Discussion Paper No. 1033

SIZE-DEPENDENT POLICIES AND EFFICIENT FIRM CREATION

Sakai Ando

June 2018

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Size-Dependent Policies and Efficient Firm Creation*

Sakai Ando†

June 27, 2018

Abstract

I study the welfare implications of size-dependent firm regulation policies (SDPs) in the presence of entrepreneurial risks. Although SDP has been considered a source of misallocation, I show that, once entrepreneurial risks are taken into account, SDP can improve efficiency. Quantitatively, I show that, based on French data, removing the SDP leads to output and welfare loss by 1.5% and 1.3%, respectively, in opposition to the output gain reported by the previous literature that abstracts from risks. Qualitatively, I solve an optimal non-linear SDP problem and show that the observed SDP shares certain features with the optimal SDP. The analysis uncovers a novel trade-off between the inefficiencies of the intensive and extensive margins. In extension, it is shown that (1) whether SDPs improve efficiency depends on the level of financial development and (2) capital accumulation and consumption-smoothing motive further justify SDPs.

JEL: D52, D61, H21, J08, L11, L24, L51

Keywords: misallocation, firm creation, size-dependent policy, financial development, incomplete markets

*I thank Yu Awaya, Tania Babina, Pierre-Andre Chiappori, Andres Drenik, Matthieu Gomez, Tomohiro Hirano, Harrison Hong, Takatoshi Ito, Atsushi Kajii, Jennifer La'O, Misaki Matsumura, Christian Moser, Andreas Mueller, Emi Nakamura, Yuji Sakurai, Jose Scheinkman, Stephanie Schmitt-Grohe, Jon Steinsson, Martin Uribe, Michael Woodford, Makoto Yano, Pierre Yared, and all participants in the seminars at Australian National University, Bank of Japan, City University of Hong Kong, Columbia University, Hitotsubashi University, Kyoto University, Melbourne University, National Taiwan University, Peking University, Purdue University, University of Tokyo, and Waseda University. This paper is awarded the 18th Moriguchi prize from Osaka University Institute of Social and Economic Research.

†Email: sa3016@columbia.edu.

1 Introduction

Size-dependent policies (SDPs) that preferentially treat small firms are ubiquitous. For instance, in France, firms that hire more than 50 workers have to pay additional regulatory costs, such as higher tax rates and more stringent labor regulations. Such SDPs naturally create the bunching behavior in the firm size distribution as in Fig.1, since some big firms rationally remain small to save regulatory costs. The literature has focused on the misallocation from the bunching behavior, reporting that the removal of SDPs leads to output gain by $.02 \sim 4\%$ (Gourio and Roys [2014], Garicano et al. [2016]).

In this paper, I argue that our understanding of SDPs can change drastically if we take into account the fact that firm creation is risky. Specifically, I show that, once the analysis incorporates uninsurable entrepreneurial risks, SDPs can generate higher output and welfare, contributing to net efficiency gains despite the remarkable bunching behavior. The key observation is that, when entrepreneurial risks cause insufficient firm creation relative to the Pareto efficient level, SDPs can mitigate the market failure by supporting firm creation. Thus, SDPs balance the trade-off between the inefficiencies in the intensive (bunching) and extensive margins (insufficient firm creation). This is the novel trade-off that this paper uncovers and is at the heart of the following quantitative and qualitative analyses that generate the opposite policy implications to the previous literature.

The key assumption is that entrepreneurial risks deter firm creation. This assumption is not only intuitive but also has been confirmed by several empirical studies. For instance, Hombert et al. [2014] study the French unemployment insurance reform in 2002 and find a positive effect of insuring the downside risks of entrepreneurship. Gottlieb et al. [2018] argue that the Canadian maternity leave policy reform mitigates the downside risks of losing safe job options and increases entrepreneurship. In addition to these causal studies, we can also observe the cross-sectional negative correlation between the downside risks and firm creation as in Fig.2. Thus, in this paper, I take the assumption as given and focus on its implications.

To illustrate the trade-off and quantify the welfare consequence of SDPs, I extend the standard occupation choice model a la Lucas [1978] and more directly Garicano et al. [2016] by adding entrepreneurial risks. Specifically, the model has three features. First, firms' production technology exhibits decreasing returns to scale. This is the same trick as the span of control model a la Lucas [1978] and generates a well-defined number of firms in equilibrium, the concept necessary to discuss the efficiency of firm creation. Second and most importantly, agents face uninsurable idiosyncratic entrepreneurial risks in the sense that, when agents choose to become either entrepreneurs or workers, they are not sure about their entrepreneurial productivities. Such downside risks lead to insufficient firm creation

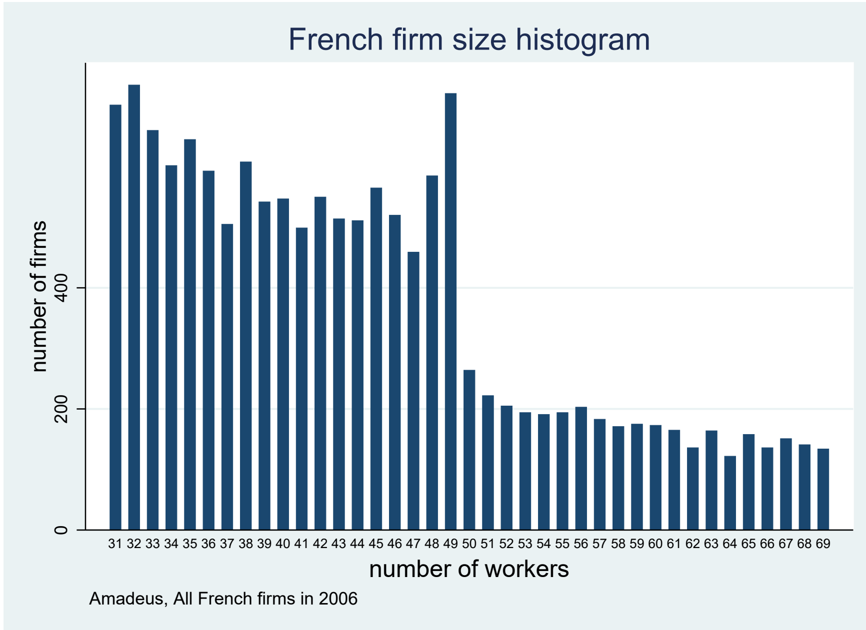

Figure 1: The number of firms by employment size in France. The rise and drop at the employment size of 50 reflect the endogenous reaction of firms to the regulation threshold 50, above which firms are subject to higher tax and heavier labor regulations. The sample includes all the firms in Amadeus 2006 with employment size between 31 and 69.

as a result of incomplete markets, opening up the scope for policy interventions. Finally, entrepreneurs face SDPs modeled as the tax payment based on the number of employees. This formulation is the same as Garicano et al. [2016] and parsimoniously creates the bunching behavior by treating small firms preferentially.

The mechanism by which SDPs improve efficiency is based on the two general equilibrium channels through which SDPs facilitate firm creation. The first channel is market clearing. Specifically, since big firms pay regulatory costs under the SDPs, they hire fewer employees. The lower labor demand from big firms implies more firm creation since agents are either entrepreneurs or workers. In other words, SDPs prevent big firms from absorbing human resources that could otherwise be devoted to entrepreneurship. The second channel is risk mitigation through wage decline. In particular, since big firms make fewer profits, entrepreneurship is less attractive. To offset the resulting increase in the supply of salary workers and reflect weaker labor demand from big firms, wage declines. This is good news for small firms because they do not face regulations and can enjoy lower labor costs. As a consequence, the profitability of small firms improves. From the ex-ante point of view, this means that entrepreneurs face lower downside risks since, even when businesses end up being small or unprofitable, they can earn more than in the laissez-faire economy. As a result of these channels, the economy with SDPs features more firm creation.

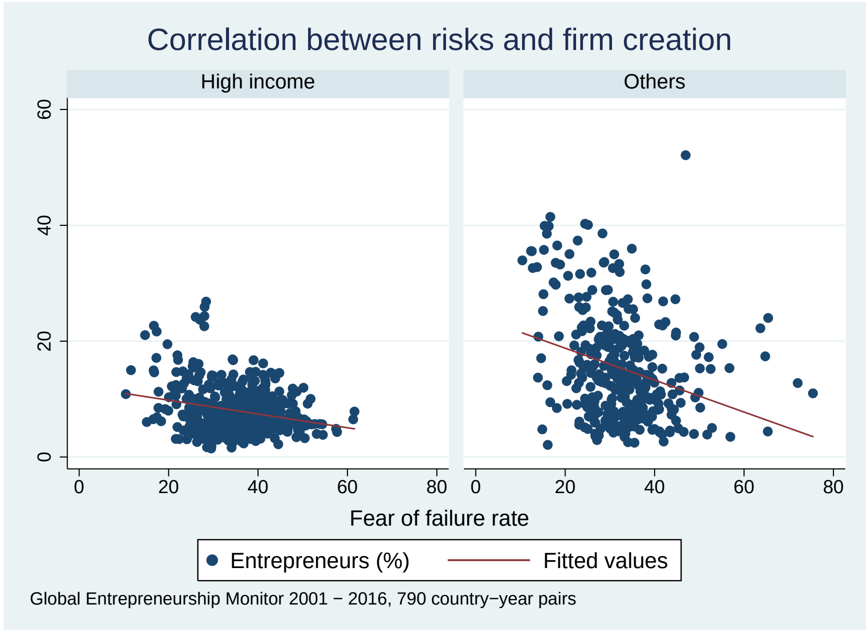

Figure 2: A negative correlation between fear of failure rate and entrepreneurship. Fear of failure rate is the percentage of $18-64$ population perceiving good opportunities to start a business who indicate that fear of failure would prevent them from setting up a business. Entrepreneurship is the percentage of the population of $18-65$ years old who is either a nascent entrepreneur or owner-manager of a new business. The plot pools yearly data $2001-2016$ for all available countries from 28 to 76 depending on years. The definition of high-income countries and others follow the categorization of the World Bank. Data: Global Entrepreneurship Monitor.

The increase in the number of firms can raise output and welfare in the presence of entrepreneurial risks. Intuitively, in the laissez-faire economy with entrepreneurial risks, marginal entrepreneurs produce and consume more on average for the compensation of risk takings. Therefore, SDPs that change the occupation of agents from workers to entrepreneurs can generate positive net output gain, as long as the output loss from bunching is not too large. Accordingly, the higher output and the lower entrepreneurial risks can Pareto improve the laissez-faire economy. Whether the efficiency gain from these channels outweighs the efficiency loss of bunching is a non-trivial question and therefore necessitates the quantification exercise.

To quantify the welfare impact, I conduct the same counterfactual analysis as Garicano et al. [2016] except that my model has entrepreneurial risks. Importantly, I use the same specification and the same French data so that I can isolate the pure implication of entrepreneurial risks. The counterfactual analysis reveals that removing SDPs decreases output by 1.5% and welfare in wage unit by 1.27%. The result is in sharp contrast to Garicano et al. [2016] that report that removing SDPs increases output by $.02 \sim 4\%$ in risk-free settings.

In terms of implication, the result not only uncovers the novel trade-off behind SDPs but also poses a caveat to the misallocation measurement literature in general including Hsieh and Klenow [2009]. Specifically, misallocation measurement exercises typically assume marginal product equalization in the efficient benchmark and measure the deviation from it. My result implies that using such benchmark can be misleading since it might be better to have some wedges among marginal products when the economy has uninsurable entrepreneurial risks. In other words, the trade-off between the inefficiencies of the intensive and extensive margins uncovered in the analysis highlights a limitation of the core logic of the typical misallocation measurement that focuses only on the former.

To explore the novel trade-off, I study the non-linear SDP that optimally creates wedges among firms. More concretely, I extend the piecewise linear taxation schedule used in the quantitative exercise to the set of arbitrary differentiable functions and solve the optimal SDP problem using mechanism design techniques. The main result is that the optimal SDP and the observed threshold type SDP share qualitative features; both of them (1) distort medium-sized firms more than the smallest and biggest and (2) feature a larger number of firms compared to the laissez-faire economy. In addition to these analytical results, I also show that, calibrated at the same parameter values as the quantification exercise, the optimal SDP subsidizes small firms and taxes big firms. Thus, the observed SDP has efficiency-enhancing properties.

Once the novel trade-off is established, I extend my analysis by examining its generality. A natural question is whether entrepreneurial risks always justify SDPs. I argue that the answer is negative, especially when the laissez-faire economy creates an excessive number of firms. As an example of excessive firm creation, I study the financial frictions that prevent firms from expanding to their profit-maximizing sizes.

It is shown that removing SDPs improves efficiency even when there are entrepreneurial risks if existing firms face severe financial frictions. To be concrete, one can imagine a developing country in which firms cannot expand their employment due to financial frictions. Since firms cannot hire people, those unemployed have to do businesses on their own. As a result, one can observe many small businesses in the economy including people engaged in food truck businesses on the streets. I model such intuition and show that severe financial frictions can lead to too many firms compared to the Pareto efficient level. In this situation, SDPs should be removed since the laissez-faire economy already generates an excessive number of firms in the first place. In fact, the efficiency gains are doubled under excessive firm creation since removing SDPs fixes both the intensive and extensive margins.

The analysis of financial frictions implies that whether SDPs improve efficiency depends on the level of financial development. In financially developed countries where entrepreneurial risks cause insufficient firm creation, SDPs should be kept, while in countries with immature financial infrastructure and therefore excessive firm creation, SDPs should be removed.

Finally, I investigate the implications of capital accumulation. To isolate the implication of inter-temporal decisions in a tractable dynamic environment, I extend the static model by introducing Kreps-Porteus-Epstein-Zin preference (Kreps and Porteus [1978], Epstein and Zin [1989]) and capital accumulation a la Krebs [2003]. These modeling techniques allow me to derive a closed form solution despite the heterogeneous-agent incomplete market dynamic model.

It is shown that consumption-smoothing motive provides a further justification for keeping SDPs. In particular, I show that the laissez-faire economy creates insufficient firms, even if agents are risk neutral. The idea is that, when agents prefer smooth-consumption, they want to avoid asset fluctuation, so they do not want to take entrepreneurial risks. Such dynamic trade-off discourages firm creation and therefore justifies SDPs. This result not only confirms the robustness of the analysis in the static model but also highlights a force specific to the dynamic environment.

An implication of the dynamic extension is that SDPs should be kept in an economy with patient agents. This is because patient agents value future consumption more and therefore have stronger consumption-smoothing motives. Theoretically, this reflects the observation that consumption-smoothing motive is controlled not just by the inter-temporal elasticity of substitution but also by the discount factor. Thus, other things being equal, in countries with patient agents such as ones with the culture of patience or long life-expectancy, it is more likely that removing SDPs exacerbates the market failure of firm creation.

1.1 Literature

This paper contributes to several strands of literature. First, it provides a novel insight into the conventional wisdom about SDPs. Restuccia and Rogerson [2008], Guner et al. [2008] and Garicano et al. [2016] measure the output gain from removing SDPs. While the details of the models are different across those papers, they share the same feature that the laissez-faire economy without SDPs is efficient. Hence, SDPs decrease output by construction. My contribution is to show that the introduction of entrepreneurial risks might justify SDPs, altering the understanding of SDPs being bad. Another important paper in this strand of literature is Gourio and Roys [2014], which conduct a counterfactual analysis using a firm dynamics model with risky TFP and entry cost. Although they obtain an output loss from removing SDPs, the aggregate consumption and welfare increase, so their policy implication is still to remove SDPs. In contrast, I show a case in which SDPs should be kept.

More broadly, my paper poses a caveat to the misallocation measurement exercises based on marginal product equalization (Hsieh and Klenow [2009], Hsieh et al. [2013]). In particular, I show that, with uninsurable entrepreneurial risks, there is a trade-off between the efficiencies of the intensive and extensive margins, so the efficient benchmark does not necessarily feature marginal product equalization.

The second strand of literature that I contribute studies the efficiency of firm formation. Kihlstrom and Laffont [1979] and Kanbur [1981] are early references for risky firm formation. These papers assume both risky occupation choice and non-contingent labor choice so the resulting market failure is not specific to the friction due to risks. To the best of my knowledge, my paper is the first to study the market failure specific to risky occupation choice. For other environments, Mankiw and Whinston [1986] and Suzumura and Kiyono [1987] are early references that study the optimality of the firm entry in strategic settings. Jaef [2012] studies the optimality of entry in the firm dynamics model of Hopenhayn [1992]. All of them emphasize excessive entry, so the results are the opposite of mine.

The third strand of literature that this paper contributes to is the one on financial constraints and entrepreneurship as surveyed by Quadrini [2008] and Buera et al. [2015]. The literature has studied the impact of financial frictions on the market equilibrium but has not studied the impact of financial frictions on the efficiency, i.e., the difference between market equilibrium and efficient allocation. I study both in a simple model with a closed form solution.

Finally, my paper contributes to dynamic macroeconomics modeling by offering a tractable dynamic general equilibrium model with heterogeneous-agent, incomplete markets and risky occupation choice. The trick behind the tractability is from Krebs [2003]. Toda [2015] and Gottardi et al. [2016] study the efficiency of the descendants of Krebs [2003]. I extend the framework to include risky occupation choice.

2 Model

In this section, I present the baseline model. After discussing the existence and uniqueness of the equilibrium, I use an efficiency analysis to illustrate the market failure that the SDP mitigates. In particular, I show that the laissez-faire economy without SDP generates an insufficient number of firms if and only if there are entrepreneurial risks.

There is a continuum of ex-ante identical risk-averse agents indexed by $i \in [0, 1]$. Each agent is endowed with one unit of indivisible labor that can be spent in running a firm or working for a firm.

If agent i chooses to be a worker, she receives wage w independent of her entrepreneurial productivity z_i. If she chooses to be an entrepreneur, she observes her entrepreneurial productivity z_i and then decides the size of the firm measured by the number of employees n to maximize the profit

$$[\pi(z_i, w), n(z_i, w)] = \max_{n \geq 0} \begin{cases} z_i f(n) - wn & n \leq N \\ z_i f(n) - w\tau n - F & n > N \end{cases} \tag{1}$$

where f is the production function, $n(z_i, w)$ is the associated policy function and $(\tau, F, N) \in [1, \infty) \times \mathbb{R} \times \mathbb{R}_+$ is the SDP modeled as the variable and fixed tax that entrepreneurs have to pay when they hire more than N workers. Note that once the firm hires more than N workers, the variable cost $\tau \geq 1$ applies to not just the net additional workers $n - N$ but also all the workers n. Therefore, firms have the incentive to shrink the size as in Fig.1 even if the fixed component is negative $F < 0$ as long as the total tax payment $(\tau - 1) wN + F$ is positive.

Given the payoffs of the two occupations, each agent i observes the signal s_i about the entrepreneurial productivity z_i and chooses the occupation that maximizes her expected utility. The joint distribution of the signal and productivity (s, z) is exogenously given and denoted by G. Formally, each agent i solves

$$e(s_i, w) = \arg\max_{e \in [0,1]} e\mathbb{E}[u(\pi(z_i, w)) | s_i] + (1 - e) u(w) \tag{2}$$

where u is the utility function, and $\mathbb{E}[\cdot | s_i]$ is the expectation with respect to the productivity $z_i \in \mathbb{R}_+$ conditional on the observed signal $s_i \in \mathbb{R}$. If the agent chooses $e = 1$, she becomes an entrepreneur. Accordingly, $e = 0$ indicates that she becomes a worker.

I make two observations about the occupation choice problem. First, note that the choice variable e, representing whether to become an entrepreneur, can take a continuous value $e \in [0, 1]$. This formulation allows agents to take mixed strategies when entrepreneurship and salary job are indifferent. Second, the occupation choice is risky because the agents have to decide the occupation before observing the productivity z_i. Since the salary job is risk-free, depending on the realization of the productivity z_i, entrepreneurs might consume less than workers in equilibrium. In this sense, entrepreneurship involves downside risks.

Finally, the wage w clears the labor market

$$1 - \int e(s_i, w)\, di = \int e(s_i, w)\, n(z_i, w)\, di \tag{3}$$

where the left-hand side is the aggregate labor supply and the right-hand side is the aggregate labor demand from firms.

The following definition summarizes the description of the equilibrium as well as other objects of interests.

Definition 1. Fix the fundamentals (u, f, G) and the SDP (τ, F, N). The set of wage, occupation choice, and production decisions $\{w, e, \pi, n\}$ is an equilibrium if it satisfies (1), (2) and (3). The number of firms and the aggregate output in the equilibrium is defined as

$$\phi = \int e(s_i, w)\, di,\ Y = \int e(s_i, w)\, z_i f(n(z_i, w))\, di. \tag{4}$$

The laissez-faire economy is defined as the one with $(\tau, F) = (1, 0)$.

Throughout the paper, I attach LF and SDP to the equilibrium objects whenever the distinction of the laissez-faire economy and the economy with the SDP needs to be made explicit.

To ensure that the equilibrium is well-behaved, I make three assumptions on the fundamentals (u, f, G). First, utility and production functions (u, f) are strictly increasing, strictly concave and satisfy the Inada condition. As will be clear, the strict concavity of the production function ensures the optimal level of firm creation. Second, the joint distribution of the signal and productivity $G(s, z)$ is continuous, has bounded productivity $P(0 < z_{min} \leq z \leq z_{max} < \infty) = 1$ for some (z_{min}, z_{max}), and reflects positively informative signals, i.e., the conditional distribution $G(z|s)$ first-order stochastically dominates $G(z|s')$ whenever $s > s'$. The continuity allows the system to adjust continuously, the bounded productivity circumvents the complexity of infinite utility, and the positively informative signal provides a normalization so that a higher signal is associated with a higher productivity. Finally, I assume the signal and productivity $(s_i, z_i)_i$ are drawn *i.i.d.* from G. The *i.i.d.* assumption makes it possible to invoke the law of large numbers.

I call these assumptions (A). Under these assumptions, the existence and the uniqueness can be guaranteed as stated in the next proposition. The proof clarifies the structure of the equilibrium.

