WORKING PAPER

N° 2024-8

CORPORATE TAXATION AND FIRM HETEROGENEITY

JULIEN ALBERTINI, XAVIER FAIRISE, ANTHONY TERRIAU

www.tepp.eu

TEPP – Theory and Evaluation of Public Policies - FR CNRS 2042

Corporate taxation and firm heterogeneity

Julien Albertini[*] Xavier Fairise[†] Anthony Terriau[‡]

September 18, 2024

Abstract

This paper explores the differentiated effects of corporate tax changes based on firm characteristics and evaluates the potential impact of a tax system modulated by both firm size and age. Using tax rate variations across U.S. states and comparing adjacent counties across state borders, we find that corporate taxes significantly reduce employment in small and young firms, while having no notable impact on large and older firms. We then develop a model to analyze firm dynamics throughout their life cycle under different tax regimes. Our simulations show that a corporate tax system adjusted by both firm size and age is more effective than one based solely on size (and even more so than a system with a single rate). This approach lightens the tax burden on highly productive young firms and shifts it toward less productive older firms, ultimately boosting employment and welfare without reducing the fiscal surplus.

JEL Classification: H25, H32, J21, J23, E61, E62

[*]Julien Albertini, Université Lumière Lyon 2, CNRS, Université Jean Monnet Saint-Etienne, emlyon business school, GATE, 69007, Lyon, France

[†]GAINS, Le Mans Université

[‡]GAINS, Le Mans Université

1 Introduction

The impact of corporate taxation on business activity has been widely debated over the past decades, both among economists and policymakers (Adelman, 1957; Due, 1961; Auerbach et al., 1983). As noted by Jacob (2021), a large body of the literature focuses on the effect of corporate income tax on investment (Sandmo, 1974; Djankov et al., 2010; Edgerton, 2010; Ohrn, 2018), while studies analyzing the employment response are relatively scarce (Mertens and Ravn, 2013). Harden and Hoyt (2003), employing state-level data spanning 1977 to 1994, find that corporate tax increases have a detrimental impact on employment. Giroud and Rauh (2019), utilizing firm-level data from 1977 to 2011, show that firms decrease their workforce in response to increases in state corporate income tax. Mertens and Ravn (2013), adopting a narrative approach and quarterly data covering 1950 to 2006, find no evidence of a significant effect of corporate tax cuts on employment.

Identifying the effects of state corporate income tax on employment is a challenging task. Distinguishing the effects of state policy from those of state characteristics is difficult, and state policy is likely to be endogenous to the state's labor market conditions. Ljungqvist and Smolyansky (2014) propose a compelling identification strategy. Following the approach developed by Dube et al. (2010), they exploit policy discontinuities at state borders, and compare all contiguous county pairs in the United States that are located on opposite sides of a state border. This strategy offers two advantages. First, contiguous border counties exhibit similar characteristics and employment trends (Hagedorn et al., 2015). Second, since each state encompasses numerous counties and must consider the potentially divergent interests of all counties, not just a specific one, it is unlikely that labor market conditions in a particular county would significantly influence state policy (Huang, 2008). The first argument makes the common trend assumption much more credible while the second argument minimizes endogeneity issues. Based on this empirical framework, Ljungqvist and Smolyansky (2014) find that a one percentage point increase in corporate income tax leads to a 0.2% decline in employment in the affected county relative to the neighboring unaffected county located on the other side of the state border.

In this paper, we take a further step by analyzing the employment response based on firm size and age. As emphasized in Haltiwanger et al. (2013)'s seminal paper, there is a widespread belief that small firms create the most jobs. However, this perception is somewhat misleading. The pool of small businesses comprises two types of firms: unproductive older firms that have failed to grow, and young firms that are, on average, more productive (conditional on survival) but still in a growth phase. The latter, often referred to as "startups", represent the primary source of job creation in the United States. Their initial small size has contributed to the prevailing notion that small firms are a source of employment. This also explains why some countries, particularly in Europe, have implemented tax reductions targeted at small firms to stimulate employment. In this paper, we examine whether tax cuts would be more effective (and to what extent) if targeted at young firms (which are mostly small) rather than small firms (which can be young or old), in order to prevent these tax reductions from benefiting unproductive older firms.

