model). Thus, the interest rates must fall substantially for the asset market to clear. The

[28]Midrigan and Xu (2014) report about a 17$\%$ drop in measured TFP in the open economy due to self-financing borrowing constraint. However, most of this decline (about three quarters of it) occurs at the extensive margin, because the number of operating firms goes down. In contrast, in the model in this paper, the comparable in size drop in measured TFP arises along the intensive margin only, once the endogenous productivity growth is taken into account.

[29]The losses due to misallocation alone are slightly smaller in the (NG) model than in the (XG) model, 2% vs. 3%, because firms with more assets can afford to make more innovative investment and become more productive. This increases the correlation between the amount of inputs used by the firms and their productivity levels, thereby reducing the extent of misallocation.

[30]The percentage change in the (NG) model is almost identical to that of the percentage change of the output because the share of innovative investment in output is barely affected after the borrowing constraint is introduced.

[31]Due to self-financing, firm owners rely on their own savings and cannot borrow from those who save for precautionary reasons or those who owned firms in the past and have accumulated large amounts of capital.

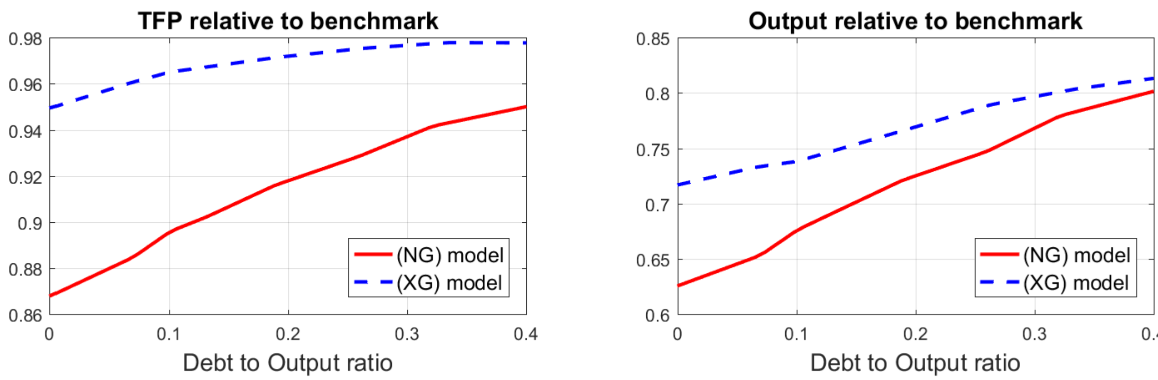

Figure 2: The effects of relaxing the borrowing constraint (namely, varying b_0 in $k \leq a - \underline{a} + b_0$) on TFP and GDP in the (NG) and (XG) models.

last two columns in Table 3 report the stationary equilibrium effects of the borrowing constraints when the asset market clears (closed economy).

The equilibrium interest rate falls by slightly less in the (NG) model than in the (XG) model because, for a given interest rate, the firm owners in the (NG) model are less profitable and accumulate less assets. Because the firm owners receiving new ideas hold less assets, the labor demand decreases and, therefore, the equilibrium wages must fall further than in the open economy. The wage is still considerably lower in the (NG) model than in the (XG) model, for the reasons discussed above. Likewise, the measured TFP, capital, and output decline compared to their open economy counterparts. Quantitatively, the effects are still much bigger in the model with endogenous productivity growth: measured TFP falls by 13% in the (NG) model compared to the 5% in the (XG) model, and output falls by 37% in the (NG) model compared to 28% in the (XG) model.