Proposition 1. *Suppose that the fundamentals (u, f, G) satisfy the assumptions (A). An equilibrium exists and is unique almost surely.*

Proof. Given the signal structure, the occupation choice follows a threshold strategy $e(s_i) = 1_{s_i \geq \bar{s}}$. The "almost surely" qualification reflects the indeterminacy of optimal occupation for the marginal entrepreneurs $\{i : s_i = \bar{s}\}$. Given the individual occupation choice $e(\cdot)$, the market equilibrium can be characterized by the equilibrium wage w and the threshold $\bar{s}$ that satisfy the following indifference condition and the market clearing condition

$$\mathbb{E}[u(\pi(z, w))\,|s = \bar{s}] = u(w), \tag{5}$$

$$\mathbb{E}\left[n\left(z,w\right)1_{s\geq\bar{s}}\right]=G_{s}\left(\bar{s}\right). \tag{6}$$

where Eq.(3) invokes the law of large numbers a la Uhlig [1996].

To show the existence and the uniqueness, note that, for a fixed SDP (τ, F, N), the indifference condition (5) specifies a positive relationship between the wage w and the threshold $\bar{s}$, while the market clearing condition (6) provides a negative relationship. Both of these relationships are continuous since $n\left(z,w\right)$ jumps at most at one point. Moreover, according to the market clearing condition, $w\to\infty$ as $G\left(\bar{s}\right)\to 0$, and $w\to 0$ as $G\left(\bar{s}\right)\to 1$. Therefore, the two loci must cross each other once. □

The proof of the existence and the uniqueness highlights the two forces that govern the relationships between the wage w and the number of firms $\phi=1-G_{s}\left(\bar{s}\right)$. One can graphically see them in Fig.3. First, the indifference condition implies that if the wage increases, more people find it attractive to become workers, so the number of firms declines. This is the labor supply side intuition from the individual perspective and generates the negative relationship between the wage and the number of firms. Second, the market clearing condition implies that, if the wage increases, firms' labor demand declines, so there will be fewer workers, ending up with more firms. This labor demand side story gives the positive relationship between the wage and the number of firms. Since these two forces bring the system in the opposite directions, there exists a unique equilibrium irrespective of the specific form of the productivity and the signal structure G.

2.1 Riskiness of firm creation

Since the key departure from the literature is entrepreneurial risks, I devote this section to discuss the riskiness of entrepreneurship in the model.

The riskiness of entrepreneurship is defined by how informative the signal s is about the productivity z. If the signal s is very informative about entrepreneurial productivity z, agents can accurately forecast future profits when they become entrepreneurs. In contrast, if agents have uninformative signal s, becoming entrepreneurs involves substantial downside risks.

One way to define informativeness mathematically is to use the information sets. If the information sets of the two random variables are identical, i.e., the sigma algebras are identical $\sigma s=\sigma z$, there are no entrepreneurial risks. In contrast, if the signal is independent of the productivity $\sigma s\perp\!\!\!\perp\sigma z$, agents are called to face full risks. Although this definition has generality, it does not come with a natural framework to think about the intermediate cases. Thus, I use a weaker condition to define riskiness.

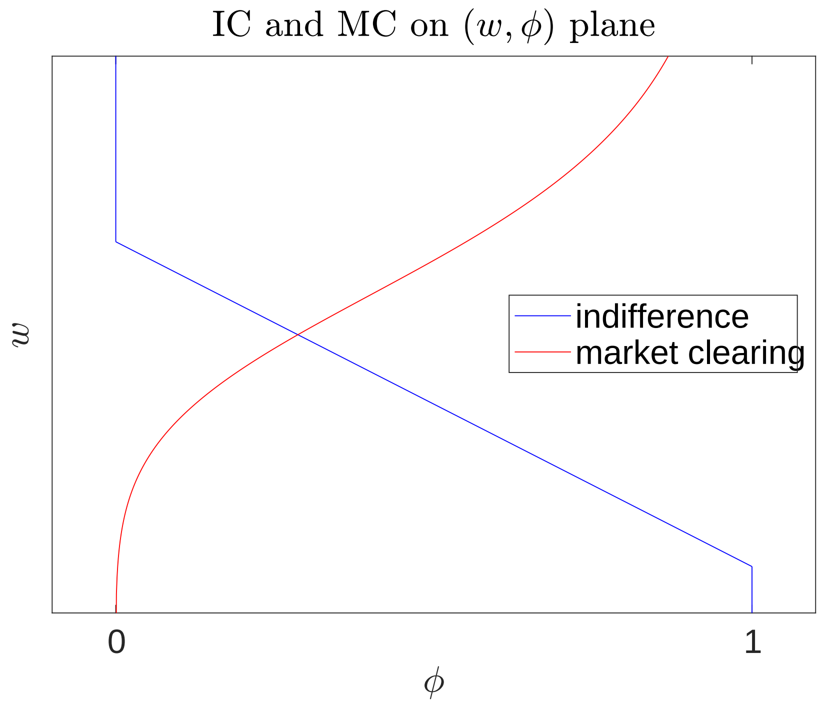

Figure 3: Equilibrium system on (w, ϕ) plane. The indifference condition gives a negative relationship, while the market clearing condition generates a positive relationship.

Definition 2. The riskiness of entrepreneurship for those who observe s is defined as the conditional variance $V(z|s) \in [0, V(z)]$. $V(z|s) = 0$ corresponds to no risks, while $V(z|s) = V(z) > 0$ corresponds to full risks.

One can use weaker conditions, defining no risks as the signal structure where each agent knows the best occupation for sure, i.e., $P(\pi(z, w) \geq w|s) \in \{1, 0\}$ for all s, and full risks as the signal structure where the signal is not informative about the occupation choice, i.e. $P(\pi(z, w) \geq w|s) = P(\pi(z, w) \geq w)$ for all s. This definition allows intuitive intermediate cases for some signal structure and is adopted in Appendix E.

In any case, the economy with no risks is characterized by $(\bar{z}, w)$ that satisfies the following indifference and market clearing conditions

$$\pi(\bar{z}, w) = w, \ \mathbb{E}\left[n(z, w) 1_{z \geq \bar{z}}\right] = G_z(\bar{z}), \tag{7}$$

and the economy with full risks is characterized by (w, ϕ) that satisfies another set of indifference and market clearing conditions

$$\mathbb{E}u(\pi(z, w)) = u(w), \ \phi \mathbb{E}n(z, w) = 1 - \phi. \tag{8}$$

One can see that, without risks, the utility function disappears from the equilibrium system.

With full risks, the equilibrium wage is determined by the indifference condition, and the number of firms is determined by the market clearing condition. Graphically, the full-risk case requires the indifference condition in Fig.3 to be horizontal for the interior number of firms. In terms of the literature, the two cases correspond to Lucas [1978] and Kanbur [1979], and Garicano et al. [2016] use the former to estimate the efficiency loss due to the SDP.

2.2 Market failure of the laissez-faire economy

Given the riskiness, I can discuss the efficiency of the laissez-faire economy. Once the market failure of the laissez-faire economy is understood, I describe the mechanism by which the SDP improves efficiency in section 3.

To discuss the efficiency, I define the planner's problem who can choose allocations at each state but faces the same informational constraints as individual agents. Formally, a state of the economy is a collection of the productivity realizations for all agents $\omega = \{z_i\}_i \in \Omega = [z_{min}, z_{max}]^{[0,1]}$. The planner chooses the contingency plan of the consumption $c = \{c_i(\omega)\}_{i\in[0,1],\omega\in\Omega}$, the employment schedule when agents become entrepreneurs $\{n_i(\omega)\}_{i\in[0,1],\omega\in\Omega}$, and the allocation of occupation $e = \{e(s_i)\}_i$. Note that the consumption and employment are contingent on each state, but the occupation has to be measurable with respect to the signal s since that is the informational constraint agents face. Such informational constraint is standard in the literature of efficiency analysis with informational frictions. (Angeletos and Pavan [2007])

Given a measurable Pareto weight $i \mapsto \Lambda_i$, the planner maximizes the weighted sum of the expected utility

$$U(c,\Lambda) = \int \mathbb{E}_\Omega u(c_i(\omega))\, d\Lambda_i \tag{9}$$

where $\mathbb{E}_\Omega$ denotes the expectation over the state Ω constructed from G_z.[1] Since the economy consists of heterogeneous agents differentiated by the observed signal $\{s_i\}_i$, I do not take a stand on the Pareto weight such as utilitarian or Rawlsian, and instead, focus on the production side. Note that since the economy does not have aggregate uncertainty and production in one firm is independent of the productivity of other firms, I can restrict the employment schedule to be measurable with respect to productivity, $n = \{n(z)\}_z$. Furthermore, since the planner maximizes output, the allocation of occupation is a threshold rule $e(s_i) = 1_{s_i \geq \bar{s}}$ without loss of generality. This is equivalent to choosing the number of firms $\phi = 1 - G_s(\bar{s})$. Hence, the planner's problem can be defined as follows.

Definition 3. Fix the fundamentals (u, f, G). A set of consumption, employment schedule,

[1] The formal construction relies on Kolmogorov extension theorem.

output and number of firms $\{c, n, Y, \phi\}$ is Pareto efficient if it solves

$$\max_{c,n,Y,\phi} U(c, \Lambda) \;\; s.t. \;\; \begin{cases} \int c_i(\omega)\, di = Y \\ Y = \phi \mathbb{E}\left[z f(n(z)) | s \geq G_s^{-1}(1-\phi)\right] & \forall \omega \in \Omega \\ \phi + \phi \mathbb{E}\left[n(z) | s \geq G_s^{-1}(1-\phi)\right] = 1 \end{cases} \tag{10}$$

for some Pareto weight Λ.[2]

The first constraint requires that individual consumption adds up to the aggregate output. Since the economy does not have aggregate uncertainty, it implies that the planner can provide full insurance, $c_i(\omega) = c(\omega')$ for all $\omega, \omega' \in \Omega$. An immediate implication is that the planner chooses employment schedule n and the number of firms ϕ to maximize output. The second constraint illustrates the production technology as a function of (n, ϕ). There are ϕ firms and each of them produces $\mathbb{E}\left[z f(n(z)) | s \geq G_s^{-1}(1-\phi)\right]$ on average. The conditional expectation reflects the selection of entrepreneurs $\bar{s} = G_s^{-1}(1-\phi)$. The third constraint describes the human resource constraint. ϕ agents become entrepreneurs and each of them hires $\mathbb{E}\left[n(z) | s \geq G_s^{-1}(1-\phi)\right]$ workers, which have to add up to the total population 1.

In general, the planner's solution depends on the Pareto weight Λ. The production side (n, Y, ϕ), however, can be uniquely determined independently of the Pareto weight.

Proposition 2. *Suppose that the fundamentals (u, f, G) satisfy the assumptions (A). Then, there is a unique interior solution $\phi^P \in (0,1)$ to the planner's problem (10).*

Proof. See Appendix A.1. □

To understand the planner's trade-off that pins down the interior solution, let $Y(\phi)$ denote the aggregate output as a function of the number of firms, obtained by maximizing out employment schedule $\{n(z)\}_z$ from the last two constraints of (10). To be concrete, let the production function be Cobb-Douglas $f(n) = n^\alpha$ with $\alpha \in (0,1)$. The aggregate output $Y(\phi)$ takes the following form

$$Y(\phi) = A(\phi)\, \phi^{1-\alpha} (1-\phi)^\alpha, \;\; A(\phi) = \left(\mathbb{E}\left[z^{\frac{1}{1-\alpha}} | s \geq G_s^{-1}(1-\phi)\right]\right)^{1-\alpha}. \tag{11}$$

One can see the two channels through which the number of entrepreneurs impacts the aggregate output.

The first channel is the average productivity $A(\phi)$. It describes a typical concern about entrepreneurship promotion. If the number of entrepreneurs increases $\phi \nearrow$, since the marginal

[2] Although all variables are functions of states, the assumption of no aggregate uncertainty saves the notation for aggregate variables.

entrepreneur is less productive $G_s^{-1}(1-\phi)\searrow$, the average productivity declines $A(\phi)\searrow$. However, the average productivity $A(\phi)$ is not the main technological trade-off that pins down the unique interior solution. After all, the average productivity is bounded $A(\phi)\in[z_{min}, z_{max}]$ and does not even react to the number of firms ϕ when the signal s is uninformative about the productivity z.

The second channel is the allocation of occupations $\phi^{1-\alpha}(1-\phi)^{\alpha}$. This is the key trade-off that pins down the unique interior solution. When most agents are workers $\phi\to 0$, the small number of firms employs a large number of workers. Due to the decreasing returns to scale assumption on the production technology f, the worker's marginal product is low. In this case, dividing the firm into two and shifting some of the workers to higher marginal product activities raises the aggregate output. However, such firm creation is not costless since it requires a decrease in the number of workers engaged in production. In the limit, if everyone becomes an entrepreneur $\phi\to 1$, no workers make production, resulting in 0 aggregate output despite the worker's marginal product being ∞.

The second channel is present even if the productivity distribution G_z is degenerate or the signal is not informative $s \perp\!\!\!\perp z$. In these cases, the optimal number of firms is $\phi^P = 1-\alpha$. It also differentiates the occupation choice model from the firm dynamics models following Hopenhayn [1992], in which labor supply is fixed and no entrepreneurs are needed to create additional firms.

Now I am ready to state the efficiency result. Let ϕ^{LF} and ϕ^P be the number of firms in the laissez-faire economy and the planner's solution. The following proposition states that entrepreneurial risks create the market failure of insufficient firm creation. Note that the assumptions on the riskiness $G(z|s)$ affect both the market equilibrium and the planner's solution.

Proposition 3. *Suppose that the fundamentals (u, f) satisfy the assumptions (A).*

1. *If there are no entrepreneurial risks, i.e., $V(z|s)=0$ for all s, the laissez-faire economy is Pareto efficient.*

2. *If there are entrepreneurial risks, i.e., $V(z|s)>0$ for all s, the laissez-faire economy is not Pareto efficient. In particular, the laissez-faire economy creates insufficient number of firms $\phi^{LF}<\phi^P$.*

Proof. See Appendix A.2. □

The result is not surprising if one notices that the economy without risks has complete markets so the analogy of the first welfare theorem holds, while the economy with risks features incomplete markets and uninsurable idiosyncratic risks. As is standard in the literature of incomplete markets, the result can be extended to constrained inefficiency as shown in Appendix F.1 and discussed in details in Ando and Matsumura [2017].

The intuition of the results is transparent if one considers a marginal increase in firms and aggregate output. Without risks, the entrepreneur with productivity z produces $\pi(z, w)$ and workers produce w. If one worker changes the occupation to entrepreneur, the net output gain is $\pi(z, w) - w$. In the laissez-faire economy (7), the marginal entrepreneurs are indifferent to the workers $\pi(\bar{z}, w) = w$. Therefore, increasing firms results in zero net output gain $\pi(\bar{z}, w) - w = 0$. With risks, the marginal entrepreneurs produce more than workers on average in the laissez-faire economy $\mathbb{E}[\pi(z, w) | s = \bar{s}] > w$. This is because, for risk-averse agents to become entrepreneurs, they have to be able to produce and consume more than workers on average. As a result, switching the marginal workers into entrepreneurs raise aggregate output by $\mathbb{E}[\pi(z, w) | s = \bar{s}] - w > 0$.

In the next section, I explain why the SDP can increase the number of firms.

3 Mechanism

I have shown in the previous section that Pareto improvement is possible only if there are entrepreneurial risks. This section uses the full-risk model to describe how the SDP could make Pareto improvement in three steps. The same argument carries on for the intermediate risk case, but I keep the exposition minimal for simplicity. At the end of the section, I provide discussions of the mechanism.

3.1 SDP can increase the number of firms

The first step is that the SDP can increase the number of firms by (1) preventing big firms from absorbing potential entrepreneurs and (2) reducing the entrepreneurial risks through general equilibrium.

To see the first channel, note that the SDP imposes regulatory costs on big firms, so they hire fewer workers. Since agents are either entrepreneurs or workers, the decline in the number of workers implies the number of firms increases in equilibrium. This is a partial equilibrium impact with fixed wage and is illustrated as the shift of the market clearing condition in Fig.4.

The second channel relies on the wage decrease in general equilibrium. To see this, note that the SDP makes entrepreneurship less attractive in partial equilibrium, i.e., $\pi^{SDP}\left(z, w^{LF}\right) \leq \pi^{LF}\left(z, w^{LF}\right)$ for all $z \in [z_{min}, z_{max}]$. Since the inequality is strict for big firms, the indifference

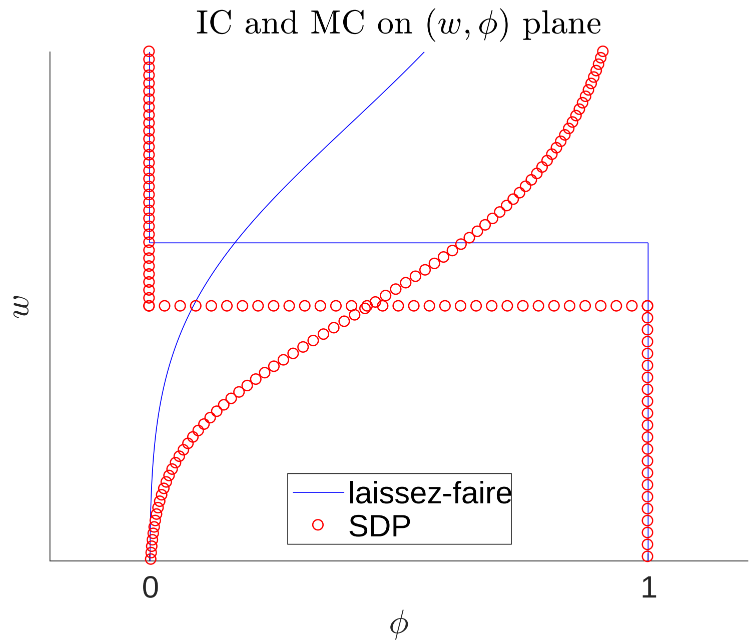

Figure 4: Indifference condition and market clearing condition.

condition (8) implies that the wage declines $w^{SDP} < w^{LF}$. Intuitively, since entrepreneurship is less attractive, agents want to become workers. As a result of the increase in worker supply, the wage declines. This force is illustrated as the shift of the indifference condition in Fig.4.[3]

The wage decline mitigates the downside risks of entrepreneurship. This is because, from the ex-ante point of view, even when businesses end up being small or unprofitable, entrepreneurs do not have to pay the regulatory costs and at the same time can enjoy cheaper labor costs. The mitigation of the downside risks is illustrated in the right figure of Fig.5. In partial equilibrium, the profit in the laissez-faire economy is higher for all productivity levels. However, such profit shift makes entrepreneurship less attractive. To regain equilibrium, the wage has to decline. As a result, the small firms that do not face regulation can hire workers cheaply, resulting in higher profit.

Both direct regulations on big firms and indirect subsidies to small firms contribute to firm creation. The former prevents big firms from absorbing potential entrepreneurs, and the latter reduces the downside risks of entrepreneurship.

[3]When the signal s is informative about the productivity z, the lower labor demand from big firms also pushes down the wage. Mathematically, the indifference condition is not horizontal, so the shift of the market clearing condition itself causes a wage decline.

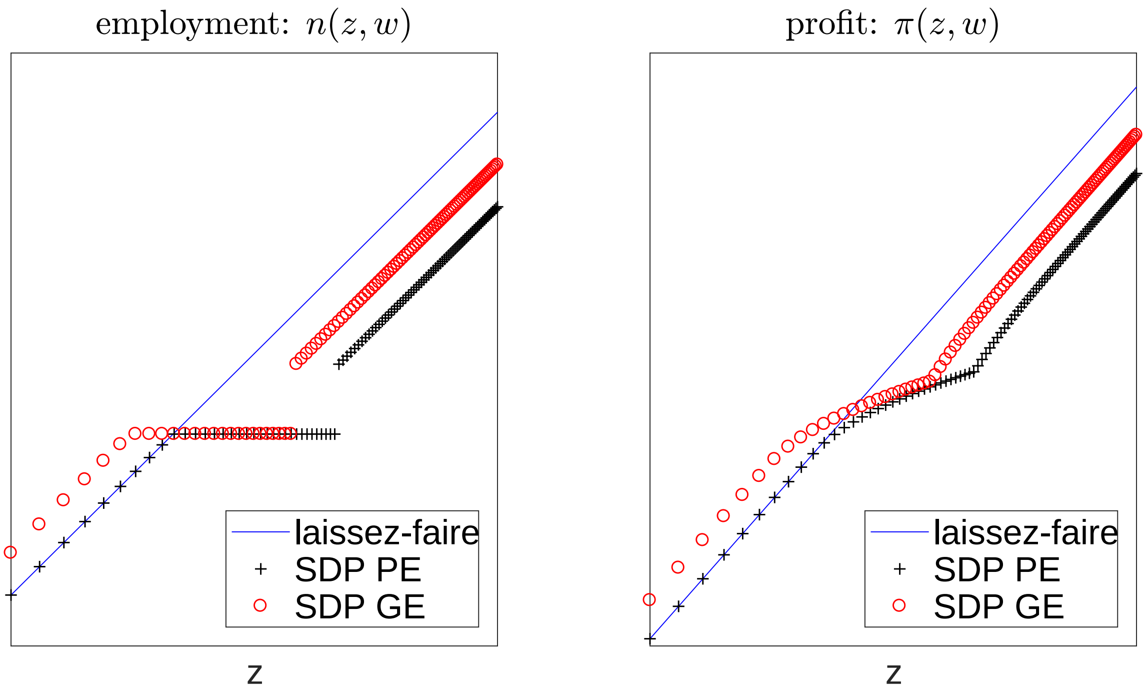

Figure 5: Employment schedule $n(z,w)$ and profit $\pi(z,w)$ for the laissez-faire economy and the economy with the SDP. PE denotes partial equilibrium where the wage is fixed at laissez-faire level w^{LF}. GE means general equilibrium where the wage is w^{SDP}.

3.2 Increase in the number of firms can raise output

The increase in the number of firms can lead to output increase. The intuition can be obtained by the perturbation argument. Suppose the government implements a small value of SDP around the laissez-faire $T = (\tau, F) \approx (1, 0)$. Appendix B shows that the marginal impact on the aggregate output is

$$\partial_T Y = \partial_T \phi \cdot (\mathbb{E}\pi(z,w) - w). \tag{12}$$

The expression illustrates that the policy impact around the laissez-faire economy is the multiplication of how many firms the policy increases $\partial_T \phi$ and how much additional output each firm produces $\mathbb{E}\pi(z,w) - w$. Intuitively, since entrepreneurs produce $\mathbb{E}\pi(z,w)$ and workers w, if an agent changes the occupation from worker to entrepreneur, the net gain is $\mathbb{E}\pi(z,w) - w$ on average.