We first provide empirical evidence of the impact of corporate income tax on employment. In addition to the federal corporate tax, many states have introduced their own corporate tax, which varies across states and years. Unlike most existing studies that employ a conventional state and year fixed effects model, we follow the approach developed by Ljungqvist and Smolyansky (2014). We compare contiguous county pairs in neighboring states with different corporate tax rates. Using the Business Dynamics Statistics for the period 2002 to 2009, we estimate the effect of state corporate income tax on employment by firm size and age. We show that corporate tax has a significant and negative impact on employment in small (young) firms, whereas there is no evidence of a significant effect on large (old) firms. When comparing the effect on small versus young firms, our estimates suggest that a one percentage point increase in the state corporate income tax rate reduces employment in small (young) firms by 1.35% (4.42%). Given that more than 99% of young firms are small and 40% of small firms are young, these results suggest that a significant portion of the employment response of small firms to corporate tax is driven by young firms. This also suggests that tax reductions implemented in some countries could be more effective if targeted at young firms (rather than small firms).

We then develop a model of firm dynamics in the spirit of Elsby and Michaels (2013). Following Sedláček and Sterk (2017) and Sedláček (2020), we consider that firms differ in terms of age and productivity. This allows us to distinguish between firms that are small because they are young and firms that are small because they are unproductive. Our model replicates the main features of the labor market, including the distribution of firms and employment by firm size and age, and the response of each type of firm to a corporate tax increase. In the spirit of Weinzierl (2011), we then conduct counterfactual experiments to determine the optimal design of the tax based on both the age and size of firms.[1] We show that a tax modulated according to either age or size is more effective in increasing employment and welfare than a single tax rate. A tax based on size achieves better performance than a tax based on age, but the most favorable outcomes emerge when both size and age criteria are combined. Our simulations show that the optimal tax rate is (i) increasing with firm age, (ii) decreasing with firm size, and (iii) is significantly lower than the benchmark rate for most firm sizes and ages. This latter configuration allows for a substantial increase in employment and welfare without diminishing the fiscal surplus. These results advocate for greater modulation of corporate taxation. While some countries, particularly in Europe, have implemented a reduced tax rate for small businesses, we show that it would be more effective to modulate corporate tax based on the size and age of firms. These results provide new insights into how firms respond to taxes and offer guidance to policymakers on how to design taxes effectively.

The remainder of the paper is structured as follows. Section 2 describes the data and the empirical strategy used to estimate the effect of corporate tax on employment, explores the heterogeneous effects by firm size and age, and presents various robustness tests. Section 3 develops the mechanisms in a toy model and provides analytical results. Section 4 presents the quantitative model. Section 5 presents the estimation of model parameters, compares model-generated moments with empirical data, and examines the effects of corporate tax modulation on firm dynamics, employment, and welfare. The final section concludes.

[1] Note that Weinzierl's work focuses on the modulation of labor income tax (based on the age of workers), not on the modulation of corporate tax (based on the size and age of firms).

2 Empirical framework

2.1 State corporate income tax

Corporate income tax is a tax on business profits. In addition to the federal corporate income tax, many U.S. states impose their own corporate income tax, known as the state corporate income tax. We exploit changes in state corporate income tax rates[2] to identify the effect of corporate taxation on employment.

2.2 Empirical strategy

Unlike existing studies that use standard state-year difference-in-differences models, we adopt the identification strategy proposed by Dube et al. (2010) and Ljungqvist and Smolyansky (2014). This strategy involves comparing contiguous counties situated on opposite sides of a state border, known as contiguous border counties. This strategy has two advantages. First, contiguous counties have very similar characteristics and exhibit similar employment trends, which makes the common trend assumption much more credible (Hagedorn et al., 2015). Second, states comprise a large number of counties, which may have divergent interests. Therefore, it is unlikely that the policy of a state is affected by the economic conditions of a particular county, which mitigates endogeneity issues (Huang, 2008).

2.3 Data

We take advantage of the Business Dynamics Statistics (BDS) for the period 2002 to 2009. The BDS is a longitudinal business database, provided by the U.S. Census Bureau, that tracks firms over time and provides annual measures of business dynamics at the county level (see Appendix A.2). Interestingly, the BDS stands out as one of the few databases in the United States that offers insights into both firm size (permitting the differentiation between small and large firms) and firm age (allowing for the differentiation between young and old firms). Appendix B displays the number of firms and employment by firm size and age.

[2] We use the *Book of the States*, published annually by *The Council of State Governments*, to compile data on corporate taxation at the state level over the period 2002-2009. See Appendix A.3 for more details.

Several noteworthy observations emerge:

i) Most young firms are small (more than 99% of young firms have fewer than 500 employees)

ii) A significant portion of small firms are young (approximately 40% of small firms are 5 years old or younger)

We then use the Book of the States, published by the Council of State Governments, to gather data on corporate taxation at the state level (see Appendix A.3). Finally, we use data from the Population Estimates Program (PEP), provided by the U.S. Census Bureau, which produces county-level estimates of the population by age, sex, and race (see Appendix A.4). All these datasets are merged based on year and state/county FIPS codes.