Figure 2 illustrates the equilibrium effects of the borrowing constraints on TFP and GDP per capita in the (XG) and (NG) economies if some borrowing is allowed (the benchmark TFP and GDP values are normalized to 1). In particular, I assume that $k \leq a - \underline{a} + b_0$ is imposed, and vary b_0. Increasing b_0 leads to higher debt to GDP ratio, and, simultaneously, raises TFP and GDP per capita.[32] As can be seen, the drop on equilibrium TFP in the (NG) model remains

[32]For the (NG) economy, the effects on GDP and GDP net of total innovative investments are quantitatively very similar, only the former is plotted on Figure 2 because the benchmark calibration equalizes its value in the (NG) model with the value of GDP in the (XG) model.

to be roughly two and a half times as big as the corresponding drop in the (XG) model as b_0 rises (conditional on leading to the same debt-to-GDP ratio). The gap between the GDP effects gets somewhat smaller as more borrowing is occurring: it falls from 33% to 20% and then to 10% and the debt-to-GDP ratio rises from 0 (self-financing) to 0.2 and then to 0.35.

The quantitative results summarized above suggest that accounting for firms' endogenous productivity growth significantly magnifies the adverse effects of borrowing constraints on economic activity. One of the predictions of this quantitative analysis, however, may appear questionable. Despite the large drop in equilibrium wages, the equilibrium employment remains unchanged (as reported in the last row of Table 3). This is because the labor supply in the model is inelastic (every agent supplies one unit of labor). Had the labor supply been elastic and upward sloping, the equilibrium wages would not decrease as much, and the effects of borrowing constraints on TFP and GDP would be more pronounced. Additionally, due to inelastic labor supply and the fact that all agents receiving productive ideas choose to operate them, the average firm size is pinned down by the exogenous parameters guiding the arrival and death of productive ideas, and thus is not affected by the borrowing constraints. It is well known, however, that the average firm size is negatively correlated with GDP and excess to external financing. To address these shortcomings, the next Section extends the model by introducing a traditional sector that offers the workers an alternative production opportunity which becomes valuable when the wage rate falls.

6 An Extension: a Two-Sector Model

Suppose that, in addition to receiving productive ideas described in the previous Section – for brevity, let us call them *modern* ideas – the agents may also get *traditional* ideas which they can operate if they forego employment in the modern sector. In contrast to modern ideas, such traditional ideas do not require any input in production except for the labor services of their owners, and their productivity does not grow over time. Intuitively, these ideas stand for routine labor-intensive activities that exhibit little productivity growth like house cleaning, child or senior care, shoe shining, low-tech farming, etc. Agents choose to undertake these activities because their wages in the modern productive sector are low.

Formally, assume that if an agent does not operate a modern idea (either because he abandoned or lost one, and has not drawn any in the beginning of the period), he draws a traditional idea of quality $h_t \sim F_T(h)$, i.i.d. across agents, with probability η_T. If a traditional idea is drawn, the agent chooses whether to pursue it or not.[33] If no ideas are pursued, the agent works in a modern firm. The labor productivity shock ξ_t is realized after all the decisions regarding idea operations are made.[34] Traditional ideas disappear with probability χ_T in the end of the period. If an existing traditional idea does not disappear, its quality remains the same in the next period. As mentioned earlier, an agent must abandon a modern idea in order to draw a traditional idea, but does not have to abandon an existing traditional idea in order to have a chance to draw a modern one; only one idea at a time can be operated though.[35]

The possibility of operating traditional ideas leads to some modifications in the recursive setup of the agents' decision problems. As in the previous Section, denote by $V_W(a)$ the expected value of the agent pursuing no productive ideas in the current period and by $V_M(a, z)$ the expected value of the agent pursuing a modern idea of quality z, with a being the asset level in the beginning of the period. Additionally, let $V_T(a, h)$ be the expected value of the agent pursuing a traditional idea of quality h. Denote also by $V_0(a)$ and $V_{00}(a)$ the auxiliary value functions corresponding to the value of the agent who has no ideas in the beginning of the period and the value of the agent who starts without ideas and draws no modern ideas (but still has a chance to draw a traditional idea). Then the expected value of the worker $V_W(a)$ is defined as in (21) in the previous Section. Likewise, the expected value of operating a modern firm $V_M(a, z)$ is still given by (22) and (23), depending on whether firm productivity growth is exogenous or endogenous. The auxiliary value functions satisfy

$$V_0(a) = (1 - \eta_M)V_{00}(a) + \eta_M \max\left\{V_{00}(a), \int V_M(a, z)dF_M(z)\right\} \tag{26}$$

[33]It is important that the quality of the traditional idea becomes known before the choice of whether to pursue it is made. It implies that if wage decreases, more traditional ideas would be taken, resulting in elastic labor supply in the modern sector (which the the sole purpose for including traditional ideas in the model).