The former $\partial_T \phi$ is positive from the previous section 3.1. The latter term $\mathbb{E}\pi(z,w) - w$ is also positive since, in the laissez-faire economy, risk aversion requires entrepreneurs to produce and consume more in equilibrium for the compensation of the risk-taking. Mathematically, this is a consequence of the Jensen's inequality

$$\mathbb{E}u\left(\pi\left(z,w\right)\right)=u\left(w\right)\Rightarrow\mathbb{E}\pi\left(z,w\right)>w \tag{13}$$

Thus, the SDP that increases the number of firms can increase the aggregate output.

Although I have used the full-risk model to explain the mechanism, the same argument holds for all the intermediate cases. As long as there are some risks, the marginal entrepreneurs produce and consume more than workers for the compensation of the residual risks

$$\mathbb{E}\left[\pi\left(z,w\right)|s=\bar{s}\right]>w. \tag{14}$$

In fact, even if I extend the model by assuming workers are risky, as long as entrepreneurs are riskier than workers, the same logic still applies.

However, the output does not increase if there are no risks. When there are no risks, the marginal entrepreneurs consume the same amount as the workers $\pi\left(\bar{z},w\right)=w$. As a result, even though the SDP increases the number of firms, it does not lead to output increase. In this sense, risks are essential for understanding the policy implications.

3.3 Increase in output can lead to Pareto improvement

The increase in output can lead to Pareto improvement if and only if the government returns the tax revenue. To see this, note that, since agents are ex-ante identical in the full-risk model (8), the welfare can be measured by the worker's consumption. As a result, if the government throws away the tax revenue, all agents are worse off due to the lower wage $w^{SDP}<w^{LF}$. However, if the government collects enough tax revenue and return it to agents, the disposable wage can increase. How much of the SDP generates tax revenue or ends up being non-pecuniary labor regulation is an empirical question. I show in the quantification section 4 that only a small amount of tax revenue is sufficient for the SDP to make Pareto improvement.

In addition to the aggregate output increase, the SDP also improves welfare through risk reduction. This channel is not specific to the full risk model but is also at work in the intermediate risk case. However, when agents are heterogeneous, the welfare criterion has to reflect the evaluation of inequality. If the welfare criterion values equality, the SDP raises welfare through both risk and inequality reduction.

Note that each of the above three steps involves non-trivial assumptions and necessitates the quantification exercise in section 4. For the first step, the increase in the number of firms hinges on the assumption that the wage does not decline too much. If wage declines to $w \approx 0$, firms can hire many workers, so the number of entrepreneurs decreases. In other words, the general equilibrium impact through wage cannot be larger than the partial equilibrium impact. For the second step, the output increase is discussed around the laissez-faire economy. This is to illustrate the intuition of the efficiency gain. For the SDP far from the laissez-faire, there is an efficiency loss due to the bunching. Therefore, whether the output increases or not depends on the balance of the two forces and has to be determined numerically. For the third step, the Pareto improvement requires the government to raise enough revenue. Although the SDP increases output and reduces risks for each entrepreneur, the number of entrepreneurs also increases. Since the entrepreneurs need to consume more on average to obtain the same utility as workers, whether the SDP can increase enough output to Pareto improve all agents is not obvious. As shown in section 4, it turns out that all of these assumptions are satisfied under the standard parametric specifications.

3.4 Discussion about the mechanism

Before moving to the quantification exercise, I provide critical discussions about three features of the mechanism.

First, since the mechanism involves the occupation change from workers to entrepreneurs, one might be concerned about unemployment. Although the model does not address unemployment, I make two heuristic observations. First, there are many countries that have both SDPs and low unemployment rates. Therefore, at least in the long run, SDPs and unemployment can be considered separate issues. Second, the thought experiment that I am interested in is the removal of the SDP. Thus, the relevant trade-off is the lower unemployment rate in the short run and the lower aggregate output in the long run. Therefore, the social cost associated with unemployment is not an issue in both the long and short run in the current context.

Second, the output increase depends on entrepreneurs producing more on average than workers in the laissez-faire economy, i.e., $\mathbb{E}\pi^{LF}\left(z, w^{LF}\right) > w^{LF}$.[4] Hamilton [2000] and Moskowitz and Vissing-Jorgensen [2002] argue that entrepreneurs earn less than workers on average. Their studies have generated a literature that tries to explain why people choose to become entrepreneurs in the first place (e.g. Vereshchagina and Hopenhayn [2009]). More recent papers, however, revisit the issue and argue that entrepreneurs indeed earn more than workers on average. For instance, Levine and Rubinstein [2017] argue that the definition of entrepreneurs used in the literature may not be appropriate. If entrepreneurs are defined to

[4] Since many countries adopt SDPs, it is not easy to empirically test whether $\mathbb{E}\pi^{LF}\left(z, w^{LF}\right) > w^{LF}$ holds in the unobservable laissez-faire economy. That being said, we can check if a sufficient condition is empirically satisfied, assuming $\mathbb{E}\pi^{SDP}\left(z, w^{SDP}\right) > w^{SDP}$ implies $\mathbb{E}\pi^{LF}\left(z, w^{LF}\right) > w^{LF}$.

be incorporated self-employed, the mean return is much higher than salaried workers. Manso [2016] points out that the estimations in the literature might be biased due to the usage of cross-sectional data to compute the mean of entrepreneurial earnings. Manso [2016] reports that, once the option value of experimentation is incorporated, the mean lifetime earnings of entrepreneurs are higher than those of salary workers in the U.S. Recognizing such debate on the measurement of entrepreneurial return, I explore the implications of the standard occupation choice model in this paper.

Finally, the Pareto improvement is based on the tax return. Since many SDPs are non-pecuniary such as working hours regulations, there might be no tax revenue in reality. In the next quantification section 4, I show that, even when all the tax revenue is thrown away, the economy with SDP generates higher output than the laissez-faire economy. In terms of welfare, I show in Appendix C that the economy with SDP generates higher welfare than the laissez-faire economy with the same amount of non-distortionary regulation costs.

4 Quantification

This section quantifies the welfare impact of removing the SDP. The main result is that removing the SDP leads to lower output and welfare. The result of output decline is the opposite of the previous literature and highlights the potential importance of the channel described in section 3.

To isolate the role of entrepreneurial risks, I closely follow the setup of Garicano et al. [2016]. In particular, I impose the parametric assumptions on the fundamentals, with utility and production functions being constant relative risk aversion and Cobb-Douglas

$$u(c) = \frac{c^{1-\gamma} - 1}{1-\gamma}, \ \gamma > 0, \ f(n) = n^{\alpha}, \ \alpha \in (0,1) \tag{15}$$

and the distribution of the productivity being the power law

$$g_z(z) = \frac{1-\beta_z}{z_{max}^{1-\beta_z} - z_{min}^{1-\beta_z}} z^{-\beta_z}, \ z_{min} \leq z \leq z_{max}. \tag{16}$$

This productivity distribution leads to a broken power law of the firm size distribution as detailed in Appendix D.

I focus on the full risks model (8) and relegate a specific intermediate signal case in Appendix E. In this way, the quantitative exercise in this section is the exact opposite polar case of Garicano et al. [2016] with the same set of fundamentals. Therefore, one can transparently see the implications of the entrepreneurial risks without being affected by the specific choice of the signal structure.

Formally, the economy with the SDP (τ, F, N) and the non-distortionary income tax t has the equilibrium $\left(w^{SDP}, \phi^{SDP}\right)$ that satisfies the indifference condition

$$\mathbb{E}u\left(\left(1-t\right)\pi\left(z,w\right)\right)=u\left(\left(1-t\right)w\right), \tag{17}$$

profit maximization (1), and the market clearing condition in (8). The income tax t is non-distortionary because, when the utility function is CRRA, $1-t$ drops out of the equilibrium system. As a result, all the equilibrium objects except for welfare are independent of the income tax rate t. To determine the welfare, I set the income tax rate $t = t\left(\tau, F\right)$ by imposing the balanced budget constraint

$$\begin{aligned}\phi^{SDP}\mathbb{E}\left[\left\{\left(\tau-1\right)w^{SDP}n\left(z,w^{SDP};\tau,F\right)+F\right\}1_{n(z,w^{SDP};\tau,F)>N}\right]\\ +t\left\{\phi^{SDP}\mathbb{E}\pi\left(z,w^{SDP};\tau,F\right)+\left(1-\phi^{SDP}\right)w^{SDP}\right\}=0.\end{aligned} \tag{18}$$

I compare the economy with SDP against the counterfactual laissez-faire economy where $(t, \tau, F) = (0, 1, 0)$. Specifically, I compare aggregate output, welfare, number of firms, and wage in the two economies. Note that the welfare is given by the disposable income $\left(1-t\right)w^{SDP}$ and w^{LF}.

There are two differences from Garicano et al. [2016]. The first is the existence of entrepreneurial risks, and the second is the welfare criterion. The welfare criterion of Garicano et al. [2016] is aggregate output minus tax revenue. To make the comparison consistent, I also report output minus tax revenue in section 4.2. An alternative welfare specification assuming all the SDP being non-pecuniary is available in Appendix C. Irrespective of the specification of the welfare criterion, all other equilibrium objects (Y, w, ϕ) are directly comparable.

4.1 Parameter values

In this section, I specify the parameter values for quantification. The main takeaway is that I can use the same parameter values as Garicano et al. [2016], and therefore all the differences in the results are driven by the existence of entrepreneurial risks, not other elements such as the data used and specifications of the model.

Table 1 summarizes the choice of the parameter values based on the same French data used by Garicano et al. [2016]. Note that the fixed tax F is negative. This does not mean the firms at size N get subsidies. Since the variable tax τ applies to all workers, the tax payment for firms at N is positive $\left(\tau-1\right)wN+F=.186w>0$.

The main parameter of interest is the SDP (τ, F, N). The threshold N is determined

parameter	value	explanation
γ	2	CRRA parameter
α	.8	Cobb-Douglas parameter
τ	1.023	marginal tax
F/w	$-.941$	normalized fixed tax
β_z	5	power law parameter of G_z
z_{max}	8.07	productivity upper bound
z_{min}	1.4	productivity lower bound

Table 1: Parameter values for counterfactual analysis. All parameters are taken from Garicano et al. [2016] except for (γ, z_{min}). Wage w is not identified so is normalized to $w = 1$ without loss of generality.

by the law so there is no other choice than $N = 49$. The variable and fixed components (τ, F), as pioneered by Garicano et al. [2016], can be estimated jointly with the power law parameter β_z by using the data of firm size distribution. To be more concrete, I derive the equilibrium firm size distribution $n \mapsto G_n(n; \tau, F, \beta_z)$. This firm size distribution follows a broken power law and its derivation is given in Appendix D. The maximum likelihood method then fits the empirical firm size distribution data $\{n_i\}_i$ by choosing (τ, F, β_z) and other auxiliary parameters. In the estimation, I follow Garicano et al. [2016] to truncate the data and likelihood function ($n \leq 10$ and $n \geq 10,000$) so that the estimation is not impacted by the two tails that do not fit power law well.

In principle, since I use a different model with entrepreneurial risks, I need to recollect the same data and redo the same estimation. However, I can use the same estimates as Garicano et al. [2016] because the implied firm size distributions after truncation are identical. To see this, recall that the only difference in the equilibrium firm size distributions of the two models (7) and (8) is whether low productivity agents become entrepreneurs or not. With no risks, only high productivity agents $\{i : \pi(z_i, w) \geq w\}$ become entrepreneurs, while with full risks, all productivity levels are possible. When the truncated distribution does not capture the low productivity region, the fact that the broken power law is invariant to truncation implies that the two models generate an identical truncated firm size distribution.[5]

Other parameters follow the calibration and normalization of Garicano et al. [2016]. For instance, the equilibrium wage is unidentified so is normalized to be 1. The parameter that

[5] The likelihood function in Garicano et al. [2016] also incorporates measurement error. Although this complication might potentially create discrepancies between the estimates, I show they are numerically negligible in Appendix D using Amadeus data and the models with and without entrepreneurial risks.

Equilibrium objects	Full risks	No risks	Concepts
$\ln Y^{LF} - \ln Y^{SDP}$	-1.5%	$.02\%$	output
$\ln \phi^{LF} - \ln \phi^{SDP}$	-8%	-7.2%	number of firms
$\ln Y^{LF} - \ln \left(Y^{SDP} - \text{tax}\right)$	$-.3\%$	1.3%	output - tax
$\ln w^{LF} - \ln \left\{(1-t)\, w^{SDP}\right\}$	-1.25%	NA	welfare
$\ln w^{LF} - \ln w^{SDP}$	$.002\%$	1.8%	wage

Table 2: Results of the counterfactual analysis. The column "Full risks" is the result of the counterfactual analysis using full-risk model. The one with "No risks" is the result of Garicano et al. [2016]. Welfare in the model without risks is NA because agents are heterogeneous, so the welfare depends on the Pareto weight. Specifically, big firms and workers lose, while small firms win.

is not in Garicano et al. [2016] is the risk aversion γ. I choose the standard number $\gamma = 2$ and discuss the robustness of the results to it in section 4.2. Other parameters $(\alpha, z_{min}, z_{max})$ also follow Garicano et al. [2016]. For the details, see Appendix D.

4.2 Results

The results of the counterfactual analysis are summarized in table 2. The first column shows the results of the counterfactual analysis based on my model with full entrepreneurial risks. The second column is a copy from Garicano et al. [2016] which are based on the model with no entrepreneurial risks.

The first row, which is the main result, shows that the economy with SDP produces the higher output than the laissez-faire economy. This might be surprising if one just sees that the marginal products equalize among all firms in the laissez-faire economy, yet the laissez-faire economy produces less than the economy in which the SDP creates wedges in the marginal products. Fig.6 summarizes the intuition. The red line represents the production possibility frontier of the economy with the SDP $Y^{SDP}(\phi)$ as a function of the number of firms ϕ. For each fixed number of firms, the removal of the SDP improves production efficiency so the production possibility frontier of the laissez-faire economy is located above. Thus, for a given number of firms, removing SDP does improve resource allocation. However, removing the SDP exacerbates the market failure of firm creation. In particular, as discussed in section 3, big firms suck up human resources from entrepreneurial activities and the entrepreneurship becomes riskier. In consequence, the number of firms declines as shown in the second row. What the first row shows is that this efficiency loss in the extensive margin can quantitatively

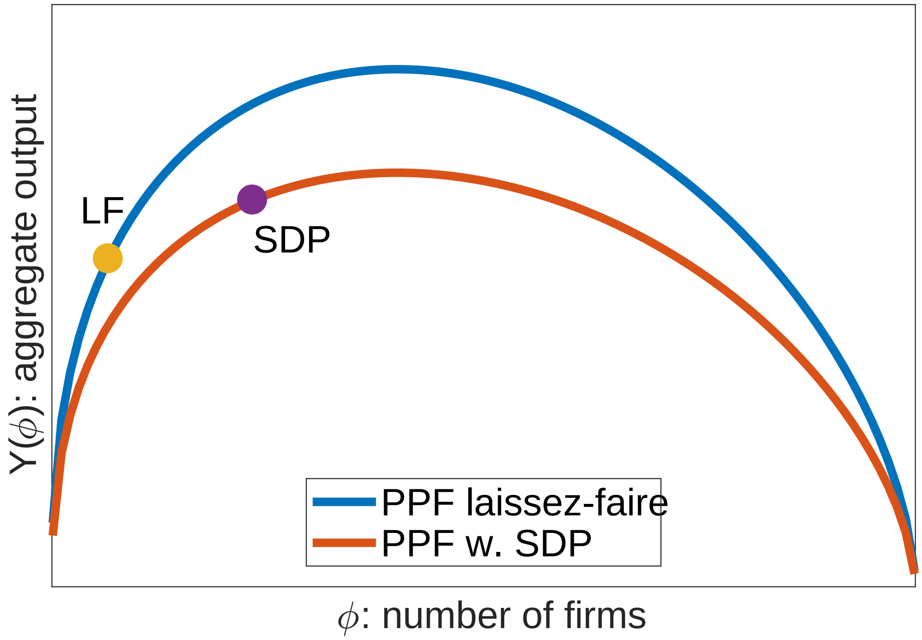

Figure 6: Aggregate output $Y(\phi)$. Points A, B, and C are the counterfactual allocation without the SDP, the allocation in the economy with the SDP and the efficient allocation.

outweigh the gain in the intensive margin.[6]

The third row of table 2 highlights the implication of the entrepreneurial risks in a stronger sense. It shows that, even if the laissez-faire economy is given the tax advantage, the economy with the SDP still produces more when there are entrepreneurial risks. In fact, this output concept is used as the main welfare criterion in Garicano et al. [2016]. In this sense, my result is the opposite of Garicano et al. [2016].

The fourth row of table 2 shows that the welfare is also higher in the economy with the SDP. As discussed in section 3, the result implies that the SDP increases the aggregate output by a non-trivial amount. Quantitatively, the tax revenue

$$\mathcal{T} = \phi^{SDP}\mathbb{E}\left[\left\{(\tau-1)\, w^{SDP} n\left(z, w^{SDP}; \tau, F\right) + F\right\} 1_{n(z,w^{SDP};\tau,F)>N}\right] \tag{19}$$

amounts to 1.14% of aggregate output in the current setting. Thus, the result highlights the first order importance of the channel highlighted in this paper.

The fifth row confirms the decline in wage. As discussed in section 3, this is an implication of the specific form of the SDP (τ, F) and therefore it is not necessarily a robust implication of SDPs. In fact, I show in section 5 that, if the SDP is allowed to be an arbitrary non-linear

[6] This is not the case with Garicano et al. [2016] because the laissez-faire economy is at the top of the production possibility frontier when there are no risks. Therefore, no matter how the number of firms changes, the aggregate output always declines.

function $n \mapsto T(n)$, wage can actually increase. Hence, even if the government throws away the tax revenue, the SDP can achieve Pareto improvement.

Finally, I argue these results are robust to the change in the risk aversion parameter γ. For instance, the break-even point for the output $Y^{SDP} = Y^{LF}$ and welfare $(1-t)\, w^{SDP} = w^{LF}$ is $\gamma = .01$. This small risk aversion is not surprising if one notes that the output gain from removing the SDP in the model with no entrepreneurial risks is only .02%. Similarly, the break-even point for $Y^{SDP} - tax = Y^{LF}$ is $\gamma = .7$, still way below the standard parameter value $\gamma = 2$.

5 Optimal SDP

The analysis so far has demonstrated that an economy with the threshold type SDP (τ, F) can be better than the laissez-faire economy. In other words, an economy with wedges among firms' marginal products can lead to a better outcome than the one with equal marginal products. A natural question is "what are the optimal wedges implied by SDPs and is the threshold type SDP (τ, F) similar to the optimal SDP?" To answer this question, I study the optimal non-linear SDP and show that the observed threshold type SDP (τ, F) shares qualitative features with the optimal SDP.

The optimal SDP is defined as the solution to an optimal taxation problem. Let $n \mapsto T(n)$ be the tax imposed on the firms that hire n workers. The government is benevolent and maximizes the welfare

$$U = \phi \mathbb{E} u(\pi(z, w)) + (1-\phi)\, u(w)\,. \tag{20}$$

Although the expression is the same as utilitarian welfare, any Pareto weight leads to the same welfare function since the full-risk economy consists of homogeneous agents.

The government faces the market equilibrium constraints and the budget constraint. The market equilibrium constraints consist of the two equations (8) and the firm's profit maximization problem

$$[\pi(z, n), n(z, w)] = \max_{n \geq 0} z f(n) - wn - T(n)\,. \tag{21}$$

Note that the threshold SDP (τ, F) is a special case where $T(n) = \{(\tau - 1)\, wn + F\}\, 1_{n>N}$. The budget constraint takes the following form

$$\phi \mathbb{E} T(n(z, w)) = \mathcal{T} \tag{22}$$

where $\mathcal{T} \geq 0$ is exogenously given. When $\mathcal{T} = 0$, the government budget is balanced. Later, I conservatively set $\mathcal{T} > 0$ and investigate whether SDPs improve welfare even if resources are taken out of the economy.

Formally, the optimal taxation problem can be formulated as follows.

Definition 4. Given the budget $\mathcal{T} \geq 0$, the government chooses (w, ϕ, π, n, T) to maximize wage w subject to (8), (21), and (22).

The optimal SDP problem has a similar structure as Mirrlees [1971] except that it has a general equilibrium wage. I reformulate it as a mechanism design problem and solve the problem within the set of differentiable functions.

Proposition 4. *The solution to the optimal SDP problem can be obtained as the solution to*

$$\max_{w, \{n(z), \pi(z)\}_z} w \ s.t. \begin{cases} \mathbb{E}u\left(\pi\left(z\right)\right) - u\left(w\right) = 0 \\ \pi'\left(z\right) = f\left(n\left(z\right)\right) \\ n'\left(z\right) \geq 0 \\ \mathbb{E}\left[zf\left(n\left(z\right)\right) - wn\left(z\right) - \pi\left(z\right)\right] - \mathcal{T}\left(1 + \mathbb{E}n\left(z\right)\right) = 0 \end{cases} . \tag{23}$$

Proof. See Appendix A.3. □

The first constraint is the indifference condition. The second is the envelope condition of the truth-telling constraint, and the third is the monotonicity constraint that corresponds to the second order condition. The last constraint is the government budget. Mathematically, this is an optimal control problem.

It turns out that the optimal SDP and the observed threshold type SDP share some qualitative features. To see it, let $(w^o, \phi^o, n^o, \pi^o, T^o)$ be the solution to the optimal SDP problem.

Proposition 5. *Suppose that the fundamentals (u, f, G_z) satisfy the assumptions (A), there are risks $V(z) > 0$, and the solution does not have bunching*

$$\frac{d}{dz} n^o\left(z\right) > 0, \ \forall z \in \left(z_{min}, z_{max}\right).$$

Then, the optimal SDP distorts medium-sized firms, i.e., for all $z \in (z_{min}, z_{max})$,

$$w^o = z_{min} f'\left(n^o\left(z_{min}\right)\right) = z_{max} f'\left(n^o\left(z_{max}\right)\right) < zf'\left(n^o\left(z\right)\right).$$

For a small budget $\mathcal{T} \geq 0$, the optimal SDP increases the number of firms $\phi^o > \phi^{LF}$.