2.4 Sample

Our empirical strategy requires restricting the sample to states that have a corporate tax system similar to the federal system, and that share a border with at least one other state. For these reasons:

i) We exclude Michigan, Ohio, and Texas, which have a corporate tax system that is not comparable to the federal system

ii) We exclude Alaska and Hawaii, as they do not share a border with another state

Our sample contains 46 states (including the District of Columbia). There are substantial differences in treatment intensity between states and over time. For example, in 2009, the state corporate income tax rates ranged between 0 and 12% (see Figure 1). We then restrict our analysis to contiguous border counties in neighboring states with different corporate income tax rates (see Figure 2). Our final sample comprises 1,022 contiguous border counties. Since each county may belong to multiple county pairs, we have a total of 1,120 distinct county pairs. Some descriptive statistics by state are available in Appendix C.

Figure 1: State corporate income tax rates, 2009

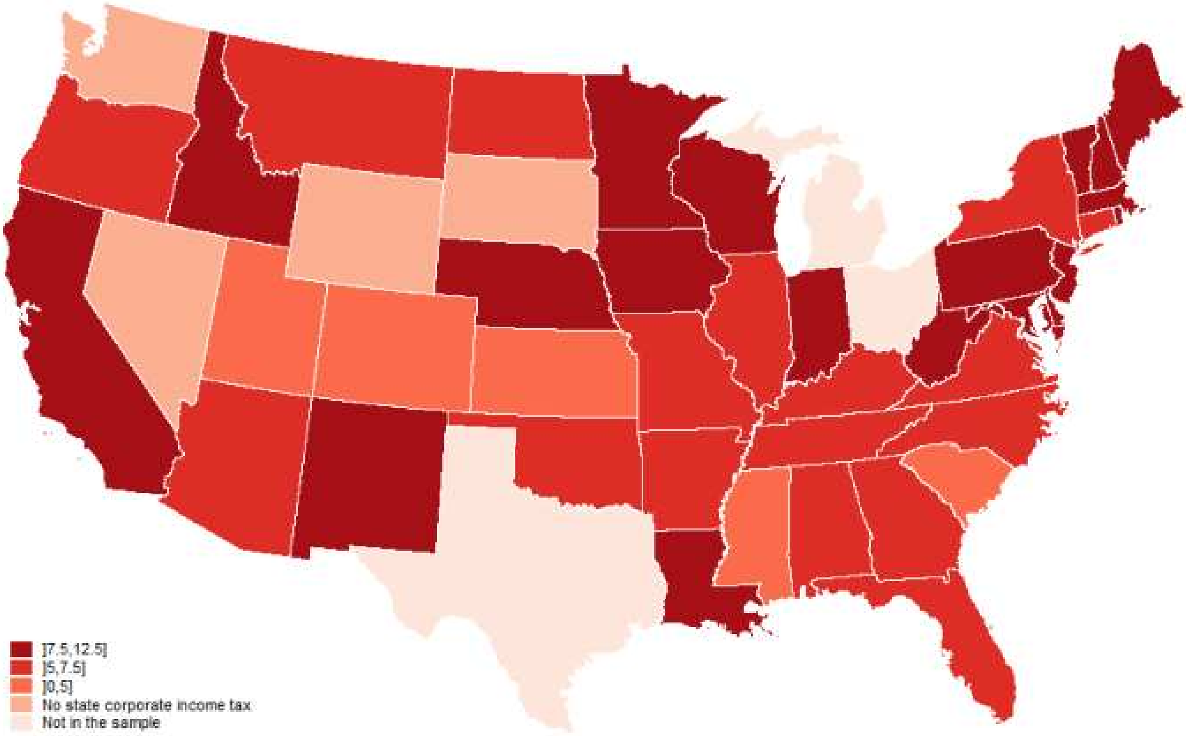
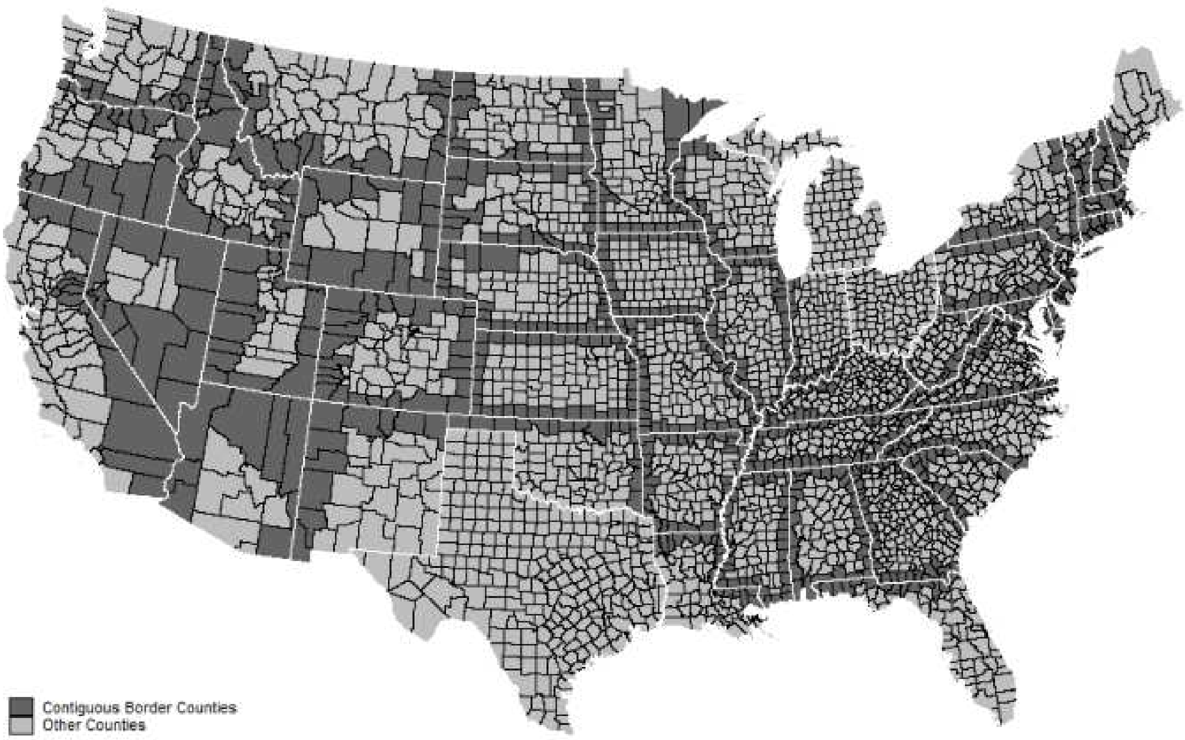