[34]This assumption plays no role, but slightly simplifies the calibration process because the total labor supply in the modern sector is equal to the mass of agents not pursuing traditional firms scaled by $\int \xi \, dF_W(\xi)$.

[35]The requirement that modern ideas must be abandoned before traditional ideas can be drawn is imposed in order to make the problem more computationally manageable; otherwise the owners of the modern ideas would have three state variables, a, z and h, which would have made the already demanding computation even more time-consuming. Assuming that modern ideas can be drawn by the owners of traditional ideas ensures that some traditional ideas are operated for more than one period, and their resulting age distribution is non-degenerate.

and

$$V_{00}(a) = (1-\eta_T)V_W(a) + \eta_T \int \max\{V_W(a), V_T(a,h)\}dF_T(h). \tag{27}$$

The value of operating a traditional firm of quality h is determined as

$$\begin{aligned} V_T(a,h) = \max_{a' \geq \underline{a}} \Big\{ & u(h + a - \frac{a'}{1+r}) + \beta \cdot \chi_T V_0(a') \\ & + \beta \cdot (1-\chi_T) \cdot (1-\eta_M) \max\{V_T(h,a), V_0(a')\} \\ & + \beta \cdot (1-\chi_T) \cdot \eta_M \max \Big\{ V_T(a',h), \int V_M(a',z) dF_M(z) \Big\} \Big\} \end{aligned} \tag{28}$$

In a stationary equilibrium, the optimal policies induce the distributions over agents' asset levels, idea ownership and idea quality. Denote by $\mu_W(a)$, $\mu_M(a,z)$, $\mu_T(a,h)$ the stationary distributions across assets and idea productivities for the workers, the modern firm owners and the traditional firm owners, respectively. Then in a competitive equilibrium prices w and r clear the labor and capital market:

$$\int n(a,z)\, d\mu_M(a,z) = 1 - \int d\mu_T(a,h) \tag{29}$$

$$\int k(a,z)\, d\mu_M(a,z) = \int a\, d\mu_W(a) + \int a\, d\mu_M(a,z) + \int a\, d\mu_T(a,h), \tag{30}$$

where $n(a,z)$ and $k(a,z)$ represent the labor and capital demand of the modern firm of quality z whose owner has asset level a. The key distinction of (29) from (24) is that in (29) only those who do not operate traditional ideas supply labor to the modern sector. Since lower wages in the modern sector induce more agents to take on traditional ideas, this feature of the environment results in the elastic labor supply in the modern sector.

While the terminology of *modern* vs. *traditional* sector is borrowed from Midrigan and Xu (2014), the structure of the environment in this paper is different in some key dimensions. In particular, in Midrigan and Xu (2014) the firm owners choose whether to operate in the productive modern sector or in the unproductive traditional sector, implying that the number of firms in the modern sector is endogenous (and goes down in response to tightening borrowing constraints, which, due to decreasing returns to scale, becomes a major driver of the productivity decline). In contrast, in this model, the number of modern firms is not affected by the borrowing constraints, and the additional effects (compared to the one-sector model) appear because the possibility of pursuing traditional ideas results in the elastic labor supply in the modern sector.