Proof. See Appendix A.4. □

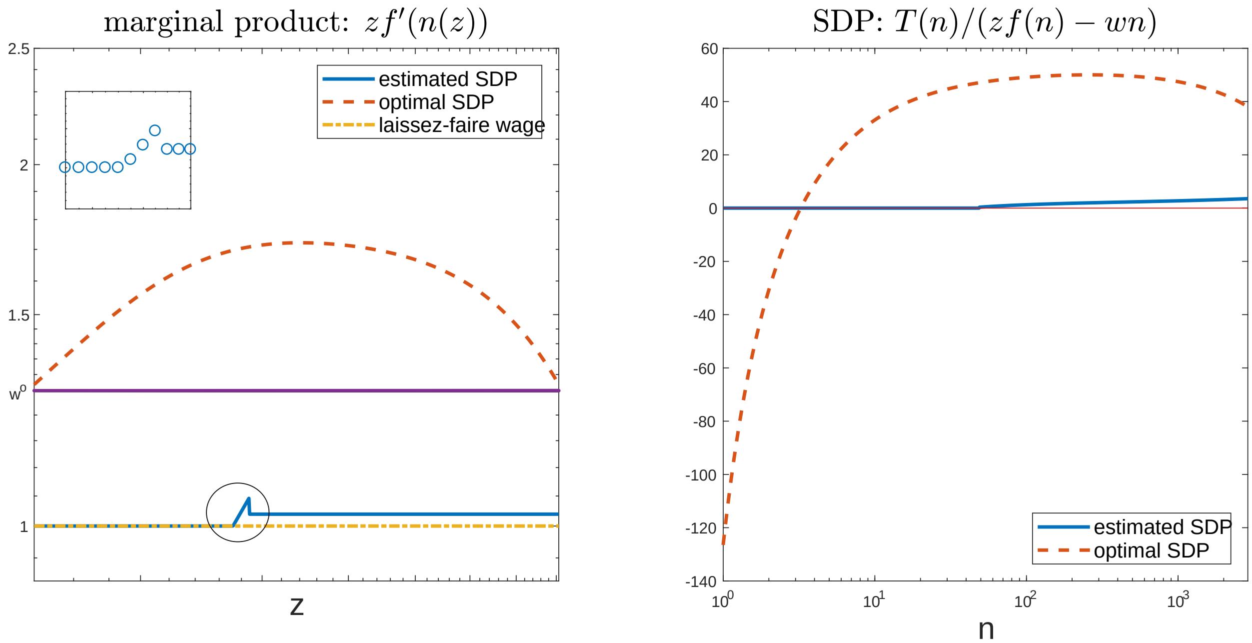

Figure 7: The left panel shows the marginal product as the measure of distortion. The right panel shows the optimal SDP as the percentage of the before-tax profit. The wage in the economy with the estimated threshold SDP is normalized to be $w^{SDP} = 1$.

Recall that, as Fig.5 suggests, the bunching firms are most distorted in the sense that their marginal products are higher than other firms. Moreover, in section 3, we have seen that the threshold type SDP increases the number of firms. Proposition 5 states that the optimal SDP shares these features under fairly weak conditions.

To visualize this point, I show a numerical solution in Fig.7. The exercise is based on the same parameterization as section 4, i.e., (15) and (16), the same parameter values as in table 1, and the same tax revenue $\mathcal{T} > 0$ as (19). Under these conditions, it is straightforward to check that no bunching condition is satisfied.

The distortion, measured by the deviation of marginal product equalization, can be seen in the left panel of Fig.7. One can see that, under the optimal SDP, the marginal products are highest in the middle. This is also the case with the threshold type SDP (τ, F), under which the bunching firms with the highest productivity $\sup\{z : n(z, w) = N\}$ are most distorted. One way to understand this property of the optimal SDP is to note the role of each firm. Highest productivity firms are socially important since they produce goods efficiently. Lowest productivity firms are also important since they represent the downside risk of entrepreneurship. Thus, the distortion is concentrated on the middle firms.

The left panel of Fig.7 also shows that the wage is higher under the optimal SDP than the threshold SDP by more than 30%. Therefore, the wage level under the optimal SDP is higher than the laissez-faire economy $w^{SDP} < w^{LF} < w^o$. This is because the optimal SDP subsidizes

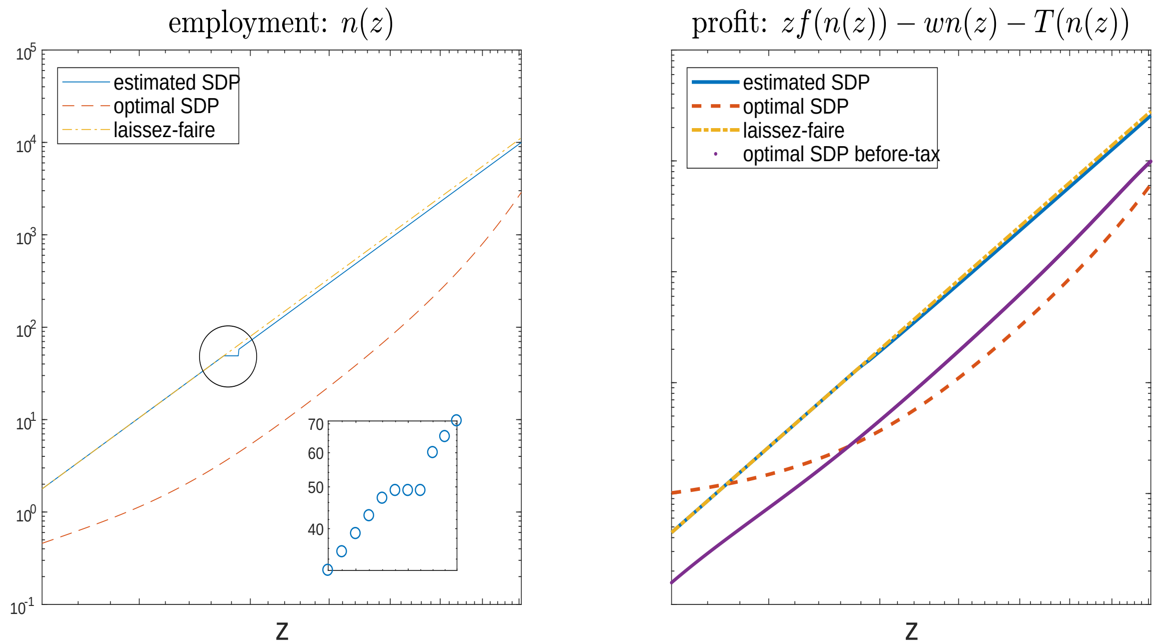

Figure 8: Employment and profit schedule under the estimated threshold SDP, optimal SDP and laissez-faire economy.

small firms as shown in the right panel of Fig.7, and therefore the welfare is higher by the insurance effect. This feature of the optimal SDP is not unrealistic since there are various subsidies available for small firms in practice. More importantly, the observation suggests that the result $w^{SDP} < w^{LF}$ in section 4 is not a robust implication of SDPs in general but is a consequence of the unrealistically parsimonious modeling of the SDP (τ, F) that abstracts from small firm subsidies. Thus, if the quantification exercise includes subsidies, the wage could go up and the welfare could be higher. In this sense, the quantification result is conservative and should be interpreted as saying that the SDP could enhance efficiency even when the SDP does not subsidize small firms.

For the optimal SDP in the right panel of Fig.7, I make two more observations. First, the tax paid by the largest firms is smaller than the middle firms in percentage. This is not saying that the absolute amount of the tax is lower. Actually, the absolute amount of the tax is monotonically increasing in productivity. It is just that the percentage of tax out of the before-tax profit is lower. The second observation is that the subsidy is not mechanical. This is where the government budget $\mathcal{T}$ matters. If I impose the balanced budget $\mathcal{T} = 0$, some firms have to be subsidized by construction. But here, I set the tax revenue $\mathcal{T}$ to be the same number as in the previous quantification exercise, which is positive, so the subsidy is not mechanical. In other words, the mechanism designer does not have to but still chooses to subsidize the unproductive small firms since it affects ex-ante firm creation.

Now let me use Fig.8 to explain that the optimal SDP increases the number of firms. In the left panel that plots the employment schedule for each productivity level z, one can see that under the optimal SDP every firm hires fewer than both the laissez-faire and estimated SDP, so the number of workers declines and firms increases. This is due to both the regulation and higher wage. Thus, the laissez-faire economy generates too few firms and they are too big.

Finally, the right panel of Fig.8 describes the insurance effect of the optimal SDP. The purple line is the before-tax profit under the optimal SDP and the red dashed line is the after-tax profit. The fact that the after-tax profit schedule is flatter reflects insurance. This is also true in the threshold type SDP (τ, F), although the insurance effect of the estimated threshold SDP comes from general equilibrium wage effect instead of the direct insurance as the optimal SDP.

6 Financial frictions

The analysis so far has shown that the insufficient firm creation due to entrepreneurial risks might justify SDPs. A natural question is whether this is always the case. I argue the answer is negative by showing that, under severe financial frictions, removing SDPs might improve efficiency. As a policy implication, whether to remove SDPs depends on the level of financial development.

In principle, there are three cases where SDPs should be removed. The first case is when entrepreneurial risks are not quantitatively important. If individual agents obtain a precise signal s_i about their entrepreneurial productivity z_i, SDPs are harmful as shown by Garicano et al. [2016] and Gourio and Roys [2014]. The second case is when SDPs are so radical that the efficiency loss dominates the gain from more efficient firm creation. In French data, this is not the case, but data from other countries might find a large estimate of SDP (τ, F). Both of these cases rely on the idea that the cost of SDPs is larger than the benefit.

The third case, which is theoretically more interesting, is when there are too many firms in the laissez-faire economy. This case is different from the previous two because removing SDPs incurs no cost. Fig.9 illustrates this situation. The SDPs that increase the number of firms not only distort production by creating wedges among firms but also worsen the market failure of firm creation by adding more firms to the economy that already has excessive firms. Thus, removing SDPs enhances efficiency through two channels.

Potentially, there are several frictions that could generate excessive firms. As an example, I argue that severe financial frictions can lead to excessive firm creation in the next section.

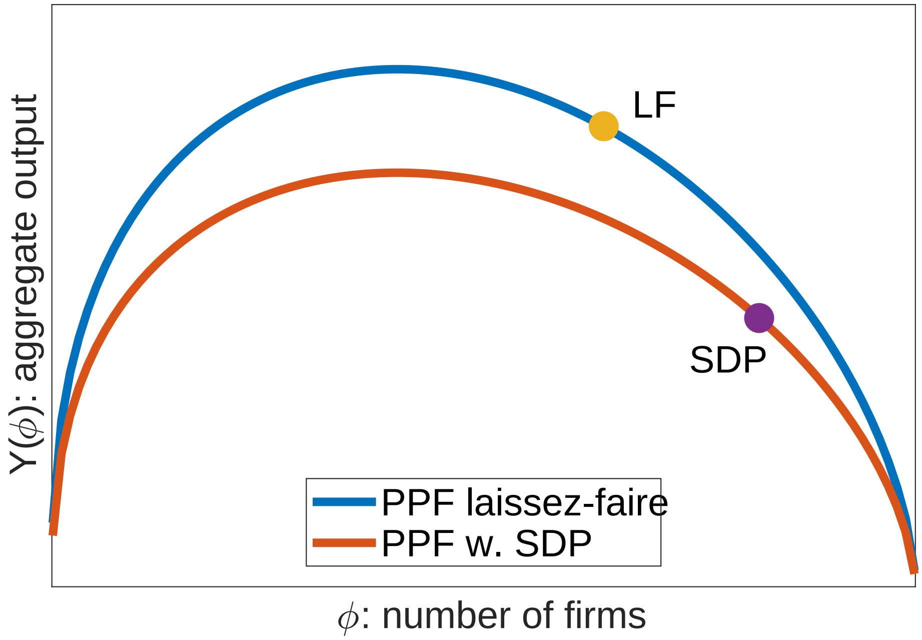

Figure 9: Aggregate output $Y(\phi)$ when market generates too many firms. Points A, B, and C are the market equilibrium without SDPs, the market equilibrium with SDPs, and the unconstrained efficient allocation.

6.1 Financial frictions

One situation that leads to excessive firm creation is severe financial frictions that prevent firms from expanding their employment to their profit-maximizing sizes. In this section, I show that, under severe financial frictions, the laissez-faire economy creates more firms than the Pareto efficient level.

To illustrate this point parsimoniously, I introduce a simple form of financial friction to the firm's profit maximization problem as follows.

$$\pi(z,w) = \max_{n} zf(n) - wn \ s.t. \ wn \leq \lambda zf(n) \tag{24}$$

where $\lambda^{-1} > 0$ represents the severity of the financial friction. If the financial frictions are not severe $\lambda > \alpha$, the constraint does not bind. As the financial constraint becomes more severe, i.e., λ gets close to 0, the constraint becomes more binding. Such form of financial constraint has been used in the literature, e.g., Bigio and La'O [2016], and is of particular interest in this paper since the marginal products among firms are equalized and therefore the resulting misallocation cannot be understood by the standard marginal product argument

$$z_i f'(n_i) = \max\left\{1, \frac{\alpha}{\lambda}\right\} w, \ \forall i. \tag{25}$$

One way to micro-found the constraint is limited enforcement. Each entrepreneur owes the workers their salaries. However, due to the limited enforcement, the entrepreneur can divert $1 - \lambda$ fraction of the sales and run without paying the salaries. In this case, workers are not willing to work for the firm if their salaries exceed λ percent of the sales. One can also attribute λ to institutional immaturity or other non-financial obstacles. After all, the formulation is a reduced-form, so I call it financial friction for simplicity.

With this financial constraint, I conduct the efficiency analysis using the full-risk model. Specifically, let $\left(w^{LF}, \phi^{LF}\right)$ be the equilibrium in the laissez-faire economy characterized by (8) and (24), ϕ^P be the planner's solution that solves (10) with the uninformative signal $s \perp\!\!\!\perp z$. The following proposition states that under severe financial friction $\lambda \approx 0$, the laissez-faire economy generates too many firms so that the SDPs that increase the number of firms exacerbate human resource misallocation.

Proposition 6. *Fix CRRA utility, CD production function (u, f), and an arbitrary risk distribution G_z with bounded support. Under the severe financial friction, i.e., for small $\lambda \geq 0$, the laissez-faire economy generates an excessive number of firms*

$$\phi^P < \phi^{LF} \leq 1 = \lim_{\lambda \to 0} \phi^{LF}.$$

Proof. See Appendix A.5. □

The proposition imposes particular functional form assumptions for analytical cleanliness, but they can be relaxed at the cost of heavier notation. The inefficiency result can be strengthened to constrained inefficiency. As is the case for Proposition 3, I relegate the result to Appendix A.5.

To understand the intuition behind the result, recall that the planner is free from financial frictions. As a result, the efficient allocation is invariant to the severity of financial frictions λ. Therefore, the inefficiency follows if the number of firms in market equilibrium can get arbitrarily close to 1. The intuition comes from the firm side. If financial frictions are severe $\lambda \to 0$, due to the cash shortage or limited enforcement, firms cannot expand their employment. In the extreme case $\lambda = 0$, no firm can hire workers. Since the labor demand from firms is limited, those who are not employed have to do their own businesses, leading to an increase in the equilibrium number of firms $\phi \to 1$. Thus, under severe financial friction $\lambda \approx 0$, the number of firms in the laissez-faire economy ϕ^{LF} exceeds that of the efficient allocation ϕ^P.

The analysis of the financial friction implies that whether SDPs should be removed or not depends on the level of financial development. For instance, in developing countries with immature financial infrastructure, SDPs should be removed and resources should be devoted to relaxing financial frictions, even if firm creation is risky.

This implication might be surprising if one considers that entrepreneurial risks and financial frictions are often cited as two big obstacles to entrepreneurship. In fact, Proposition 3 and 6 suggest that the two frictions cause market failure in the opposite directions, i.e., entrepreneurial risks cause the market to create too few firms while financial frictions cause too many. There are two ways to reconcile the conventional wisdom and my result.

The first one is about the modeling. If one models the financial friction as the fixed cost necessary to start a firm, it does suppress firm creation. However, as Quadrini [2008] discusses, most entrepreneurs can lower the fixed cost by choosing to start small. In fact, the moment entrepreneurs face financial frictions most is when they want to expand. In this sense, typical financial frictions that entrepreneurs face might be better captured by Eq.(24) than the fixed entry cost.

The second one is to recognize that the conventional wisdom is based on partial equilibrium, and is not suitable to understand the impact of the financial frictions on the general equilibrium. In partial equilibrium, financial friction reduces profit so people do not want to become entrepreneurs. In general equilibrium, however, since firms cannot expand employment due to the financial friction, those unemployed have to do business on themselves. To rationalize the existence of many more small firms, the wage declines to the level consistent with such scenario.

The positive correlation between the severity of financial frictions and firm creation is also consistent with data. Fig.10 plots entrepreneurship activity against the severity of financial frictions. The entrepreneurship is measured by the percentage of the $18 - 65$ population who is either a nascent entrepreneur or owner-manager of a new business, taken from Global Entrepreneurship Monitor for $2001 - 2016$. The severity of financial frictions is measured by the value of collateral needed for 100 unit of loans, taken from the World Bank Enterprise Survey.

One can see that more severe financial frictions are associated with higher entrepreneurial activities. Fig.11 presents another data consistent with the analysis. It is based on survey data, asking country experts whether a country has enough entrepreneurship financing. Although the data are subjective, it contains more samples. One can see that higher availability of financing is associated with lower entrepreneurial activities.

What these figures do not show is the efficiency of the market equilibrium. In other words, they suggest financial frictions are associated with more firm creation but do not provide information about whether the increase is too much or not. In this sense, Proposition 6 provides a normative insight into the empirical regularities.

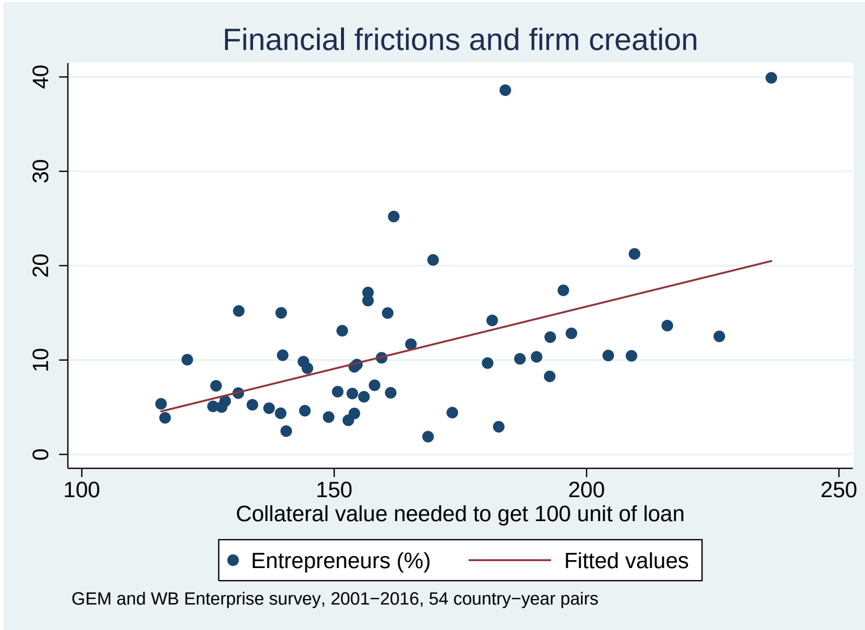

Figure 10: Entrepreneurship is positively correlated with financial frictions. x axis represents the collateral value required for 100 unit of loans. y axis represents the percentage of the $18-65$ population who is either a nascent entrepreneur or an owner-manager of a new business. The plot pools all available countries and years $2001-2016$. Sources: World Bank, Enterprise Survey. Global Entrepreneurship Monitor.

As a caveat, financial frictions are not the entire picture. As one can see from Fig.11, there is a difference between high- and low-income countries unexplained by the financial frictions. Indeed, as Gollin [2008], Seshadri and Roys [2014], Poschke [2017] discuss, various explanations can be consistent with the negative correlation between GDP per capita and entrepreneurship. That being said, since the financial friction λ in section 6.1 is a reduced form formulation, it can also be interpreted as anything that prevents firm's expansion. In this sense, the unexplained variation does not necessarily limit the scope of the insights.

7 Capital accumulation

So far, I have developed analyses in static environments. In this section, I extend the analysis to a tractable dynamic environment with inter-temporal optimization. It is shown that SDPs are further justified in a dynamic environment, so the analyses in the static environments are robust.

To study the implication of inter-temporal trade-off in a parsimonious and tractable dynamic model, I extend the laissez-faire economy by introducing capital accumulation a la Krebs [2003] and Kreps-Porteus-Epstein-Zin preference. (Epstein and Zin [1989], Kreps

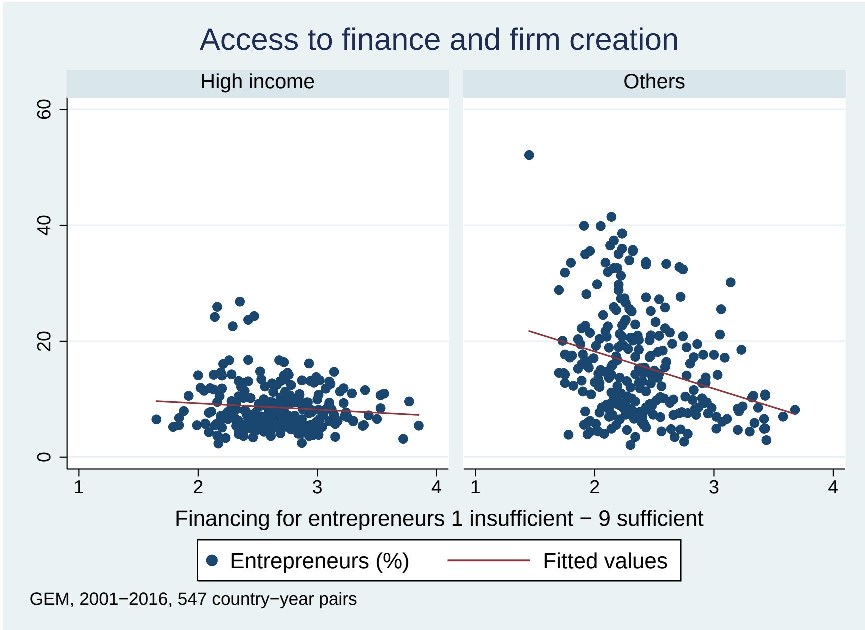

Figure 11: Entrepreneurship is negatively correlated with the availability of financing. x axis is the survey responses about the availability of financial resources - equity and debt - for SMEs including grants and subsidies. y axis is the percentage of the $18-65$ population who is either a nascent entrepreneur or an owner-manager of a new business. Country year pooled data. Sources: World Bank, Enterprise Survey. Global Entrepreneurship Monitor.

and Porteus [1978]) The "all-purpose" good a la Krebs [2003] gives dynamic tractability by invoking a similar capital accumulation structure as Brock and Mirman [1972]. The Kreps-Porteus-Epstein-Zin preference isolates the inter-temporal preference from risk aversion so that I can highlight the implications specific to dynamic environments.