Figure 2: Sample of contiguous border county pairs



2.5 Effect of corporate taxation on employment

We begin by estimating the effect of corporate income tax on total employment. Formally, we estimate the following model:

$$\ln(TotalEmployment_{cpt}) = \alpha + \beta CorporateIncomeTax_{ct} + \gamma X_{ct} + \mu_c + \nu_{pt} + \epsilon_{cpt} \quad (1)$$

where c represents the county, p the county pair, and t the year. $TotalEmployment$ is the total employment, $CorporateIncomeTax$ is the state corporate income tax rate, X is a vector of time-varying county characteristics that captures changes in the composition of the county population (for example, in terms of age, sex, or race), μ_c is a county fixed effect, ν_{pt} is a pair-time fixed effect, and ϵ_{cpt} is a random error term. To account for serial and spatial correlations, we use the two-way clustering method proposed by Dube et al. (2010) and Cameron and Miller (2015) to compute the standard errors.

The estimates are reported in Table 1, column (1). For the full sample, the coefficient associated with the $CorporateIncomeTax$ is negative and significant at the conventional 5% level. According to our estimates, a one percentage point rise in the state corporate income tax rate leads to a 0.78% reduction in total employment.

Table 1: Effect of corporate taxation on employment - Benchmark

BENCHMARK	(1) All firms sample	(2) Large firms sample	(3) Small firms sample	(4) Old firms sample	(5) Young firms sample
Corporate Income Tax	-0.0078**	-0.0010	-0.0135**	-0.0013	-0.0442***
	(0.0038)	(0.0337)	(0.0059)	(0.0058)	(0.0140)
Controls					
Time-varying county char.	YES	YES	YES	YES	YES
County fixed effect	YES	YES	YES	YES	YES
Pair-time fixed effect	YES	YES	YES	YES	YES
Number of periods	8	8	8	8	8
Number of county pairs	1,120	1,120	1,120	1,120	1,120

Table 2: Robustness test 1: Other state policies

ROBUSTNESS TEST 1	(1) All firms sample	(2) Large firms sample	(3) Small firms sample	(4) Old firms sample	(5) Young firms sample
Corporate Income Tax	-0.0077*	-0.0034	-0.0145**	0.0000	-0.0451***
	(0.0041)	(0.0312)	(0.0065)	(0.0057)	(0.0144)
Controls					
Time-varying county char.	YES	YES	YES	YES	YES
County fixed effect	YES	YES	YES	YES	YES
Pair-time fixed effect	YES	YES	YES	YES	YES
Other state policies	YES	YES	YES	YES	YES
Number of periods	8	8	8	8	8
Number of county pairs	1,120	1,120	1,120	1,120	1,120