The calibration of this specification of the model involves choosing the parameters guiding the evolution of traditional firms η_T, $F_T(h)$ and χ_T, as well as adjusting some of the previously calibrated parameters so the model matches the same targeted moments as in the one-sector specification. All the parameters fixed outside the model and reported in Table 1 remain unchanged. The parameters in Table 2 are re-calibrated to account for the presence of the traditional firms. Many of them change very little, or remain unaffacted, compared to the one-sector model because in the benchmark unconstrained version of the model the traditional firms contribute very little to the economy's aggregate activity. Table 4 reports the parameters (those that are added due to model extension, and the old ones that are recalibrated), their values, as well as the targeted moments and their counterparts generated by the benchmark model.

The functional form of the distribution of traditional ideas is chosen in such a way that its parameters directly impact the elasticity of the labor supply in the modern sector. In particular, I assume that

$$F_T(h) = Prob(h_t \leq h) = \begin{cases} Ah^{\epsilon}, & Ah^{\epsilon} \leq 1 \\ 1 & \text{otherwise.} \end{cases} \tag{31}$$

Notice that if there were no shocks to labor productivity, and no possibility of arrival of modern ideas, the agents would take productive ideas if and only if they draw h_t higher than the wage rate w, resulting in the labor supply function $N^s(w) = Aw^{\epsilon}$, in which case the labor supply elasticity would be exactly ϵ. In the full model, the uncertainty in labor productivity shocks slightly encourages the uptaking of traditional ideas, but the parameter ϵ still determines how much the amount of workers undertaking production ideas changes in response to a change in wages. I set $A = 1$ and $\epsilon = 0.5$, which falls within the range of estimated levels of macro elasticity of labor supply.[36,37]

The empirical counterpart of traditional firms in the model are the so-called 'necessity entrepreneurs' who, in contrast to 'opportunity entrepreneurs' become business owners not

[36]For review of the literature see, for example, Chetty (2012).

[37]The value of A plays little role in the analysis. It controls how many traditional firms are present in the economy in the benchmark scenario and, via it, slightly effects the values of the calibrated parameters. However, it has little impact on the results from the main experiment. Under $A = 1$, traditional firms account for about ten percent of the firms in the economy, and only for 0.2 percent of the total output in the benchmark unconstrained economy.

Table 4: Parameters calibrated jointly, two-sector (NG) model

Parameter	Value	Target	Data	Model
ϵ, in $F_T(h)$	0.5	labor supply elasticity	0.25-1	0.25
η_M, arrival of modern ideas	0.003	average firm size, BDS data	22.7	22.9
η_T, arrival of trad. ideas	0.041	share of necessity entrepreneurs	0.1	0.1
χ_T, exit rate of trad. firms	0.0738	same as χ_M		
γ, returns to scale in (15)	0.77	top 4% earnings share	0.26	0.25
$F_M(z)$, new ideas distr.		BDS data, net of 10% trad		
B, productivity in (17)	1.41	normalized equilibrium wage	1	1
δ_z, depreciation in (17)	0.051	share of int. inv. in output	0.077	0.077
θ, returns to scale in (17)	0.10	average size of 10-15 y.o. firms	15.9	15.3
β, time discount factor	0.9425	equilibrium interest rate	0.03	0.03
$\underline{a}$, borrowing limit	-1	fract. of pop. with negative assets	0.10	0.18

because they have good ideas, but because they do not have good labor market opportunities.[38] According to various estimates,[39] about 10-15$\%$ of entrepreneurs in the U.S. can be classified as such necessity entrepreneurs. In the model, I target the ratio of 0.1. I also set $\chi_T = \chi_M = 0.0738$, though differences in the exit rates can potentially be incorporated into the model in a straightforward way.[40] Once the exit rates χ_M and χ_T are set, the arrival probability η_M of modern ideas is uniquely pinned down to match the average firm size and the share of

[38] Another term used in the literature is 'push' entrepreneurs, as opposed to 'pull' entrepreneurs. Such 'push' or 'necessity' entrepreneurial firms experience very little growth, consistently with the assumptions made in the model about the evolution of the traditional firms.