7.1 Equilibrium

This section defines the equilibrium of the laissez-faire economy in a tractable dynamic environment. I derive the closed-form solution of the equilibrium objects and discuss the similarity to the static model.

The setup of the dynamic economy is a direct extension of the static laissez-faire economy. There is a continuum of agents facing risky occupational choice and dynamic consumption maximization problem. The economy has no aggregate uncertainty.

At the beginning of each period, each agent $i \in [0,1]$ is characterized by the $(a_{it-1}, j_{it-1}, z_{it})$ consisting of asset $a_{it-1} > 0$, occupation $j_{it-1} \in \{E, W\}$ where E denotes entrepreneurs and W workers, and productivity $z_{it} > 0$. The distribution of the state at t is denoted by $S_t(a_{it-1}, j_{it-1}, z_{it})$ and the initial distribution S_0 is exogenously given.

I make three comments about the state. First, the asset a_{it} is an "all-purpose good" because it can be used as a consumption good, physical capital and human capital. Second, for simplicity I assume the productivity $\{z_{it}\}_{i,t}$ is *i.i.d.* over i and t, distributed according to G_z. I assume *i.i.d.* structure for simplicity, but it is straightforward to extend it to a persistent process that corresponds to the model with the intermediate signal in the static environments. These assumptions give tractability as will be clear later. Third, the convention of the time subscript is based on the measurability. In other words, the asset and the occupation are pre-determined at $t-1$ before observing the productivity z_{it}. Therefore, in principle, both the saving and occupation decisions involve risk takings. However, as will be discussed later, the log inter-temporal preference makes sure the only decision affected by entrepreneurial risks is the occupation choice.

Within each period t, agents work first, and then make consumption/saving and occupation decisions. For the working part, each agent transforms the all-purpose asset a_{it-1} into productive capital $f_I(a_{it-1})$ to earn income. Each entrepreneur $i \in E_t$ uses the capital $k_{it-1} = f_I(a_{it-1})$ to run her firm of productivity z_{it}. In particular, given wage w_t, the entrepreneur i solves the profit-maximization problem to obtain $\pi_{it} = \pi(z_{it}, w_t, k_{it-1})$ by hiring $n_{it} = n(z_{it}, w_t, k_{it-1})$ employees

$$[\pi(z,w,k), n(z,w,k)] = \max_{n\geq 0} zf(k,n) - wn. \tag{26}$$

Each worker $i \in W_t$ sells his human capital $h_{it-1} = f_I(a_{it-1})$ to earn labor income $w_t h_{it-1}$.

After obtaining income, each agent i makes consumption-saving decision and chooses whether to become an entrepreneur or a worker next period. The consumption-saving decision (c_{it}, a_{it}) has to be chosen from the budget set

$$\begin{aligned} B_t(a,E,z) &= \{(c,a') : k = f_I(a), c + a' = \pi(z, w_t, k)\} \\ B_t(a,W,z) &= \{(c,a') : h = f_I(a), c + a' = w_t h\}. \end{aligned} \tag{27}$$

The occupation choice can be a mixed strategy $e_{it} \in [0,1]$ as in the static case, i.e., if an agent chooses e_{it}, she becomes an entrepreneur $j_{it} = E$ with probability e_{it}. Each agent i chooses $(c_{it}, a_{it}, e_{it}) \in B_t(a_{it-1}, j_{it-1}, z_{it}) \times [0,1]$ for all states to maximize the recursive utility

$$V_{it} = u_I^{-1}\left[(1-\beta)u_I(c_{it}) + \beta u_I\left(u^{-1}(\mathbb{E}u(V_{it+1}))\right)\right] \tag{28}$$

where u and u_I are intra- and inter-termporal utility functions, and the expectation $\mathbb{E}$ is taken with respect to both productivity $z_{it} > 0$ and the mixed strategy $e_{it} \in [0,1]$.

Finally, the labor market has to clear, i.e., labor demand from entrepreneurs equals to human capital supply from workers

$$\int e_{it-1}n_{it}di = \int (1 - e_{it-1})\, h_{it-1}di. \tag{29}$$

One can rewrite this equation using the distribution of states S_t, but I avoid the complication and define the equilibrium concept using the simple summation over i.

Definition 5. Fix the fundamentals $\mathcal{E} = (\beta, u, u_I, f, f_I, G_z, S_0)$. The set of individual choices and wage $\{V_{it}, a_{it}, c_{it}, e_{it}, h_{it}, k_{it}, \pi_{it}, n_{it}, w_t\}_{i,t}$ constitutes an equilibrium if the following conditions are satisfied for all $t \geq 0$.

1. At each t, given w_t, the profit and the employment decision $\{\pi_{it}, n_{it}\}$ by the entrepreneur with productivity z_{it} solves the profit maximization problem (26).

2. Given $\{w_t\}_t$, the value and policies $\{V_{it}, a_{it}, c_{it}, e_{it}\}_{i,t}$ maximize individual welfare (28) subject to $(c_{it}, a_{it}, e_{it}) \in B_t\,(a_{it-1}, j_{it-1}, z_{it}) \times [0, 1]$.

3. At each period t, labor market clears (29).

The equilibrium number of firms at t is pre-determined at $t-1$ and can be written as

$$\phi_{t-1} = \int e_{it-1}di.$$

This is a dynamic general equilibrium model with heterogeneous agents and incomplete markets, so in general, a closed form solution is not available. However, the following parameterization allows us to study the efficiency of firm creation analytically

$$u\,(c) = \frac{c^{1-\gamma} - 1}{1-\gamma},\;\; u_I\,(c) = \ln c,\;\; f\,(k, n) = k^{1-\alpha}n^{\alpha},\;\; f_I\,(a) = a^{\theta} \tag{30}$$

where $\gamma > 0$, $\alpha \in (0, 1)$ and $\theta \in (0, 1)$. The intra-temporal utility u remains the same as the static model. The logarithmic inter-temporal utility u_I separates the saving decision from the uncertainty about the productivity z_{it}. This is because the optimal saving is a constant fraction of the income and is independent of the potentially risky returns to saving. In this way, I can make sure that the only risky decision is occupation choice so that the result is directly comparable with the static environment. The production function is the standard Cobb-Douglas form with complementarity between the entrepreneurial capital and the workers' human capital. As a result, the profit becomes a linear function of entrepreneurial capital $\pi\,(z, w, k) = r\,(z, w)\,k$, and therefore can be interpreted as the return to entrepreneurial capital. The investment function features decreasing returns to scale to ensure a stationary asset distribution.

Given the parameterization, the market equilibrium takes the following form as derived in Appendix A.6

$$
\begin{aligned}
w_t^{LF} &= \left(\mathbb{E}z^{\frac{(1-\gamma)(1-\beta)}{(1-\alpha)(1-\beta\theta)}}\right)^{\frac{(1-\alpha)(1-\beta\theta)}{(1-\gamma)(1-\beta)}} \alpha^{\alpha}(1-\alpha)^{1-\alpha}, \\
\phi_t^{LF} &= \frac{1-\alpha}{1-\alpha+\alpha\mathbb{E}z^{\frac{1}{1-\alpha}}\left(\mathbb{E}z^{\frac{(1-\gamma)(1-\beta)}{(1-\alpha)(1-\beta\theta)}}\right)^{-\frac{1-\beta\theta}{(1-\gamma)(1-\beta)}}}.
\end{aligned}
\tag{31}
$$

One can see that the risk aversion $1-\gamma$ is always multiplied by $\frac{1-\beta}{1-\beta\theta} > 1$. To anticipate the result, this additional term comes from the inter-temporal decision and makes the agents effectively more risk-averse so that ϕ_t^{LF} is further away from the efficient level $1-\alpha$.

7.2 Efficiency analysis

This section presents the efficiency analysis of the dynamic model. The direction of the market failure remains the same as in the static economy but the severity is greater due to the consumption-smoothing motive. A novel insight is that the market failure is more severe in an economy with more patient agents.

The planner maximizes the welfare by choosing resource allocations at each state $\omega_t \in \Omega_t = \left([z_{min}, z_{max}]^{[0,1]}\right)^t$ facing the same technology and information constraints as the individual agent. Specifically, given the Pareto weight function $i \mapsto \Lambda_i$, the planner maximizes the welfare $\int V_{it}(\omega)\, d\Lambda_i$ where $V_{it}(\omega)$ satisfies

$$
V_{it}(\omega_t) = u_I^{-1}\left[(1-\beta)\, u_I(c_{it}(\omega_t)) + \beta u_I\left(u^{-1}\left(\mathbb{E}_{\Omega_{t+1}|\omega_t} u(V_{it+1})\right)\right)\right] \tag{32}
$$

and the expectation $\mathbb{E}_{\Omega_{t+1}|\omega_t}$ is taken over the subset of state space Ω_{t+1} accessible from ω_t. Since the problem reduces to the aggregate production maximization in an economy with no aggregate uncertainty, I omit the notation of the state ω and related measure theoretic treatment in the following.

The planner's decision can be described recursively. At each period t, the planner takes as given the set of entrepreneurs $E_{t-1} \subset [0,1]$ and workers $W_{t-1} = [0,1] \setminus E_{t-1}$ chosen at period $t-1$, and their saving of capital $\{a_{it-1}\}_i$. The capital generates entrepreneurial capital $k_{it-1} = f_I(a_{it-1})$ if stored by an entrepreneur $i \in E_{t-1}$ and human capital $h_{it-1} = f_I(a_{it-1})$ if stored by a worker $i \in W_{t-1}$. The human capital is then reallocated by the planner to make production

$$
Y_t = \int_{i\in E_{t-1}} z_{it} f(k_{it-1}, n_{it})\, di, \quad \int_{i\in E_{t-1}} n_{it} di = \int_{i\in W_{t-1}} h_{it-1} di. \tag{33}
$$

After the production, the planner chooses the occupation E_t, W_t for period $t+1$ and divides the aggregate output into consumption today and asset saved for tomorrow

$$\int c_{it} di + \int a_{it} di = Y_t. \tag{34}$$

The key informational assumption is that when the planner chooses the asset a_{it}, the productivity tomorrow z_{it+1} is not observable

$$a_{it} \perp\!\!\!\perp z_{it+1}. \tag{35}$$

In summary, the efficient allocation can be defined as follows.

Definition 6. Fix the fundamentals $\mathcal{E}$. The set of individual values, consumption-saving, occupation, and production decisions and the aggregate output $\{V_{it}, c_{it}, a_{it}, E_t, W_t, k_{it}, h_{it}, n_{it}, Y_t\}$ is efficient if it maximizes (32) subject to (34), (35), and (33). The number of firms in the planner's solution is defined as

$$\phi_{t-1}^P = \int_{i \in E_{t-1}} di.$$

It is immediate to see that the efficient allocation maximizes the aggregate output Y_t at each period given the asset $\{a_{it-1}\}$. Since the planner can control all the allocations, as in the static case, the planner's problem can be separated into production and allocation of consumption.

The following proposition states that the direction of the market failure remains the same as the static case, but the market failure gets worse due to the inter-temporal consumption smoothing motive.

Proposition 7. *The laissez-faire economy generates an insufficient number of firms if and only if there are entrepreneurial risks.*

$$\phi_t^{LF} < \phi_t^P = 1 - \alpha \Leftrightarrow V(z) > 0.$$

This is true even when all agents are risk neutral $\gamma = 0$.

Proof. See Appendix A.6. □

The inefficiency result can be extended to constrained inefficiency as discussed in Appendix F.3.

I make two observations about the statement. First, note that the planner's solution $\phi_t^P = 1 - \alpha$ is identical to the static model, making the static and dynamic models directly comparable. Together with Eq.(31), one can see that the inefficiency in the dynamic environment is worse than the static, and therefore, SDPs are more effective in a dynamic environment. In this sense, the results in the static analyses are conservative.

Second, the last line presents a new insight specific to the dynamic environment. It states that the market failure happens even when all agents are risk-neutral. This is not the case in the static environment. If agents are risk-neutral, both the laissez-faire and planner choose $\phi = 1 - \alpha$ in the static environment, making Pareto improvement impossible. The intuition is based on the consumption-smoothing motive. Since agents want to smooth consumption, they want to avoid asset fluctuation. To avoid the asset fluctuation, they can choose the salary job with a safer return. This inter-temporal motive is conceptually different from the static risk aversion at least in the framework of Kreps-Porteus-Epstein-Zin preference. Therefore, the number of firms is smaller than the efficient level even when all agents are risk-neutral $\gamma = 0$.

In terms of the intuition, observe that the consumption-smoothing motive works similarly as risk aversion. This is not surprising if one realizes that the intra-temporal and inter-temporal preferences both discount the value of fluctuations across a set of states. Mathematically, they both represent the force of Jensen's inequality over the states called in different ways.

Finally, one non-trivial insight can be obtained about the patience β. As can be seen from Eq.(31), the more patient agents are $\beta \nearrow$, the more severe the market failure is $\phi^P - \phi^{LF} \nearrow$. This is because patient agents care more about the fluctuation of the future consumption. In other words, higher patience magnifies consumption-smoothing motive. Therefore, in high patience economy, removal of SDPs can cause severe human resource misallocation.

8 Conclusion

I have shown that SDPs might enhance efficiency in the presence of entrepreneurial risks. The analysis not only changes the conventional wisdom about SDPs but also casts a caveat in general to the misallocation measurement exercises based on marginal product equalization. I have also made the analysis comprehensive by showing that severe financial frictions can cause excessive firm creation even if firm creation is risky. This result implies that developing countries with severe financial frictions might require different policy prescriptions. Finally, I have analyzed the implications of dynamic decisions. The analysis not only confirms the robustness of the findings in the static environment but also provides a novel insight into the role of consumption-smoothing motive and patience. These analyses deepen our understanding of SDPs.

I conclude the paper by mentioning three caveats to my analyses. First, there can be many other justifications for SDPs. For instance, small firms might have some externality in R&D and employment, although these arguments are highly controversial compared to entrepreneurial risks (Biggs [2002], Shane [2008, 2009]). Another justification for SDPs is that it saves the government administrative costs as discussed in Kaplow [2017]. My analysis should be considered a complement to these discussions. Second, I have assumed that individual agent maximizes profit, but the importance of non-pecuniary benefit has also been pointed out in the entrepreneurship literature (Hurst and Pugsley [2011], Gordon and Sarada [2017]). Exploring the implications of behavioral agents that have other objectives than profit maximization is an interesting future topic. Third, I have focused on SDPs that treat small firms preferentially. However, some countries treat large firms preferentially to foster international competitiveness. In this case, SDPs are adopted based on a totally different logic and therefore, their analysis needs a different framework. These points are missing in the analysis and should be explored in future research.

References

Sakai Ando and Misaki Matsumura. Constrained Efficiency of Competitive Entrepreneurship. Working Paper, 2017.

George-Marios Angeletos and Alessandro Pavan. Efficient Use of Information and Social Value of Information. *Econometrica*, 75(4):1103–1142, July 2007.

Tyler Biggs. Is Small Beautiful and Worthy of Subsidy? Literature Review. January 2002.

Saki Bigio and Jennifer La'O. Financial Frictions in Production Networks. Working Paper 22212, National Bureau of Economic Research, April 2016.

William A Brock and Leonard J Mirman. Optimal Economic Growth and Uncertainty: The Discounted Case. *Journal of Economic Theory*, 4(3):479–513, June 1972.

Francisco J. Buera, Joseph P. Kaboski, and Yongseok Shin. Entrepreneurship and Financial Frictions: A Macrodevelopment Perspective. *Annual Review of Economics*, 7(1):409–436, 2015.

Peter A. Diamond. The Role of a Stock Market in a General Equilibrium Model with Technological Uncertainty. *The American Economic Review*, 57(4):759–776, 1967.

Larry G. Epstein and Stanley E. Zin. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica*, 57(4):937–969, 1989.

Luis Garicano, Claire Lelarge, and John Van Reenen. Firm Size Distortions and the Productivity Distribution: Evidence from France. *American Economic Review*, 106(11):3439–3479, November 2016.

Douglas Gollin. Nobody's Business but My Own: Self-Employment and Small Enterprise in Economic Development. *Journal of Monetary Economics*, 55(2):219–233, March 2008.

Roger H. Gordon and Sarada. How Should Taxes Be Designed to Encourage Entrepreneurship? January 2017.

Piero Gottardi, Atsushi Kajii, and Tomoyuki Nakajima. Constrained Inefficiency and Optimal Taxation with Uninsurable Risks. *Journal of Public Economic Theory*, 18(1):1–28, February 2016.

Joshua D. Gottlieb, Richard R. Townsend, and Ting Xu. Does Career Risk Deter Potential Entrepreneurs? SSRN Scholarly Paper ID 2714577, Social Science Research Network, Rochester, NY, January 2018.

Francois Gourio and Nicolas Roys. Size-Dependent Regulations, Firm Size Distribution, and Reallocation. *Quantitative Economics*, 5(2):377–416, July 2014.

Nezih Guner, Gustavo Ventura, and Yi Xu. Macroeconomic Implications of Size-Dependent Policies. *Review of Economic Dynamics*, 11(4):721–744, October 2008.

Barton H. Hamilton. Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment. *Journal of Political Economy*, 108(3):604–631, June 2000.

Johan Hombert, Antoinette Schoar, David Sraer, and David Thesmar. Can Unemployment Insurance Spur Entrepreneurial Activity? Technical Report w20717, National Bureau of Economic Research, November 2014.

Hugo A. Hopenhayn. Entry, Exit, and firm Dynamics in Long Run Equilibrium. *Econometrica*, 60(5):1127–1150, 1992.

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

Chang-Tai Hsieh, Erik Hurst, Charles I. Jones, and Peter J. Klenow. The Allocation of Talent and U.S. Economic Growth. Technical Report w18693, National Bureau of Economic Research, January 2013.

Erik Hurst and Benjamin Wild Pugsley. What Do Small Businesses Do? *Brookings Papers on Economic Activity*, page 73, September 2011.

Roberto Fattal Jaef. Constrained Optimality and the Welfare Effects of Misallocation. *2012 Meeting Papers*, 2012.

S. M. Kanbur. Impatience, Information and Risk Taking in a General Equilibrium Model of Occupational Choice. *The Review of Economic Studies*, 46(4):707–718, 1979.

S. M. Kanbur. Risk Taking and Taxation. *Journal of Public Economics*, 15(2):163–184, April 1981.

Louis Kaplow. Optimal Regulation with Exemptions. Working Paper 23887, National Bureau of Economic Research, September 2017.

Richard E. Kihlstrom and Jean-Jacques Laffont. A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion. *Journal of Political Economy*, 87(4): 719–748, 1979.

Tom Krebs. Human Capital Risk and Economic Growth. *The Quarterly Journal of Economics*, 118(2):709–744, May 2003.

David M. Kreps and Evan L. Porteus. Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica*, 46(1):185–200, 1978.

Ross Levine and Yona Rubinstein. Smart and Illicit: Who Becomes an Entrepreneur and Do They Earn More? *The Quarterly Journal of Economics*, 132(2):963–1018, May 2017.

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

N. Gregory Mankiw and Michael D. Whinston. Free Entry and Social Inefficiency. *The RAND Journal of Economics*, 17(1):48–58, 1986.

Gustavo Manso. Experimentation and the Returns to Entrepreneurship. *The Review of Financial Studies*, 29(9):2319–2340, September 2016.

J. A. Mirrlees. An Exploration in the Theory of Optimum Income Taxation. *The Review of Economic Studies*, 38(2):175–208, 1971.

Tobias J. Moskowitz and Annette Vissing-Jorgensen. The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle? *The American Economic Review*, 92(4): 745–778, 2002.

Markus Poschke. The Firm Size Distribution Across Countries and Skill-Biased Change in Entrepreneurial Technology. Working Paper, 2017.

Vincenzo Quadrini. Entrepreneurship in Macroeconomics. *Annals of Finance*, 5(3-4):295–311, September 2008.

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

Ananth Seshadri and Nicolas Roys. Economic Development and the Organization of Production. Working Paper, 2014.

Scott Shane. Why Encouraging More People to Become Entrepreneurs Is Bad Public Policy. *Small Business Economics*, 33(2):141–149, August 2009.

Scott A. Shane. *The Illusions of Entrepreneurship : The Costly Myths That Entrepreneurs, Investors, and Policy Makers Live By.* Yale University Press, New Haven, US, 2008.

Kotaro Suzumura and Kazuharu Kiyono. Entry Barriers and Economic Welfare. *The Review of Economic Studies*, 54(1):157–167, January 1987.

Alexis Akira Toda. Asset Prices and Efficiency in a Krebs Economy. *Review of Economic Dynamics*, 18(4):957–978, October 2015.

Harald Uhlig. A Law of Large Numbers for Large Economies. *Economic Theory*, 8(1):41–50, 1996.

Galina Vereshchagina and Hugo A. Hopenhayn. Risk Taking by Entrepreneurs. *American Economic Review*, 99(5):1808–30, December 2009.

Appendices for Online Publication

A Proofs of propositions

I use the following lemma.