[39] See, for example, Caliendo and Kritikos (2009) Fairlie and Fossen (2018)

[40] The one-sector model assumes that, contrary to the empirical evidence, the exit rates do not vary with firm size or age. This shortcut that is often taken in the literature to simplify exposition and computations. Since a common exit rates was assumed in the one-sector model, it also makes sense to make the same simplifying assumption in the two-sector model. At this stage, it is hard to speculate how adopting more realistic exit rates would effect the quantitative results.

necessity entrepreneurs among all firms, for a given η_T. The arrival η_T of traditional ideas is, in turn, calibrated jointly with the rest of the model's parameters to ensure that the traditional firms account for 10 percent of all firms in the economy.

The rest of the parameters are disciplined as in Section 4. The only modification is that traditional firms should be accounted for while calibrating the distribution of modern entrant's productivity shocks. Namely, it is assumed that 10% of the BDS entrants are traditional firms of size 1. Then $F_M(z)$ is calibrated to match the entrants' size distribution net of these firms. As mentioned earlier, the rest of the parameters are barely adjusted relative to their one-sector counterparts because in the unconstrained model traditional firms contribute little to economic activity.

To calibrate the two-sector (XG) model, I use exactly the same approach as in Section 4.2. The parameters guiding the evolution of traditional firms remain the same as in the (NG) model because, by design, the evolution of firm sizes and workers' wages is the same in the two models.

Table 5 summarizes how the steady state allocations in the two-sector (NG) and (XG) models adjust after the self-financing constraint is imposed on firms' physical capital. Overall, the negative effects on output and TFP are slightly bigger than in a one-sector model because, due to elastic labor supply, wages do not fall as much. Additionally, the gap between the effects in the (NG) and (XG) models is also a bit bigger than in a one-sector environment. For example, the difference in measured TFP losses in the (NG) and (XG) models is 9.3% in the two-sector model, compared to 8.1% in the one-sector model; while the corresponding differences in GDP effects are 9.9% and 9.1%, respectively.

The large contrast with the one-sector model arises in the effects of borrowing constraints on the average firm size. Because the labor supply is inelastic in the one-sector model, the average firm size is not impacted by the presence of the borrowing constraint. In contrast, in the two-sector model the labor supply reduces as wages drop because some of the potential workers choose to operate traditional firms.[41] As a result, borrowing constraints reduce the total employment in modern firms, while simultaneously increasing the number of traditional firms. In fact, the share of traditional firm rises dramatically from 10% to $63-72\%$, depending on the

[41] The numbers in Table 5 can be used to compute the elasticity of the labor supply in the modern sector implied by the model. For example, in the closed (NG) economy, the labor supply falls by about 9% (from 25.2 per firm to 22.9 per firm), while wages fall by 36%, implying a labor supply elasticity of about 0.25

Table 5: The effects of borrowing constraints in (NG) and (XG) models

	Benchmark	Self-financing			
		open economy		closed economy	
		(NG)	(XG)	(NG)	(XG)
interest rate, r	0.03	0.03	0.03	-0.075	-0.075
wage, w	1	0.698	0.775	0.640	0.734
measured TFP, modern	1	0.886	0.962	0.857	0.95
mean $z^{1-\gamma}$	1	0.908	1	0.887	1
capital input	1	0.340	0.412	0.262	0.35
Output, all	1	0.688	0.764	0.625	0.724
Output, modern	1	0.651	0.732	0.582	0.688
Output minus inn. inv.	1	0.689	0.764	0.627	0.724
access capital supply	0	385%	395%	0	0
mean firm size, all	22.8	8.1	9.5	7.1	8.7
mean firm size, modern	25.2	23.5	23.8	22.9	23.6
std. firm size, modern	75.5	66.1	57.0	78.1	64.3
share of trad. firms	0.1	0.68	0.63	0.72	0.66

model's specification. Since all these traditional firms have only one employee (the owner), the average firm size across all the firms falls dramatically (from 22.8 to 7-10 employees). These quantitative results, of course, depend on the particular distribution function of traditional ideas, and the somewhat ad-hoc assumption that traditional firms have only one employee and do not grow over time. At the same time, they underscore an economic mechanism that can contribute to explaining the well-documented fact that the average firm size is smaller in developing countries,[42] and the quantitative analysis above suggests that accounting for endogenous innovative investments is important for quantifying the implications of this mechanism.

7 Final remarks

This paper has demonstrated that accounting for endogenous innovative investment is important for quantifying the impact of financial constraints on economic development. In particular, endogenous firm productivity growth magnifies the importance of the intensive margin via which borrowing constraints reduce economic activity: in the benchmark formulation of the model measured TFP falls by additional 8% and GDP falls by additional 9% (or, respectively, by 2.6 and 1.3 times more) when firms' productivity growth is *endogenous*. While existing studies have argued that, via the intensive margin alone, firm financing constraints have little effect on measured TFP and output, this paper demonstrates that accounting for endogeneity of firm productivity growth considerably amplifies this impact, and the intensive margin becomes a lot more significant. The key driving force behind the effect on TFP is that, in addition to misallocation emphasized in previous papers, firm financing constraints also reduce the average firm productivity.

The model, by design, focuses on the role of intensive margin only. As a result, the calculations in this paper provide a *lower bound* on the additional effects arising due to the endogeneity of firm productivity growth. The extensive margin present in the two-sector extension of the model (the endogenous number of workers employed in the modern sector) indeed amplifies the additional effects arising due to endogenous innovative investment. These effects would be even bigger had another commonly emphasized extensive margin were present: if the modern firm owners had to forego wage earnings to operate own firms, borrowing constraints would discourage more agents from starting their firms if the firm productivity growth is impeded by financial constraints.

The two-sector extension of the model also underlines a potentially important mechanism which can contribute to explaining why firm size is positively correlated with economic development. Notice also that because traditional firms exhibit no productivity growth, a dramatic increase in the number of traditional firms due to financial constraints also flattens the firm size

[42] See, for example, Hsieh and Klenow (2014).

vs. age profile. Thus, a two-sector version of the model can also contribute to explaining the firm life-cycle observations documented in Hsieh and Klenow (2014) (that the firms grow considerably slower over their life-cycle in less developed economies). Note that the one-sector version of the model actually runs counter with these observations because borrowing constraints reduce firm size in the beginning of life, but become irrelevant as the firms age because firm owners are able to accumulate sufficient amount of assets over time.

On a related note, incorporating extensive margin into the model, as in Buera, Kaboski and Shin (2011) (by assuming that firm owners must forego wage income in the spirit of the span-of-control model in Lucas (1978)), would help produce the positive correlation between economic development and firm size because of the general equilibrium effects: borrowing constraints reduce equilibrium wages and, therefore, encourage more entrepreneurs, potentially exhibiting lower productivity, to enter into the business. This effect would be bigger in a model in which firm productivity growth in endogenous because the effects on equilibrium wage are bigger. Note, however, that such a model with extensive margin would fail to reproduce the differences in firm life-cycle patterns documented in Hsieh and Klenow (2014), due to the reasons outlined above. In contrast, the model with the slow growing 'push entrepreneurs' outlined in Section 6 describes an economic mechanism that can contribute to explaining these patterns.