Lemma 1. *(Jensen's inequality) Suppose $X>0$ is a random variable with positive variance $VX>0$. Then, for any $\gamma>0$,*

$$\left(\mathbb{E}X^{1-\gamma}\right)^{\frac{1}{1-\gamma}}<\mathbb{E}X.$$

Proof. Since CRRA utility function is strictly concave and increasing when $\gamma>0$, Jensen's inequality and taking inverse function imply

$$\mathbb{E}u\left(X\right)<u\left(\mathbb{E}X\right)\Leftrightarrow u^{-1}\left(\mathbb{E}u\left(X\right)\right)<\mathbb{E}X\Leftrightarrow\left(\mathbb{E}X^{1-\gamma}\right)^{\frac{1}{1-\gamma}}<\mathbb{E}X.$$

□

A.1 Proof of Proposition 2

The planner's solution ϕ^P maximizes the output.

$$Y\left(\phi,n\right)=\max_{n(z)\geq0,\phi\in[0,1]}\phi E\left[zf\left(n\left(z\right)\right)|s\geq G_s^{-1}\left(1-\phi\right)\right]$$

$$s.t.\ \phi+\phi E\left[n\left(z\right)|s\geq G_s^{-1}\left(1-\phi\right)\right]=1.$$

Reformulate the problem using $\bar{s}$ gives

$$Y\left(\bar{s},n\right)=\max_{\bar{s},n(z)\geq0}\int zf\left(n\left(z\right)\right)1_{s\geq\bar{s}}dG\ s.t.\ \int n\left(z\right)1_{s\geq\bar{s}}dG=G_s\left(\bar{s}\right).$$

For each fixed $\bar{s}$, the sub-problem of choosing $n\left(z\right)$ is

$$\max_{n(z)\geq0}\int zf\left(n\left(z\right)\right)g\left(z|s\geq\bar{s}\right)dz\ s.t.\ \int n\left(z\right)g\left(z|s\geq\bar{s}\right)dz=\frac{G_s\left(\bar{s}\right)}{1-G_s\left(\bar{s}\right)}.$$

The objective function is a weighted sum of the strictly concave function of $\{n\left(z\right)\}_z$ and the constraint is a linear function of $\{n\left(z\right)\}_z$. Hence, the unique solution is characterized by the first order conditions.

Furthermore, let $s_{min},s_{max}\in[-\infty,\infty]$ be the smallest and largest values that the signal can take, so that $1_{s\geq s_{min}}=1$ and $1_{s\geq s_{max}}=0$ for all realizations. When $\bar{s}=s_{min}$, the value is $Y\left(s_{min},n\right)=0$ since the only feasible employment schedule is $n\left(z\right)=0$ for all z. When $\bar{s}=s_{max}$, $Y\left(s_{max},n\right)=0$ since the objective function is 0. Since 0 output is the lower bound $Y\geq 0$, the solution $\bar{s}$ is not at the boundaries. Therefore, the solution exists and is unique if there is a unique set of $\left(\bar{s},n\right)$ that solves the first order conditions.

The first order condition can be obtained by taking the derivative of the Lagrangian.

$$\begin{cases} n\left(z\right)=\left(f'\right)^{-1}\left(\frac{\lambda}{z}\right),\ \lambda>0. \\ \mathbb{E}\left[zf\left(n\left(z\right)\right)|s=\bar{s}\right]=\lambda\left(1+\mathbb{E}\left[n\left(z\right)|s=\bar{s}\right]\right) \\ \mathbb{E}\left[n\left(z\right)1_{s\geq\bar{s}}\right]=G_s\left(\bar{s}\right) \end{cases}$$

Combining the conditions gives

$$\begin{cases} F_1\left(\bar{s},\lambda\right)=\mathbb{E}\left[zf\left(\left(f'\right)^{-1}\left(\frac{\lambda}{z}\right)\right)-\lambda\left(f'\right)^{-1}\left(\frac{\lambda}{z}\right)|s=\bar{s}\right]-\lambda=0. \\ F_2\left(\bar{s},\lambda\right)=\mathbb{E}\left[\left(f'\right)^{-1}\left(\frac{\lambda}{z}\right)1_{s\geq\bar{s}}\right]-G_s\left(\bar{s}\right)=0 \end{cases}$$

Note that the integrand of F_1 is identical to the value function of the profit maximization problem with wage λ

$$zf\left(\left(f'\right)^{-1}\left(\frac{\lambda}{z}\right)\right)-\lambda\left(f'\right)^{-1}\left(\frac{\lambda}{z}\right)=\max_{n\geq 0}zf\left(n\right)-\lambda n.$$

Thus, F_1 gives a strictly positive relationship between $\bar{s}$ and λ. In contrast, F_2 gives a strictly negative relationship. Hence, there is a unique solution $\left(\bar{s},n\right)$ to the output maximization problem, and unique solution ϕ^P to the original problem.

A.2 Proof of Proposition 3

For the first statement, when there are no risks, the laissez-faire economy is characterized by $\left(\bar{z}^{LF},w^{LF},\pi^{LF},n^{LF}\right)$ that satisfy

$$\begin{aligned} \pi\left(\bar{z}^{LF},w^{LF}\right)&=w^{LF} \\ zf'\left(n^{LF}\left(z,w^{LF}\right)\right)&=w^{LF} \\ \pi\left(z,w^{LF}\right)&=zf\left(n\left(z,w^{LF}\right)\right)-w^{LF}n\left(z,w^{LF}\right) \\ \mathbb{E}\left[n\left(z,w^{LF}\right)1_{z\geq\bar{z}^{LF}}\right]&=G_z\left(\bar{z}^{LF}\right) \end{aligned}$$

Choosing the Pareto weight using these variables.

$$\lambda(z)=\begin{cases}\frac{1}{u'(w^{LF})} & z\leq\bar{z}\\ \frac{1}{u'(\pi(z,w^{LF}))} & z>\bar{z}\end{cases}$$

I show that $(\bar{z},w,\pi,n)$ is the solution to the planner's problem. Without loss of generality, $c_i(\omega)=c(z_i)$. The planner's solution can be written as

$$\max_{c,n,Y,\bar{z}}\int\lambda(z)\,u(c(z))\,dG_z\ s.t.\ \begin{cases}\int c(z)\,dG_z=Y\\ Y=\mathbb{E}\left[zf(n(z))\,1_{z\geq\bar{z}}\right]\\ \mathbb{E}\left[n(z)\,1_{z\geq\bar{z}}\right]=G_z(\bar{z})\end{cases}$$

The Lagrangian is

$$L\ =\ \int\lambda(z)\,u(c(z))\,dG_z\ +\ \mu\left\{\mathbb{E}\left[zf(n(z))\,1_{z\geq\bar{z}}\right]-\mathbb{E}c(z)+\nu\left(G_z(\bar{z})-\mathbb{E}\left[n(z)\,1_{z\geq\bar{z}}\right]\right)\right\}$$

The FOCs are

$$\begin{cases}L_c=0 & \Rightarrow u'(c(z))=\frac{\mu}{\lambda(z)}\\ L_n=0 & \Rightarrow zf'(n(z))=\nu\\ L_{\bar{z}}=0 & \Rightarrow \bar{z}f(n(\bar{z}))-vn(\bar{z})=\nu\end{cases}$$

$(\mu,\bar{z},v,n)=\left(1,\bar{z}^{LF},w,n^{LF}\right)$ and $c(z)=\pi^{LF}\left(z,w^{LF}\right)1_{z\geq\bar{z}^{LF}}+w1_{z<\bar{z}^{LF}}$ satisfies the equilibrium conditions.

For the second statement it suffices to show that the laissez-faire economy does not maximize output. With risks, the laissez-faire economy is characterized by

$$\mathbb{E}\left[u\left(zf\left(h\left(\frac{w}{z}\right)\right)-wh\left(\frac{w}{z}\right)\right)|s=\bar{s}\right]=u(w)$$
$$\mathbb{E}\left[h\left(\frac{w}{z}\right)1_{s\geq\bar{s}}\right]=G_s(\bar{s})$$

where $h(x):=(f')^{-1}(x)$. The FOCs of the planner's problem

$$\max_{\bar{s},n(z)}\int zf(n(z))\,1_{s\geq\bar{s}}dG\ s.t.\ \int n(z)\,1_{s\geq\bar{s}}dG=G_s(\bar{s}),$$

using the Lagrange multiplier λ, is

$$\mathbb{E}\left[zf\left(h\left(\frac{\lambda}{z}\right)\right)-\lambda h\left(\frac{\lambda}{z}\right)|s=\bar{s}\right]=\lambda$$
$$\mathbb{E}\left[h\left(\frac{\lambda}{z}\right)1_{s\geq\bar{s}}\right]=G_s\left(\bar{s}\right).$$

Note that the two systems are different by u. Therefore, the laissez-faire economy can Pareto improve by producing more and redistribute them.

The two systems also imply $w<\lambda$. To see this, note that Jensen's inequality implies

$$\lambda=\mathbb{E}\left[zf\left(h\left(\frac{\lambda}{z}\right)\right)-\lambda h\left(\frac{\lambda}{z}\right)|s=\bar{s}\right]>u^{-1}\left(\mathbb{E}\left[u\left(zf\left(h\left(\frac{\lambda}{z}\right)\right)-\lambda h\left(\frac{\lambda}{z}\right)|s=\bar{s}\right)\right]\right).$$

In addition,

$$h'\left(\lambda\right)<0,\ \frac{\partial}{\partial\lambda}\left\{zf\left(h\left(\frac{\lambda}{z}\right)\right)-\lambda h\left(\frac{\lambda}{z}\right)\right\}=-h\left(\frac{\lambda}{z}\right)<0.$$

so from the market clearing condition, λ and $\bar{s}$ have to be negatively correlated. Hence, $w<\lambda$ and $\bar{s}^P<\bar{s}^{LF}$, which then implies $\phi^{LF}<\phi^P$.

A.3 Proof of Proposition 4

First, I reformulate the optimal taxation problem into a mechanism design problem of choosing a set of menu.

$$\max_{w,C(z),Y(z)} w\ s.t.\ \begin{cases}\mathbb{E}u\left(zf\left(C\left(z\right)\right)-wC\left(z\right)-Y\left(z\right)\right)=u\left(w\right)\\ U\left(C\left(z\right),Y\left(z\right),z\right)\geq U\left(C\left(z'\right),Y\left(z'\right),z\right)\ \forall z,z'\\ \mathbb{E}Y\left(z\right)-\mathcal{T}\left(1+\mathbb{E}n\left(z\right)\right)=0.\end{cases}\tag{36}$$

The feasible sets of (21) and (36) are identical. To see this, fix w and choose a feasible solution $\left(n\left(z\right),T\left(n\right)\right)$ from (21). Then, $\left(C\left(z\right),Y\left(z\right)\right)=\left(n\left(z\right),T\left(n\left(z\right)\right)\right)$ satisfy

$$U\left(C,Y,z\right):=zf\left(C\right)-wC-Y$$

$$U\left(C\left(z\right),Y\left(z\right),z\right)\geq U\left(C\left(z'\right),Y\left(z'\right),z\right)\ \forall z,z'$$

To see this, pick arbitrary z and $z' \in [z_{min}, z_{max}]$. Then,

$$\begin{aligned} zf\left(C\left(z\right)\right)-wC\left(z\right)-Y\left(z\right) = & zf\left(n\left(z\right)\right)-wn\left(z\right)-T\left(n\left(z\right)\right) \\ \geq & zf\left(n\left(z'\right)\right)-wn\left(z'\right)-T\left(n\left(z'\right)\right) \\ = & zf\left(C\left(z'\right)\right)-wC\left(z'\right)-Y\left(z'\right). \end{aligned}$$

The opposite is also true. Given a $\left(C\left(z\right), Y\left(z\right)\right)$ that satisfies the truth-telling condition (36), the SDP can be constructed by

$$T\left(n\right)=\inf_{\bar{T}}\left\{\bar{Y}:U\left(C\left(z\right),Y\left(z\right),z\right)\geq U\left(n,\bar{Y},z\right)\forall z\right\}.$$

By the truth-telling condition,

$$U\left(C\left(z\right),Y\left(z\right),z\right)=U\left(C\left(z\right),T\left(C\left(z\right)\right),z\right)\geq U\left(n,T\left(n\right),z\right)\forall z,n.$$

Hence, agents facing $T\left(\cdot\right)$ chooses $n\left(z\right)=C\left(z\right)$.

Next, I reformulate the truth-telling condition into the envelope condition and the monotonicity constraint. Note that the utility function $U\left(C,Y,z\right)$ satisfies Spence-Mirrlees condition

$$MRS\left(C,Y,z\right)=-\frac{U_Y}{U_C}=\frac{1}{zf'\left(C\right)-w},\ MRS_z=-\frac{f'\left(C\right)}{\left(zf'\left(C\right)-w\right)^2}<0.$$

Hence, the truth-telling condition is equivalent to

$$\pi\left(z\right):=U\left(C\left(z\right),Y\left(z\right),z\right)$$

$$\pi'\left(z\right)=U_z\left(C\left(z\right),Y\left(z\right),z\right),\ C'\left(z\right)\geq 0$$

By substituting Y out, the problem becomes

$$\max_{w,n(z),\pi(z)} w\ s.t.\ \begin{cases}\mathbb{E}u\left(\pi\left(z\right)\right)-u\left(w\right)=0\\ \pi'\left(z\right)=f\left(n\left(z\right)\right)\\ n'\left(z\right)\geq 0\\ \mathbb{E}\left[zf\left(n\left(z\right)\right)-wn\left(z\right)-\pi\left(z\right)\right]-\mathcal{T}\left(1+\mathbb{E}n\left(z\right)\right)=0.\end{cases} \tag{37}$$

A.4 Proof of Proposition 5

To solve (37) using optimal control theory, rewrite the integral constraints by

$$B'\left(z\right)=\left\{zf\left(n\left(z\right)\right)-wn\left(z\right)-\pi\left(z\right)-\mathcal{T}\left(1+n\left(z\right)\right)\right\}g\left(z\right),\ B\left(z_{min}\right)=B\left(z_{max}\right)=0.$$

$$I'(z) = (u(\pi(z)) - u(w))\, g(z)\,,\;\; I(z_{min}) = I(z_{max}) = 0.$$

Wage is constant, so $w'(z) = 0$. Monotonicity constraint becomes $n'(z) = L(z) \geq 0$. The only control variable is L, state variables are π, w, B, I, n, and the co-state variables are $\mu_\pi, \mu_w, \mu_B, \mu_I, \mu_n$. Hamiltonian is

$$\begin{aligned} H = {} & wg(z) + \mu_\pi f(n) + \mu_w \times 0 + \mu_n L + \kappa L \\ & + \mu_B \{ zf(n) - wn - \pi - \mathcal{T}(1+n) \} g(z) + \mu_I (u(\pi) - u(w))\, g(z) \end{aligned}$$

$$H_L = 0 = \mu_n + \kappa$$

$$H_\pi = -\mu_\pi' = (-\mu_B + \mu_I u'(\pi))\, g(z)$$

$$H_w = -\mu_w' = (1 - \mu_B n - \mu_I u'(w))\, g(z)$$

$$H_B = -\mu_B' = 0$$

$$H_I = -\mu_I' = 0$$

$$H_n = -\mu_{n'} = \mu_\pi f'(n) + \mu_B \{ zf'(n) - w - \mathcal{T} \} g(z)$$

$$\kappa L = 0,\;\; \kappa \geq 0,\;\; L \geq 0.$$

Boundary conditions are

$$\begin{aligned} B(z_{min}) = B(z_{max}) = I(z_{min}) = I(z_{max}) = \mu_\pi(z_{min}) = \mu_\pi(z_{max}) \\ = \mu_w(z_{min}) = \mu_w(z_{max}) = \mu_n(z_{min}) = \mu_n(z_{max}) = 0 \end{aligned}$$

By reducing the system, the solution $(\pi, w, \mu_B, \mu_I, \mu_\pi, \mu_w, I, B, n, \mu_n)$ satisfy

$$\pi' = f(n)$$

$$w' = \mu_B' = \mu_I' = 0$$

$$\mu_\pi' = (\mu_B - \mu_I u'(\pi))\, g(z)$$

$$\mu_w' = (\mu_B n + \mu_I u'(w) - 1)\, g(z)$$

$$B'(z) = \{ zf(n(z)) - wn(z) - \pi(z) - \mathcal{T}(1 + n(z)) \} g(z)$$

$$I'(z) = (u(\pi(z)) - u(w))\, g(z)$$

$$\mu_n' = -\mu_\pi f'(n) - \mu_B \{ zf'(n) - w - \mathcal{T} \} g(z)$$

$$B(z_{min}) = B(z_{max}) = I(z_{min}) = I(z_{max}) = \mu_\pi(z_{min}) = \mu_\pi(z_{max})$$
$$= \mu_w(z_{min}) = \mu_w(z_{max}) = \mu_n(z_{min}) = \mu_n(z_{max}) = 0$$

$$\mu_n n' = 0,\ \mu_n \leq 0,\ n' \geq 0.$$

By assumption, $n' > 0$. Then, from $\mu_n' = 0$, $n(z)$ is determined by

$$zf'(n(z)) = \frac{w + \mathcal{T}}{\frac{\mu_\pi(z)}{z\mu_B g(z)} + 1} \tag{38}$$

In this case, ODE reduced to $y = (\pi, w, \mu_B, \mu_I, \mu_\pi, \mu_w, B, I)$ such that

$$\pi' = f(n)$$

$$w' = \mu_B' = \mu_I' = 0$$

$$\mu_\pi' = (\mu_B - \mu_I u'(\pi))\, g(z)$$

$$\mu_w' = (\mu_I u'(w) + \mu_B n - 1)\, g(z)$$

$$B'(z) = \{zf(n) - wn - \pi - \mathcal{T}(1+n)\}\, g(z)$$

$$I'(z) = (u(\pi) - u(w))\, g(z)$$

$$B(z_{min}) = B(z_{max}) = I(z_{min}) = I(z_{max}) = \mu_\pi(z_{min}) = \mu_\pi(z_{max}) = \mu_w(z_{min}) = \mu_w(z_{max}) = 0$$

To show that the medium-sized firms are distorted, it suffices to show $\mu_B > 0 > \mu_\pi(z)$ due to 38. To see this, note that from envelope theorem,

$$\frac{\partial H}{\partial \mathcal{T}} = -\mu_B \mathbb{E}(1 + n(z)).$$

Since higher tax takes more resources, $\frac{\partial H}{\partial \mathcal{T}} < 0$ so $\mu_B > 0$. By integrating μ_π', the boundary condition implies

$$\mu_B - \mu_I \mathbb{E} u'(\pi(z)) = 0 \Rightarrow \mu_I > 0.$$

Now, let's focus on μ_π'. which has to cross zero since $\mu_\pi(z_{min}) = \mu_\pi(z_{max}) = 0$. Since $\pi' > 0$, μ_π must be convex,

$$\frac{\partial \mu_\pi'(z)}{\partial z} = -\mu_I u''(\pi(z))\, \pi'(z) > 0.$$

Therefore, μ_π' starts from negative and monotonically cross zeros. Hence, $\mu_\pi < 0$.

The fact that the optimal SDP increases the number of firms follows from the distortion. Note that when $\mathcal{T} = 0$, the laissez-faire economy $T(n) = 0$ is a feasible choice for the government. Therefore, the optimal SDP increases wage $w^o > w^{LF}$. For small $\mathcal{T} > 0$, the marginal products are higher than the laissez-faire wage

$$zf'(n(z)) = \frac{w^o + \mathcal{T}}{\frac{\mu_\pi(z)}{z\mu_B g(z)} + 1} > w^{LF}.$$

Hence, employment is lower at each productivity level. Since the entrepreneurs and workers have to add up to one, it implies that the number of firms increases. $\phi^o > \phi^{LF}$.

A.5 Proof of Proposition 6

The equilibrium is characterized by

$$\mathbb{E}u(\pi(z,w)) = u(w)$$

$$\pi(z,w) = \max_n zn^\alpha - wn \ s.t. \ wn \leq \lambda zn^\alpha.$$

$$\phi \mathbb{E}n(z,w) = 1 - \phi$$

Let $\lambda^* = \min\{\alpha, \lambda\}$. The firm's problem gives

$$n(z,w) = \left(\frac{\lambda^* z}{w}\right)^{\frac{1}{1-\alpha}}, \ \pi(z,w) = (1-\lambda^*)\left\{z\left(\frac{\lambda^*}{w}\right)^\alpha\right\}^{\frac{1}{1-\alpha}}$$

The indifference condition leads to

$$w^{LF} = (1-\lambda^*)^{1-\alpha}(\lambda^*)^\alpha\left(\mathbb{E}z^{\frac{1-\gamma}{1-\alpha}}\right)^{\frac{1-\alpha}{1-\gamma}}$$

The market clearing condition implies that the equilibrium number of firms is

$$\phi^{LF} = \frac{1-\lambda^*}{1-\lambda^* + \lambda^*\mathbb{E}z^{\frac{1}{1-\alpha}}\left(\mathbb{E}z^{\frac{1-\gamma}{1-\alpha}}\right)^{-\frac{1}{1-\gamma}}}.$$

This is a decreasing function of λ^* satisfying $\phi^{LF} \to 1$ as $\lambda \to 0$. Since the planner does not face any friction, ϕ^P does not depend on λ. Hence, for small λ, $\phi^P < \phi^{LF}$.

A.6 Proof of Proposition 7

I first solve the laissez-faire economy and then derive the efficient allocation.