8 Appendix

8.1 Proof of Proposition 1

To verify (14), use the relationship between z_1 and k_1 in (10) to obtain

$$\frac{\partial \widehat{\Pi}^{NG}(\widehat{k_1})}{\partial \widehat{k_1}} = \nu_1 \widehat{z_1}^{\varphi_1} \widehat{k_1}^{\nu_1 - 1} - R = \nu_1 \left(\frac{\varphi_1}{1 - \beta\Omega} \right)^{\frac{\varphi_1}{1-\varphi_1}} \widehat{k_1}^{-\frac{1-\varphi_1-\nu_1}{1-\varphi_1}} - R. \tag{32}$$

Next, using calibration conditions (12),

$$\frac{\partial \widehat{\Pi}^{XG}(\widehat{k_1})}{\partial \widehat{k_1}} = \gamma_1 A_1 \widehat{k_1}^{\gamma_1 - 1} - R = R \left[\left(\frac{\nu_1}{R} \right)^{1-\varphi_1} \cdot \left(\frac{\varphi_1}{1 - \beta\Omega} \right)^{\varphi_1} \right]^{\frac{1-\gamma_1}{1-\varphi_1-\nu_1}} \widehat{k_1}^{\gamma_1 - 1} - R. \tag{33}$$

Plugging γ_1 from (12) and using (10), it is straightforward to verify that the expression in (32) exceeds the expression in (33) for all $\widehat{k_1} < k_1^{NG}$ as long as $\beta > 0$. QED

Table 6: Size distribution of age 1 firms, own calculations on BDS data

N. of workers	1-4	5-9	10-19	20-49	50-99	100-249	250-499	500-999	≥ 1000
fract. of firms	0.7269	0.1506	0.0711	0.0365	0.0094	0.0042	0.0010	0.0003	0.0001

8.2 Computing the equilibrium

The decision problem of the firm owner in the (NG) model (23) has two continuous state variables and two continuous control variables. This makes the problem demanding from the computational point of view, particularly at the calibration stage, where seven parameters are iterated on to match the targeted moments. Thus, $V_M(a, z)$ is computed on the sparse grid for asset and productivity levels (with more dense grid points at lower values). At every step of iterations, this value os interpolated on a dense grid, and the next period's (a', z') are chosen from this dense grid. Once the policy functions are computed, the steady state of the economy is found by simulating three million individuals. Since there are very few firm owners, their individual characteristics (e.g., the average size of entrants) are computed by averaging over 30 periods of the simulated economy in the steady state.

8.3 Calibration details

8.3.1 BDS data

To compute all the facts related to the dynamics of modern firms, I use publicly available BDS data, available at https://www.census.gov/ces/dataproducts/bds/data.html. I use the averages across 1998-2006 to construct all the relevant statistics. To parameterize the distribution $F_M(z)$ from which new modern ideas are drawn, I use the distribution of age 1 firms reported in Table 6; to derive these numbers one needs to use the BDS data by *Age* and *Initial Size* (as opposed to just *Size* which averages the firms' sizes across two consecutive years).

8.3.2 Labor productivity process

Given the model's specification, the variance of log earnings is

$$\begin{aligned}\sigma_\xi^2 &= (1-p)\cdot(\log\xi_L - (p\log\xi_H + (1-p)\log\xi_L))^2 \\ &\quad + p\cdot(\log\xi_H - (p\log\xi_H + (1-p)\log\xi_L))^2 \\ &= p(1-p)\log\left(\frac{\xi_H}{\xi_L}\right)^2.\end{aligned}$$

This, together with $p\xi_H + (1-p)\xi_L$ implies that

$$\begin{aligned}\xi_L &= \left(1-p+p\cdot\exp\left(\frac{\sigma_\xi}{\sqrt{p(1-p)}}\right)\right)^{-1} \\ \xi_H &= \exp\left(\frac{\sigma_\xi}{\sqrt{p(1-p)}}\right)\cdot\xi_L\end{aligned}$$

Plugging $p = 0.5$ and $\sigma_\xi = \sqrt{0.045}$ implies that $\xi_H = 1.209$ and $\xi_L = 0.791$.

References

Bento, Pedro, and Diego Restuccia. 2017. "Misallocation, Establishment Size, and Productivity." *American Economic Journal: Macroeconomics*, 9(3): 267–303.

Bhattacharyaa, Dhritiman, Nezih Guner, and Gustavo Ventura. 2013. "Distortions, endogenous managerial skills and productivity differences." *Review of Economic Dynamics*, 16: 11–25.

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

Buera, Francisco, Joe Kaboski, and Yongseok Shin. 2015. "Entrepreneurship and Financial Frictions: A Macrodevelopment Perspective." *Annual Review of Economics*, 7.

Caggese, Andrea. 2018. "Financing Constraints, Radical versus Incremental Innovation, and Aggregate Productivity." *American Economic Journal: Macroeconomics*, forthcoming.

Caliendo, Marco, and Aleksander Kritikos. 2009. "I Want to, But I Also Need to: Start-Ups Resulting from Opportunity and Necessity." *IZA Institute of Labor Economics Discussion Paper Series*, 4661.

Chetty, Raj. 2012. "Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." *Econometrica*, 80: 9691018.

Corrado, Carol, Jonathan Haskel, Cecilia Jona-Lasinio, and Massimiliano Iommi. 2018. "Intangible investment in the EU and US before and since the Great Recession and its contribution to productivity growth." 2.

Daz-Gimnez, Javier, Andy Glover, and Jos-Vctor Ros-Rull. 2011. "Facts on the Distributions of Earnings, Income, and Wealth in the United States: 2007 Update." *Federal Reserve Bank of Minneapolis Quarterly Revie*, 34.

Fairlie, Robert W., and Frank M. Fossen. 2018. "Opportunity versus Necessity Entrepreneurship: Two Components of Business Creation." *IZA Institute of Labor Economics Discussion Paper Series*, 11258.

Ferrando, Annalisa, and Alessandro Ruggieri. 2015. "Financial constraints and productivity: evidence from euro area companies." *European Central Bank Working Paper Series*, 1823.

Gorodnichenko, Yuriy, and Monika Schnitzer. 2013. "Financial Constraints and Innovation: Why Poor Countries Dont Catch Up." *Journal of the European Economic Association*, 11(5): 11151152.

Greenwood, Jeremy, Juan M. Sanchez, and Cheng Wang. 2013. "Quantifying the impact of financial development on economic development." *Review of Economic Dynamics*, 16(1): 194 – 215. Special issue: Misallocation and Productivity.

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

Hugget, Mark. 1993. "The risk-free rate in heterogeneous-agent incomplete-insurance economies." *Journal of economic Dynamics and Control*, 17: 953–969.

Hugget, Mark. 1996. “Wealth distribution in life-cycle economies.” *Journal of Monetary Economics*, 38: 469–494.

Levine, Oliver, and Missaka Warusawitharana. 2016. “Finance and Productivity Growth: Firm-Level Evidence.” *mimeo*.

Lucas, Robert. 1978. “On the Size Distribution of Business Firms.” *The Bell Journal of Economics*, 9(2): 508–523.

Manaresi, Francesco, and Nicola Pierri. 2017. “Credit Supply and Productivity Growth.” *mimeo*.

McGrattan, Ellen R., and Edward C. Prescott. 2010. “Unmeasured Investment and the Puzzling US Boom in the 1990s.” *American Economic Journal: Macroeconomics*, 2(4): 88–123.

Mestieri, Mart, Johanna Schauer, and Robert M. Townsend. 2017. “Human capital acquisition and occupational choice: Implications for economic development.” *Review of Economic Dynamics*, 25: 151 – 186. Special Issue on Human Capital and Inequality.

Midrigan, Virgiliu, and Daniel Yi Xu. 2014. “Finance and Misallocation: Evidence from Plant-Level Data.” *American Economic Review*, 104: 422–58.

Moll, Benjamin. 2014. “Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?” *American Economic Review*, 104(10): 3186–3221.

Piketty, Thomas, and Emmanuel Saez. 2003. “Income Inequality in the United States, 19131998*.” *The Quarterly Journal of Economics*, 118(1): 1–41.

Quadrini, Vincenzo. 2000. “Entrepreneurship, Saving, and Social Mobility.” *Review of Economic Dynamics*, 3(1): 1 – 40.