(Laissez-faire) Let $V_t(a,j,z)$ be the value function. Guess $V_t(a,j,z) = v_t^j(z)\,a^\rho$. For instance, the problem for entrepreneurs is

$$v_t^E(z)\,a^\rho = \max_{c,e'} \exp\Big[(1-\beta)\ln c + \beta\rho \ln\left(r_t a^\theta - c\right) \\ +\beta \ln\left\{e'\mathbb{E}v_{t+1}^E(z')^{1-\gamma} + (1-e')\,\mathbb{E}v_{t+1}^W(z')^{1-\gamma}\right\}^{\frac{1}{1-\gamma}}\Big]$$

The maximization in terms of c and coefficient matching with respect to a gives

$$c = \frac{1-\beta}{1-\beta+\beta\rho} r_t a^\theta,\;\; \rho = \frac{1-\beta}{1-\beta\theta}\theta \Rightarrow \begin{cases} c = (1-\beta\theta)\, r_t a^\theta \\ a' = \beta\theta r_t a^\theta \end{cases}$$

The same structure appears in worker's problem with r_t replaced by w_t. Since the productivity is $i.i.d.$, the productivity and asset distribution are independent. In the equilibrium where everyone takes the same strategy $e_{it} = \phi$ for all t, the market clearing condition becomes

$$\begin{cases} n(z,w,k) = \left(\frac{\alpha z}{w}\right)^{\frac{1}{1-\alpha}} k = \left(\frac{\alpha z}{w}\right)^{\frac{1}{1-\alpha}} a^\theta \\ h = a^\theta \end{cases} \Rightarrow \left(\frac{\alpha}{w_t}\right)^{\frac{1}{1-\alpha}} \mathbb{E}z^{\frac{1}{1-\alpha}}\phi = 1-\phi \tag{39}$$

so w_t and r_t is time independent. One can also see the value functions are also time independent and v^W does not depend on z.

$$\ln v^E(z) = (1-\beta)\ln(1-\beta\theta) + \beta\theta\frac{1-\beta}{1-\beta\theta}\ln\beta\theta \\ + \frac{1-\beta}{1-\beta\theta}\ln r(z,w) + \beta\ln\left\{\phi\mathbb{E}v^E(z')^{1-\gamma} + (1-\phi)\left(v^W\right)^{1-\gamma}\right\}^{\frac{1}{1-\gamma}} \tag{40}$$

$$\ln v^W = (1-\beta)\ln(1-\beta\theta) + \beta\theta\frac{1-\beta}{1-\beta\theta}\ln\beta\theta \\ + \frac{1-\beta}{1-\beta\theta}\ln w + \beta\ln\left\{\phi\mathbb{E}v^E(z')^{1-\gamma} + (1-\phi)\left(v^W\right)^{1-\gamma}\right\}^{\frac{1}{1-\gamma}} \tag{41}$$

By subtracting one from the other, I can simplify the relationship between the two value functions

$$v^E(z) = \left(\frac{r(z,w)}{w}\right)^{\frac{1-\beta}{1-\beta\theta}} v^W. \tag{42}$$

The occupation choice gives the indifference condition

$$\left(\mathbb{E}v^E(z')^{1-\gamma}\right)^{\frac{1}{1-\gamma}} = v^W.$$

These two conditions lead to the equilibrium condition that the wage has to satisfy

$$\mathbb{E}\left(\frac{r(z,w)}{w}\right)^{\frac{(1-\beta)(1-\gamma)}{1-\beta\theta}} = 1. \tag{43}$$

Since the return to capital is

$$r(z,w) = (1-\alpha)\,\alpha^{\frac{\alpha}{1-\alpha}} w^{\frac{\alpha}{\alpha-1}} z^{\frac{1}{1-\alpha}},$$

the equilibrium wage is

$$w^m = (1-\alpha)^{1-\alpha}\,\alpha^\alpha \left(\mathbb{E} z^{\frac{(1-\gamma)(1-\beta)}{(1-\alpha)(1-\beta\theta)}}\right)^{\frac{(1-\alpha)(1-\beta\theta)}{(1-\gamma)(1-\beta)}}$$

and the number of firms is given by the market clearing condition

$$\phi^m = \frac{1-\alpha}{1-\alpha+\alpha\mathbb{E}z^{\frac{1}{1-\alpha}}\left\{\mathbb{E}\left(z^{\frac{1}{1-\alpha}}\right)^{\frac{(1-\gamma)(1-\beta)}{1-\beta\theta}}\right\}^{-\frac{1-\beta\theta}{(1-\gamma)(1-\beta)}}}.$$

(Planner) To derive the optimal number of firms, it suffices to consider output maximization. The problem is recursive, so fix a time period t and the aggregate asset saving $A_t := \int a_{it} di$. This is an endogenous variable, but for the purpose of deriving the optimal number of firms, I can take it as given as will be clear. Since the investment function f_I is strictly concave and z_{it+1} is not observable at t, it is optimal to choose the same saving across agents.

$$k_{it} = k_t := f_I\left(a_t^E\right),\ h_{it} := h_t := f_I\left(a_t^W\right).$$

Conditional on the amount of capital (k_t, h_t), the problem at t reduces to

$$\max_{Y_{t+1},n_{it},\phi_t} Y_{t+1}\ s.t. \begin{cases} \phi_t \int n_{it+1} di = (1-\phi_t)\,h_t \\ Y_{t+1} = \phi_t \int z_{it+1} f\left(k_{t-1}, n_{it+1}\right) di \end{cases}.$$

By maximizing out n_{it+1}, the aggregate output can be written as

$$Y_{t+1} = \phi_t^{1-\alpha}\,(1-\phi_t)^\alpha \left(\mathbb{E}z^{\frac{1}{1-\alpha}}\right)^{1-\alpha} k_t^{1-\alpha} h_t^\alpha.$$

Hence, the planner's problem reduces to

$$\max_{a_t^E, a_t^W, \phi_t, k_t, h_t} Y_{t+1} \ s.t. \ \begin{cases} \phi_t a_t^E + (1-\phi_t)\, a_t^W = A_t \\ Y_{t+1} = \phi_t^{1-\alpha} (1-\phi_t)^{\alpha} \left(\mathbb{E} z^{\frac{1}{1-\alpha}}\right)^{1-\alpha} k_t^{1-\alpha} h_t^{\alpha} \\ k_t = f_I\left(a_t^E\right), \ h_t = f_I\left(a_t^W\right) \end{cases}.$$

By optimizing out $\left(a_t^E, a_t^W\right)$, the output takes the following form

$$Y_{t+1} = \left\{\phi_t^{1-\alpha} (1-\phi_t)^{\alpha}\right\}^{1-\theta} \left(\mathbb{E} z^{\frac{1}{1-\alpha}}\right)^{1-\alpha} (1-\alpha)^{\theta(1-\alpha)} \alpha^{\theta\alpha} A_t^{\theta}.$$

Hence, the optimal number of firms is $\phi_t^P = 1-\alpha$. One can see $\phi_t^{LF} < \phi_t^P$ iff $V(z) > 0$ since $\phi^{LF} < 1-\alpha$ is equivalent to

$$\mathbb{E} z^{\frac{1}{1-\alpha}} > \left\{\mathbb{E}\left(z^{\frac{1}{1-\alpha}}\right)^{\frac{(1-\gamma)(1-\beta)}{1-\beta\theta}}\right\}^{\frac{1-\beta\theta}{(1-\gamma)(1-\beta)}} \Leftrightarrow \gamma + (1-\theta)\frac{\beta}{1-\beta} > 0,$$

which is always true under the parameter assumptions.

B Derivation of Eq.(12)

Eq.(12) can be derived as follows. The key is that marginal product equalization holds in the laissez-faire economy.

$$Y = \phi \mathbb{E}\left[z f\left(n(z,w)\right)\right].$$

By taking derivative, noting $w = w(T)$ and $\phi = \phi(T)$ are functions of policies

$$\partial_T Y = \mathbb{E} z f\left(n(z,w)\right) \partial_T \phi + \phi \mathbb{E}\left[z f'\left(n(z,w)\right) \partial_T n(z,w)\right].$$

Since the derivative is evaluated at laissez-faire, $z f'(n(z,w)) = w$, and $\partial_T Y$ becomes

$$\partial_T Y = \mathbb{E} z f\left(n(z,w)\right) \partial_T \phi + w\phi \mathbb{E}\left[\partial_T n(z,w)\right]$$

We can simplify the second term by substituting out the derivative of the market clearing condition

$$\partial_T \phi \mathbb{E} n(z,w) + \phi \mathbb{E}\left[\partial_T n(z,w)\right] = -\partial_T \phi.$$

Hence, the marginal impact can be written as

$$\partial_T Y = \partial_T \phi \mathbb{E} z f(n) - w \partial_T \phi (1 + \mathbb{E} n) = \partial_T \phi \left(\mathbb{E}\pi(z,w) - w\right).$$

Note that the derivation does not use the indifference condition $\mathbb{E}u\left(\pi\left(z,w\right)\right)=u\left(w\right)$.

C Non-pecuniary SDP

If the SDP is non-pecuniary and does not generate tax revenue, the fair comparison is the economy with non-pecuniary SDP

$$\mathbb{E}u\left(\pi^{SDP}\left(z,w^{SDP}\right)\right)=u\left(w^{SDP}\right)$$

$$\left[\pi^{SDP}\left(z,w^{SDP}\right),n^{SDP}\left(z,w^{SDP}\right)\right]=\max_{n}\begin{cases}zf\left(n\right)-w^{SDP}n & n\leq N\\ zf\left(n\right)-w^{SDP}\tau n-F & n>N\end{cases}$$

$$\phi^{SDP}\mathbb{E}n\left(z,w^{SDP}\right)=1-\phi^{SDP}$$

and (2) the laissez-faire economy with the same amount of non-distortionary regulation costs

$$\mathbb{E}u\left(\left(1-t\right)\pi^{LF}\left(z,w^{LF}\right)\right)=u\left(\left(1-t\right)w^{LF}\right)$$

$$\left[\pi^{LF}\left(z,w^{LF}\right),n^{LF}\left(z,w^{LF}\right)\right]=\max_{n}zf\left(n\right)-w^{LF}n$$

$$\phi^{LF}\mathbb{E}n^{LF}\left(z,w^{LF}\right)=1-\phi^{LF}$$

where t is chosen to equate the tax revenue in the first economy, *i.e.*, equation (18).

With this specification, the policy implication remains the same.

$$\ln\left\{\left(1-t\right)w^{LF}\right\}-\ln w^{SDP}=-1.27\%<0.$$

D MLE estimation

The effective tax rates (τ,F) are estimated from the firm size distribution. Given the labor demand $n\left(z,w;\tau,F\right)$ and the the power law density of productivity z, the model-implied firm size distribution $G_{n}\left(n;\tau,F\right):=P\left(n\left(\frac{z}{w};\tau,\frac{F}{w}\right)\leq n\right)$ follows a broken power law and can be fit with the empirical firm size distribution using the maximum likelihood method.

To be more concrete, following the same derivation as Garicano et al. [2016], the likelihood function is constructed as follows. Given the parametric assumptions $f\left(n\right)=n^{\alpha}$, the demand function can be written as

$$n\left(z,w;\tau,F\right)=\begin{cases}\left(\frac{z\alpha}{w}\right)^{\frac{1}{1-\alpha}} & z_{min}\leq z\leq \underline{z}\\ N & \underline{z}\leq z\leq \bar{z}\\ \left(\frac{z\alpha}{w\tau}\right)^{\frac{1}{1-\alpha}} & \bar{z}<z\leq z_{max}\end{cases},\quad \begin{cases}N=\left(\frac{\underline{z}\alpha}{w}\right)^{\frac{1}{1-\alpha}}\Leftrightarrow \frac{\underline{z}}{w}=\frac{N^{1-\alpha}}{\alpha}\\ \bar{z}N^{\alpha}-wN=\bar{z}\left(\frac{\bar{z}\alpha}{w\tau}\right)^{\frac{\alpha}{1-\alpha}}-w\tau\left(\frac{\bar{z}\alpha}{w\tau}\right)^{\frac{1}{1-\alpha}}-F\\ \Leftrightarrow \frac{F}{w}=N-\frac{\tau}{\alpha}\bar{n}\left(\left(\frac{N}{\bar{n}}\right)^{\alpha}+\alpha-1\right),\ \bar{n}=\left(\frac{\bar{z}\alpha}{w\tau}\right)^{\frac{1}{1-\alpha}}\end{cases}$$

Given the density of the productivity $g_z\left(z\right)$, I can derive the distribution function of $n_i=n\left(z_i,w\right)$ by using the change of variable.

$$G_n\left(x\right)=P\left(n\leq x\right)=\begin{cases}0 & x<n_{min}\\ \frac{n_{min}^{1-\beta}-x^{1-\beta}}{n_{min}^{1-\beta}-Tn_{max}^{1-\beta}} & n_{min}\leq x<N\\ \frac{n_{min}^{1-\beta}-T\bar{n}^{1-\beta}}{n_{min}^{1-\beta}-Tn_{max}^{1-\beta}} & N\leq x\leq \bar{n}\\ \frac{n_{min}^{1-\beta}-Tx^{1-\beta}}{n_{min}^{1-\beta}-Tn_{max}^{1-\beta}} & \bar{n}\leq x\leq n_{max}\\ 1 & n_{max}<x\end{cases},\quad \begin{cases}n_{min}=\left(\frac{z_{min}\alpha}{w}\right)^{\frac{1}{1-\alpha}}\\ n_{max}=\left(\frac{z_{max}\alpha}{w\tau}\right)^{\frac{1}{1-\alpha}}\\ 1-\beta=\left(1-\alpha\right)\left(1-\beta_z\right)\\ T=\tau^{1-\beta_z}\end{cases}$$

To reconcile the model in which there are no firms over the region $N\leq n\leq \bar{n}$ with the actual data (Fig.1), the observed firm size distribution is assumed to be generated with measurement error, i.e., $\tilde{n}_i=n\left(z_i,w\right)e^{-\sigma\epsilon_i}$, $\epsilon_i\sim N\left(0,1\right)$. The distribution of $\tilde{n}_i$ follows

$$\begin{aligned}G_{\tilde{n}}\left(\tilde{n}\right)=&P\left(ne^{-\sigma\epsilon}\leq \tilde{n}\right)=E\left[P\left(n\leq \tilde{n}e^{\sigma\epsilon}|\epsilon\right)\right]\\ =&\frac{1}{n_{min}^{1-\beta}-Tn_{max}^{1-\beta}}\left[n_{min}^{1-\beta}\Phi\left(\frac{\ln\frac{\tilde{n}}{n_{min}}}{\sigma}\right)-Tn_{max}^{1-\beta}\Phi\left(\frac{\ln\frac{\tilde{n}}{n_{max}}}{\sigma}\right)\right.\\ &-\tilde{n}^{1-\beta}e^{\frac{\sigma^2(1-\beta)^2}{2}}\left\{\Phi\left(\left(1-\beta\right)\sigma-\frac{\ln\frac{n_{min}}{\tilde{n}}}{\sigma}\right)-\Phi\left(\left(1-\beta\right)\sigma-\frac{\ln\frac{N}{\tilde{n}}}{\sigma}\right)\right\}\\ &-T\bar{n}^{1-\beta}\left\{\Phi\left(\frac{\ln\frac{\bar{n}}{\tilde{n}}}{\sigma}\right)-\Phi\left(\frac{\ln\frac{N}{\tilde{n}}}{\sigma}\right)\right\}\\ &\left.-T\tilde{n}^{1-\beta}e^{\frac{\sigma^2(1-\beta)^2}{2}}\left\{\Phi\left(\left(1-\beta\right)\sigma-\frac{\ln\frac{\bar{n}}{\tilde{n}}}{\sigma}\right)-\Phi\left(\left(1-\beta\right)\sigma-\frac{\ln\frac{n_{max}}{\tilde{n}}}{\sigma}\right)\right\}\right]\end{aligned}$$

In addition, to utilize the bunching around N and avoid the deviation of the model from the data at the extremes, the sample $\left\{\tilde{n}_i\right\}_i$ is truncated between $\tilde{n}_{min}=10$ to $\tilde{n}_{max}=10,000$. Hence, the likelihood function of the parameters $\theta=\left(\alpha,n_{min},n_{max},\bar{n},\beta,\sigma,T\right)$ is

$$L\left(\theta;\left\{\tilde{n}_i\right\}_i\right)=\frac{\prod_i G'_{\tilde{n}}\left(\tilde{n}_i;\theta\right)}{G_{\tilde{n}}\left(\tilde{n}_{max};\theta\right)-G_{\tilde{n}}\left(\tilde{n}_{min};\theta\right)}\tag{44}$$

The parameters are subject to the constraint specified by the indifference condition. Given the CRRA utility parameterization $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$, the indifference condition is

$$\mathbb{E}\left[\left(\frac{z}{w}f(n) - n - \left(\tau - 1 + \frac{F}{w}\right)1_{n>N}\right)^{1-\gamma}\right] = 1. \tag{45}$$

By applying the parametric assumptions, this condition can be rewritten as

$$\begin{aligned}
1 = &\left(\frac{1-\alpha}{\alpha}\right)^{1-\gamma}\frac{\beta-1}{2-\gamma-\beta}\frac{N^{2-\gamma-\beta} - n_{min}^{2-\gamma-\beta}}{n_{min}^{1-\beta} - Tn_{max}^{1-\beta}} \\
&+ \frac{\beta_z - 1}{\left(\frac{N^{1-\alpha}}{\alpha}\right)^{1-\beta_z} - \left(\frac{\bar{n}^{1-\alpha}\tau}{\alpha}\right)^{1-\beta_z}}\int_{\frac{N^{1-\alpha}}{\alpha}}^{\frac{\bar{n}^{1-\alpha}\tau}{\alpha}}\left(xN^{\alpha} - N\right)^{1-\gamma}x^{-\beta_z}dz \\
&+ \frac{(\beta-1)T}{n_{min}^{1-\beta} - Tn_{max}^{1-\beta}}\int_{\bar{n}}^{n_{max}}\left(\frac{1-\alpha}{\alpha}\tau n - \frac{F}{w}\right)^{1-\gamma}n^{-\beta}dn.
\end{aligned}$$

Note that since the likelihood is based on truncation, the extreme values (n_{min}, n_{max}) cannot be estimated accurately. Garicano et al. [2016] avoids this problem by fixing $(\alpha, n_{max}) = (.8, \infty)$ and maximizing the likelihood L subject to their version of the indifference condition $n_{min} = \frac{\alpha}{1-\alpha}$, so the problem becomes unconstrained maximization of the likelihood (44) over $(\bar{n}, \beta, \sigma, T)$. Similarly, I fix $(\alpha, n_{max}) = (.8, \tilde{n}_{max})$ and maximize the likelihood (44) subject to the indifference condition (45).

The two models generate almost identical parameter values except for n_{min}. For n_{min}, Garicano et al. [2016] obtain $n_{min} = 4$ and my model implies $n_{min} = 1.7796$. For other parameters, the estimates under the two models are listed in the following table.

Table 3: Comparison

	Baseline	GLV with Amadeus	GLV
β	1.848	1.849	1.8
	(0.031)	(0.028)	(0.054)
$\bar{n}$	57.903	57.839	59.271
	(0.98)	(0.901)	(2.051)
τ -1	0.021	0.021	0.023
	(0.005)	(0.006)	(0.008)
σ	0.1	0.099	0.121
	(0.014)	(0.013)	(0.033)
Number of Firms	62549	62549	41067

The first two columns use the same data from Amadeus 2006. The first column assumes entrepreneurial risks and the second does not. For the purpose of comparison, the third column lists the estimates from Garicano et al. [2016] that use FICUS data.

One can see that the first two columns are almost identical. This is not surprising since the

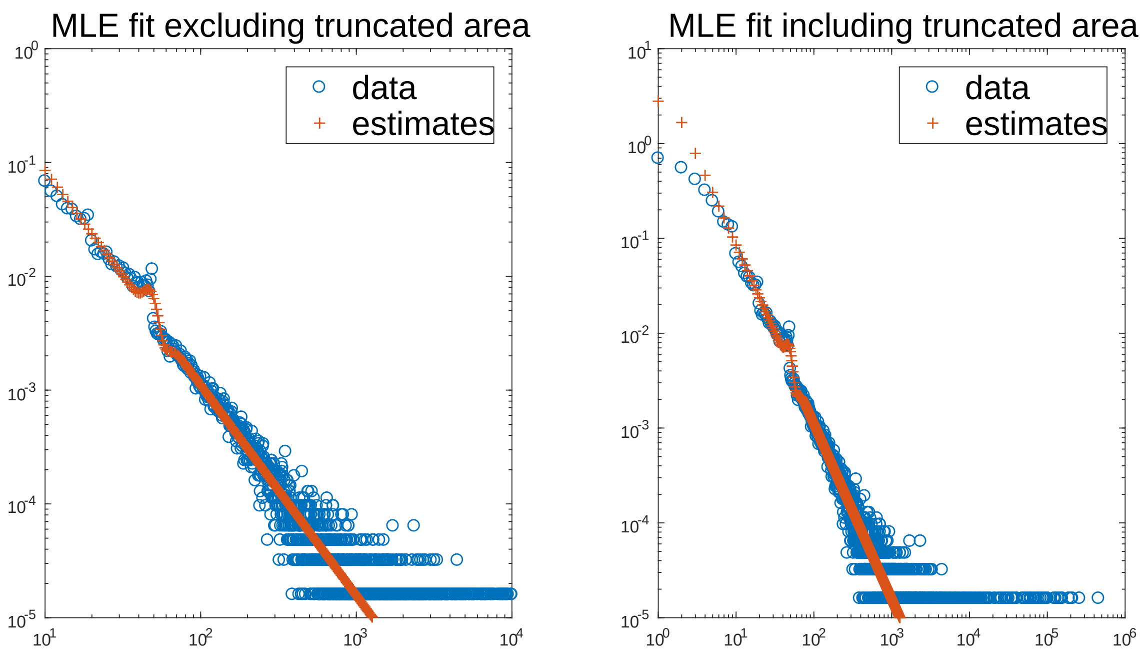

Figure 12: MLE fit. The left panel shows the fit on the truncated region. The right panel shows the fit on the entire region.

only difference between the two estimates is the shape of the indifference condition. Although the indifference condition affects n_{min}, since the likelihood is truncated, it does not contain much information for the estimation of $(\bar{n}, \beta, \sigma, T)$. In fact, if there is no observation error $\sigma = 0$, the truncation makes the likelihood estimation with and without entrepreneurial risks exactly identical. Therefore, in terms of the estimation, the existence of the entrepreneurial risks dose not matter.

One can also see that the second and the third columns are not that different. Therefore, Amadeus 2006 is a good approximation of the administrative FICUS data. Hence, even if I use FICUS data to estimate all the model parameters, I will get almost the same values as Garicano et al. [2016].

To see how well MLE estimation fits the data, I refer readers to Figure 8 of Garicano et al. [2016] for FICUS data and Fig.12 for Amadeus.

E Intermediate signal

In this section, I show one way to introduce intermediate signal. The main finding is that the efficiency gain might not be monotonic over riskiness.

To study the intermediate signal, I need to take a stance on the welfare criterion and the parameterization of the signal structure. For the welfare criterion, since the model becomes heterogeneous agents whenever there are informative signals, I focus on the aggregate output instead of focusing on a particular Pareto weight. For the parametric assumption on the signal, I assume that each individual observes

$$s_i = 1_{\pi(z_i,w)\geq \eta w},\ \eta \in [0,1]\,. \tag{46}$$

I could choose other signal structures such as log normal, but this signal structure has at least two several advantages in the current setting. First, the parameter η has a natural interpretation, i.e., agents who observe $s_i = 1$ know that they can earn at least $\eta\%$ of their salary if they are engaged in entrepreneurship. The interpretation makes it clear that $\eta = 0$ corresponds to full-risk case, and $\eta = 1$ to no-risk case. Specifically, when $\eta = 0$, the signal is uninformative, so the equilibrium is characterized by (8), while when $\eta = 1$, the wage and signal threshold satisfying (7) is an equilibrium. Therefore, $\eta \in [0,1]$ is a legitimate parameter that connects the full-risk and no-risk cases. Second, the signal structure (46) preserves power law. Since the productivity z follows a power law and the signal gives a truncation

$$\pi\left(z_i,w\right) \geq \eta w \Leftrightarrow z_i \geq z\left(\eta,w\right),$$

the distribution of the productivity conditional on the signal is still a power law. This allows me to use the same parameter values as in 4.1.

Under the signal structure (46), the equilibrium system is a function of η. For small η, there are plenty of agents observing $s_i = 1$, so the equilibrium (w,ϕ) is determined by

$$\mathbb{E}\left[u\left(\pi\left(z_i,w\right)\right)|s_i=1\right] = u\left(w\right),\ \phi\mathbb{E}\left[n\left(z_i,w\right)|s_i=1\right] = 1-\phi,\ P\left(s_i=1\right) > \phi.$$

Mathematically, the third condition is slack for small η. As η increases, the third condition becomes tighter, and when it binds, the indifference condition has to be relaxed. Hence, for large η,

$$\mathbb{E}\left[u\left(\pi\left(z_i,w\right)\right)|s_i=1\right] > u\left(w\right),\ \phi\mathbb{E}\left[n\left(z_i,w\right)|s_i=1\right] = 1-\phi,\ P\left(s_i=1\right) = \phi.$$

The results are shown in Fig.13. For each η, the wage is normalized to $w^{SDP} = 1$. At the two extremes, the efficiency gain from removing the SDP is -1.5% and $.02\%$. For low value of η, neither economy reacts to the increase in η. This is because $z_{min} > 0$, so for the small value of η, the information about the productivity is not informative $\pi\left(z_{min},w\right) \geq \eta w$. As η crosses some threshold, the outputs of both economies start to go up since the average entrepreneurs are more productive. However, the speed of the output growth might be different, so the difference might not be monotone. In the current setting, $\eta = 45\%$ is the threshold above

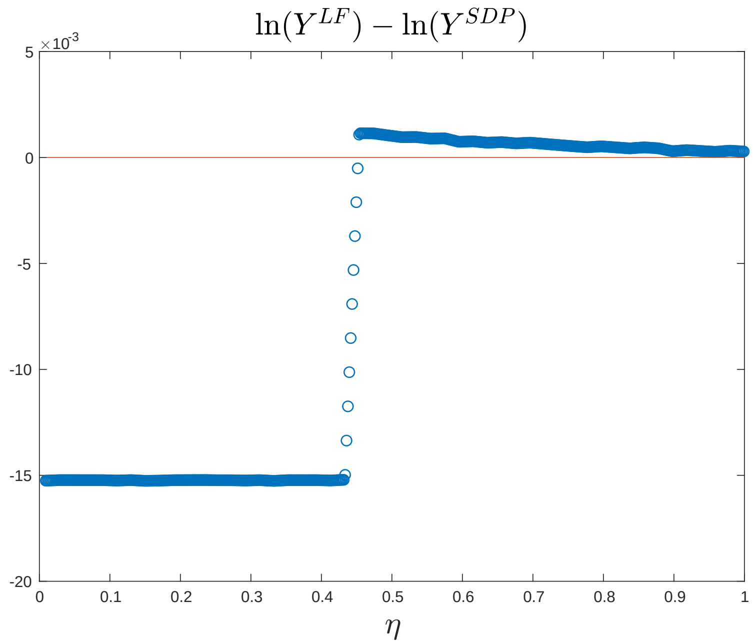

Figure 13: Counterfactual analysis with intermediate signal.

which the laissez-faire economy produces higher output. For different signal structure, the efficiency gain can be smoother.

F Constrained efficiency

F.1 Constrained efficiency with intermediate signal

The constrained planner's problem is

$$\max_{\bar{s},w} \int_{\bar{s}}^{\infty} \mathbb{E}\left[u\left(\pi\left(z,w\right)\right)|s\right] dG_s + G_s\left(\bar{s}\right) u\left(w\right) \;\; s.t. \;\; \int_{\bar{s}}^{\infty} \mathbb{E}\left[n\left(z,w\right)|s\right] dG_s = G_s\left(\bar{s}\right).$$

The idea is that the planner's choice set is restricted to the number of firms ϕ, or equivalently $\bar{s}$. All other allocations, such as employment schedule and consumption, are determined by the market. This is a stronger result of efficiency and follow the spirit of Diamond [1967].

Denote the solution by ϕ^{CP}. By restricting the functional form, a similar result as unconstrained efficiency can be obtained.

Claim. Fix CRRA utility and Cobb-Douglas production functions (u, f). The laissez-faire economy generates insufficient number of firms $\phi^{LF} < \phi^{CP}$ compared to the constrained planner iff there are entrepreneurial risks.

Proof. The equilibrium of the laissez-faire economy is characterized by $\left(\bar{s}^{LF}, w^{LF}\right)$ that satisfies

$$\frac{1-G_s(\bar{s})}{G_s(\bar{s})}\mathbb{E}\left[z^{\frac{1}{1-\alpha}}|s\geq\bar{s}\right]=\frac{1-\alpha}{\alpha}\left(\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s=\bar{s}\right]\right)^{\frac{1}{1-\gamma}}.$$

The planner solve

$$\max_{\bar{s}}\frac{1-G_s(\bar{s})}{1-\gamma}\left((1-\alpha)\,\alpha^{\frac{\alpha}{1-\alpha}}w(\bar{s})^{\frac{\alpha}{\alpha-1}}\right)^{1-\gamma}\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s\geq\bar{s}\right]+\frac{G_s(\bar{s})}{1-\gamma}w(\bar{s})^{1-\gamma}$$

$$s.t.\ w(\bar{s})=\alpha\left(\frac{1-G_s(\bar{s})}{G_s(\bar{s})}\mathbb{E}\left[z^{\frac{1}{1-\alpha}}|s\geq\bar{s}\right]\right)^{1-\alpha}$$

Substituting out the constraint with respect to wage, the planner's problem becomes

$$\max_{\bar{s}}\frac{1}{1-\gamma}\left[(1-\alpha)^{1-\gamma}\,\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}1_{s\geq\bar{s}}\right]\left(\frac{\mathbb{E}\left[z^{\frac{1}{1-\alpha}}1_{s\geq\bar{s}}\right]}{G_s(\bar{s})}\right)^{-\alpha(1-\gamma)}\right.$$
$$\left.+G_s(\bar{s})\,\alpha^{1-\gamma}\left(\frac{\mathbb{E}\left[z^{\frac{1}{1-\alpha}}1_{s\geq\bar{s}}\right]}{G_s(\bar{s})}\right)^{(1-\alpha)(1-\gamma)}\right]$$

Taking derivative with respect to $\bar{s}$ and evaluating the FOC at $\left(\bar{s}^{LF},w^{LF}\right)$ gives

$$\begin{aligned}D\left(\bar{s}^{LF},w^{LF}\right)&=\left\{\int_{\bar{s}^{LF}}^{\infty}\mathbb{E}_{z|s}\left[\frac{\partial u\left(\pi\left(z,w^{LF}\right)\right)}{\partial w}\right]dG_s+G_s\left(\bar{s}^{LF}\right)u'\left(w^{LF}\right)\right\}w'\left(\bar{s}^{LF}\right)\\&=\left\{\underbrace{\frac{G_s\left(\bar{s}^{LF}\right)}{1-G_s\left(\bar{s}^{LF}\right)}-(1-\alpha)^{-\gamma}\,\alpha^{\frac{1-\alpha\gamma}{1-\alpha}}\left(w^{LF}\right)^{\frac{1-\gamma}{\alpha-1}}\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s\geq\bar{s}^{LF}\right]}_{=:\mathcal{D}\left(\bar{s}^{LF},w^{LF}\right)}\right\}\\&\quad\times\left(w^{LF}\right)^{-\gamma}\left(1-G_s\left(\bar{s}^{LF}\right)\right)w'\left(\bar{s}^{LF}\right).\end{aligned}$$

If $D\left(\bar{s}^{LF},w^{LF}\right)<0$, it means that decreasing $\bar{s}$ increases welfare, so $\phi^{LF}<\phi^{CP}$. Since $w'\left(\bar{s}^{LF}\right)<0$, it suffices to show $\mathcal{D}\left(\bar{s}^{LF},w^{LF}\right)>0$. By applying the indifference condition and the market clearing condition,

$$\begin{aligned}\mathcal{D}(\bar{s},w)&=\frac{G_s(\bar{s})}{1-G_s(\bar{s})}-\frac{\alpha}{1-\alpha}\frac{\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s\geq\bar{s}\right]}{\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s=\bar{s}\right]}\\&=\frac{\alpha}{1-\alpha}\left(\frac{\mathbb{E}\left[z^{\frac{1}{1-\alpha}}|s\geq\bar{s}\right]}{\left(\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s=\bar{s}\right]\right)^{\frac{1}{1-\gamma}}}-\frac{\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s\geq\bar{s}\right]}{\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s=\bar{s}\right]}\right)\end{aligned}$$

When $\gamma \geq 1$, $\mathcal{D}\left(\bar{s}^{LF}, w^{LF}\right) > 0$. To see this, note that, by lemma 1, the first term is larger than 1.

$$\frac{\mathbb{E}\left[z^{\frac{1}{1-\alpha}}|s \geq \bar{s}\right]}{\left(\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s = \bar{s}\right]\right)^{\frac{1}{1-\gamma}}} > \frac{\mathbb{E}\left[z^{\frac{1}{1-\alpha}}|s = \bar{s}\right]}{\left(\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s = \bar{s}\right]\right)^{\frac{1}{1-\gamma}}} > 1$$

The second term is smaller than 1 since $z \mapsto z^{\frac{1-\gamma}{1-\alpha}}$ is a decreasing function, so the higher the signal is the lower the integrand is. When $\gamma \in (0,1)$, $\mathcal{D}\left(\bar{s}^{LF}, w^{LF}\right) > 0$ still holds. To see this, note that by Jensen,

$$\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s \geq \bar{s}\right] \leq \left(\mathbb{E}\left[z^{\frac{1}{1-\alpha}}|s \geq \bar{s}\right]\right)^{1-\gamma} \Rightarrow \frac{\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s \geq \bar{s}\right]}{\mathbb{E}\left[z^{\frac{1}{1-\alpha}}|s \geq \bar{s}\right]} \leq \left(\mathbb{E}\left[z^{\frac{1}{1-\alpha}}|s \geq \bar{s}\right]\right)^{-\gamma}$$

$$\begin{aligned}\frac{\left(\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s = \bar{s}\right]\right)^{\frac{1}{1-\gamma}}}{\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s = \bar{s}\right]} &= \left(\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s = \bar{s}\right]\right)^{\frac{1}{1-\gamma}-1} \\ &\leq \left(\left(\mathbb{E}\left[z^{\frac{1}{1-\alpha}}|s = \bar{s}\right]\right)^{1-\gamma}\right)^{\frac{1}{1-\gamma}-1} = \left(\mathbb{E}\left[z^{\frac{1}{1-\alpha}}|s = \bar{s}\right]\right)^{\gamma}\end{aligned}$$

Multiplying these two equations gives

$$\frac{\left(\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s = \bar{s}\right]\right)^{\frac{1}{1-\gamma}}}{\mathbb{E}\left[z^{\frac{1}{1-\alpha}}|s \geq \bar{s}\right]} \frac{\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s \geq \bar{s}\right]}{\mathbb{E}\left[z^{\frac{1-\gamma}{1-\alpha}}|s = \bar{s}\right]} \leq \left(\frac{\mathbb{E}\left[z^{\frac{1}{1-\alpha}}|s = \bar{s}\right]}{\mathbb{E}\left[z^{\frac{1}{1-\alpha}}|s \geq \bar{s}\right]}\right)^{\gamma} \leq 1.$$

Since all I use is Jensen's inequality, $\phi^{LF} = \phi^{CP}$ when there are no entrepreneurial risks. □

The functional form assumption is necessary. If we relax the assumptions to just $u' > 0 > u''$, it is possible to find f, G such that $\phi^{CP} < \phi^{LF}$. For the detail, see Ando and Matsumura [2017].

F.2 Constrained efficiency with financial frictions

The constrained planner solves

$$\max_{\phi, w} \phi \mathbb{E} u\left(\pi\left(z, w\right)\right) + \left(1 - \phi\right) u\left(w\right)$$

$$s.t. \begin{cases} \pi\left(z, w\right) = \max_n zn^{\alpha} - wn \ s.t. \ wn \leq \lambda z n^{\alpha} \\ \phi \mathbb{E} n\left(z, w\right) = 1 - \phi \end{cases}.$$

The idea is the same as section F.1. The constrained planner can only choose the number of firms and has to let all other allocations be determined by the market.

Claim. Fix CRRA utility and Cobb-Douglas production functions (u, f) and arbitrary risk structure G_z with bounded support. If the financial constraints are severe $\lambda \approx 0$, the laissez-faire economy generates an excessive number of firms $\phi^{CP} < \phi^{LF}$.

By substituting out ϕ using the market clearing condition and applying an increasing transformation, the objective function can be written as

$$U(w) := \left\{\phi(w)\,\mathbb{E}\pi(z,w)^{1-\gamma} + (1-\phi(w))\,w^{1-\gamma}\right\}^{\frac{1}{1-\gamma}} = w\left\{\frac{\frac{\mathbb{E}\pi(z,w)^{1-\gamma}}{w^{1-\gamma}} + \mathbb{E}n(z,w)}{1+\mathbb{E}n(z,w)}\right\}^{\frac{1}{1-\gamma}}$$

By taking derivative with respect to w,

$$\frac{\partial}{\partial w}\ln U(w) = \frac{1}{w} + \frac{1}{1-\gamma}\left\{\frac{\frac{\partial}{\partial w}\frac{\mathbb{E}\pi(z,w)^{1-\gamma}}{w^{1-\gamma}} + \frac{\partial}{\partial w}\mathbb{E}n(z,w)}{\frac{\mathbb{E}\pi(z,w)^{1-\gamma}}{w^{1-\gamma}} + \mathbb{E}n(z,w)} - \frac{\frac{\partial}{\partial w}\mathbb{E}n(z,w)}{1+\mathbb{E}n(z,w)}\right\}$$

The constrained efficient allocation is characterized by $\frac{\partial}{\partial w}\ln U\left(w^{CP}\right) = 0$. Evaluating the FOC at $w = w^{LF}$ gives

$$\frac{\partial}{\partial w}\ln U\left(w^{LF}\right) = \frac{1}{w^{LF}} + \frac{1}{1-\gamma}\left\{\frac{\frac{1-\gamma}{w^{LF}}\frac{1}{\alpha-1}}{1+\mathbb{E}n\left(z,w^{LF}\right)}\right\} = \frac{1}{w^{LF}}\left\{1 - \frac{\phi^{LF}}{1-\alpha}\right\}$$

Since $\phi^{LF} \to 1$ as $\lambda \to 0$, the FOC at $w = w^{LF}$ can be negative. Hence, $w^{LF} > w^{CP}$. By the market clearing condition $\phi^{CP} < \phi^{LF}$.

F.3 Constrained efficiency with capital accumulation

The formulation of the constrained efficiency problem inherits the ex-ante view of the planner in the static model. In particular, I make the following thought experiment. Suppose all agents are in a symmetric equilibrium in which they face the same lottery and the economy is in a stationary equilibrium. If, at some period t, a planner shows up and offers a new job lottery, will everyone prefer it?

Such a thought experiment can be formulated as follows. Let $\phi = \{\phi_t\}_t$ be the probability of becoming entrepreneurs at each period chosen by the constrained planner. This is the only choice that the planner can control. Other decisions are made by the market, i.e., each individual i takes ϕ as given and maximizes the recursive welfare by making consumption/saving decision. Formally, the agent i with the state $(a, j, z) = (a_{it-1}, j_{it-1}, z_{it})$ solves

$$V_t(a, j, z; \phi) = \max_{(c,a') \in B_t(a,j,z)} u_I^{-1} \left[(1-\beta) u_I(c_{it}) + \right.$$
$$\left. \beta u_I \left(u^{-1} \left(\phi_t \mathbb{E} u \left(V_{t+1}(a', E, z'; \phi) \right) + (1 - \phi_t) \mathbb{E} u \left(V_{t+1}(a', W, z'; \phi) \right) \right) \right) \right] \quad (47)$$

Note that the budget set reflects that the production schedule $\{\pi(z, w, k), n(z, w, k)\}$ is determined by the market, not by the planner.

Since all the equilibrium objects are functions of the planner's choice ϕ, the wage process $w = \{w_t\}$ also has to be a function of ϕ. As in the static model, the labor market clearing condition

$$\phi_{t-1} \int n_{it} di = (1 - \phi_{t-1}) \int h_{it-1} di$$

specifies the mapping from the process of the number of the firms ϕ to wage w, denoted by $w(\phi)$. I note that this mapping is implicitly included in the value function $V_t(a, j, z; \phi)$.

In summary, the constrained planner can be defined as follows. Fix the fundamentals $\mathcal{E} = (\beta, u, u_I, f, f_I, G_z, S_0)$. The path of the number of firms $\left\{ w_t^{CP}, \phi_t^{CP} \right\}_t$ is constrained efficient if it solves

$$\max_{\phi, w} \int u^{-1} \left(\phi_t \mathbb{E} u \left(V_t(a_{it}, E, z_{it}; \phi) \right) + (1 - \phi_t) \mathbb{E} u \left(V_t(a_{it}, W, a_{it}; \phi) \right) \right) di$$

subject to $\{w_t\} = w(\phi)$.

Claim. Fix CRRA intra-temporal utility, log inter-temporal utility, and Cobb-Douglas production and asset transformation functions. Then, the laissez-faire economy generates an insufficient number of firms $\phi^{LF} < \phi^{CP}$.

Proof. The Bellman equations of the constrained planner's problem after maximizing out consumption-saving decision are the same, i.e., Eq.(40), (41) and (42) all hold. However, the indifference condition no longer holds. Still, I can derive the constrained planner's objective function. Note that by (42), the value of the next period can be written as

$$\left\{ \phi \mathbb{E} v^E(z)^{1-\gamma} + (1 - \phi) \left(v^W \right)^{1-\gamma} \right\}^{\frac{1}{1-\gamma}} = \left\{ \phi \mathbb{E} \left(\frac{r(z, w)}{w} \right)^{\frac{(1-\beta)(1-\gamma)}{1-\beta\theta}} + 1 - \phi \right\}^{\frac{1}{1-\gamma}} v^W. \quad (48)$$

By substituting this condition into the worker's Bellman equation (41), I can solve the worker's value function v^E in closed form

$$\ln v^W = \ln(1-\beta\theta) + \frac{\beta\theta}{1-\beta\theta}\ln\beta\theta + \frac{1}{1-\beta\theta}\ln w$$
$$+ \frac{\beta}{1-\beta}\ln\left\{\phi\mathbb{E}\left(\frac{r(z,w)}{w}\right)^{\frac{(1-\beta)(1-\gamma)}{1-\beta\theta}} + 1-\phi\right\}^{\frac{1}{1-\gamma}} \tag{49}$$

As a result, the planner maximizes (48) subject to (49) and the market clearing condition (39). Eq.(48) is equivalent to the objective function because the pre-determined asset distribution and the current productivity are independent $a_{it} \perp\!\!\!\perp z_{it}$ and the value functions are multiplicative with respect to asset $V_t(a,j,z) = v_t^j(z)\,a^\rho$. In other words, the constrained planner solves

$$\max_{\phi,w}\left\{\phi\mathbb{E}\left(\frac{r(z,w)}{w}\right)^{\frac{(1-\beta)(1-\gamma)}{1-\beta\theta}} + 1-\phi\right\}^{\frac{1}{1-\gamma}} v^W \;\; s.t. \;\; \begin{cases}\alpha^{\frac{1}{1-\alpha}}\mathbb{E}z^{\frac{1}{1-\alpha}}w^{\frac{1}{\alpha-1}}\phi = 1-\phi \\ (49)\end{cases}$$

By taking log of the objective function, the FOC is

$$U'(w)\,w = \frac{1}{1-\beta\theta} + \frac{1}{(1-\gamma)(1-\beta)(1-\alpha)}\left\{1-\phi-\frac{\frac{(1-\beta)(1-\gamma)}{1-\beta\theta}\mathbb{E}\left(\frac{r(z)}{w}\right)^{\frac{(1-\beta)(1-\gamma)}{1-\beta\theta}} + \frac{1}{\phi}-1}{\mathbb{E}\left(\frac{r(z)}{w}\right)^{\frac{(1-\beta)(1-\gamma)}{1-\beta\theta}} + \frac{1}{\phi}-1}\right\}.$$

Since the indifference condition $\mathbb{E}\left(\frac{r(z)}{w}\right)^{\frac{(1-\beta)(1-\gamma)}{1-\beta\theta}} = 1$ holds in the equilibrium of the laissez-faire economy, (see Eq.(43))

$$U'\left(w^{LF}\right)w^{LF} = \frac{1}{1-\beta\theta}\left(1-\frac{\phi^{LF}}{1-\alpha}\right).$$

Hence, $\phi^{CP} \gtreqless \phi^{LF}$ if and only if $\phi^{LF} \lesseqgtr 1-\alpha$, which is then equivalent to $\gamma+(1-\theta)\frac{\beta}{1-\beta} \gtreqless 0$. Since the inequality is always positive, the laissez-faire economy always generates insufficient firms. $\square$