Table 3: Robustness test 2: Spillover effects

ROBUSTNESS TEST 2	(1) All firms sample	(2) Large firms sample	(3) Small firms sample	(4) Old firms sample	(5) Young firms sample
Corporate Income Tax	-0.0078**	-0.0022	-0.0115***	-0.0076	-0.0426***
	(0.0034)	(0.0235)	(0.0040)	(0.0060)	(0.0157)
Controls					
Time-varying county char.	YES	YES	YES	YES	YES
County fixed effect	YES	YES	YES	YES	YES
Pair-time fixed effect	YES	YES	YES	YES	YES
Number of periods	8	8	8	8	8
Number of county pairs	1,868	1,868	1,868	1,868	1,868

Source: BDS, Book of the States, and PEP (2002-2009)

Robust standard errors, in parentheses, are clustered on the state and border segment levels.

Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

2.6 Heterogeneous effects by firm size and age

We evaluate the heterogeneous impact of corporate income tax on employment by conducting separate analyses based on firm size and age. First, we examine the effect by firm size, considering two subsamples: the *large firms sample*, comprising all firms with 500 employees or more, and the *small firms sample*, encompassing all firms with fewer than 500 employees. We re-estimate the model presented in Equation 1 separately for each subsample and report the results in Table 1, columns (2)-(3). We find no evidence that corporate income tax affects employment in large firms. In contrast, we observe a significant and negative effect on employment in small firms. Our estimates suggest that a one percentage point increase in the state corporate income tax rate decreases employment in small firms by 1.35%.

Next, we examine the impact by firm age. For this purpose, we divide our sample into two subsamples: the *old firms sample*, consisting of all firms aged 6 years or more, and the *young firms sample*, comprising all firms aged 5 years or less. We re-estimate the model presented in Equation 1 separately for each subsample and report the results in Table 1, columns (4)-(5). We do not find any meaningful effect of corporate income tax on employment in old firms. Conversely, we observe a substantial and statistically significant effect on employment in young firms. Our estimates indicate that a one percentage point increase in the state corporate income tax rate leads to a 4.42% reduction in employment among young firms. This result suggests that the effect of corporate income tax on employment mainly comes from young firms. Haltiwanger et al. (2013) demonstrate that young firms contribute disproportionately to net job growth. We go one step further by showing that young firms are more sensitive to changes in corporate income tax rates.

This is an important contribution to the debate about encouraging private sector employment. Based on the belief that small firms create the most jobs, several countries have implemented policies targeted at small firms, such as a reduced corporate income tax rate. As young firms tend to be small, such policies may have a significant (though limited) effect on employment. However, our results show that it is not the small firms but rather the young firms that respond to taxes. Indeed, the pool of small firms contains two types: low-productivity firms that have reached their "adult size", offering limited employment growth potential, and young firms that are in a "growth phase" and represent a major source of job creation. To assess the potential impact of a corporate tax modulated based on both the age and size of firms, we develop a model of firm dynamics in Section 4, where firms differ in terms of age and productivity.

2.7 Robustness tests

2.7.1 Other state policies

Other state policies may compete with changes in corporate taxation in explaining the employment dynamics observed in the data. For example, a state might choose to cut not only corporate income taxes but also other taxes to stimulate economic activity. As shown by Harden and Hoyt (2003), some states adjust the mix of corporate income, individual income, and sales taxes to maintain budget balance. If various components of state tax policy change simultaneously, regressions that include corporate income tax changes but exclude changes in other components of the tax system may be biased. To address the potentially confounding effects of other state policies, we compile data on individual income taxation and sales taxation at the state level[3] over the period 2002-2009.[4] We then incorporate the state tax rates on individual income and sales as additional control variables. As shown in Table 2, the results closely align with the benchmark findings. Consistent with findings from Harden and Hoyt (2003), our estimates indicate that corporate income tax has a significant negative impact on employment while individual income and sales taxes do not. Note that similar results are obtained when including other state policies likely to impact employment, such as minimum wage.

2.7.2 Spillover effects

Our results may be affected by spillover effects if, in response to a state policy, firms and workers move to a county located on the other side of the state border (Tiebout, 1956; Cebula and Alexander, 2006). If individuals are sensitive to tax cuts, they are expected to move to states where taxes are lower. This interstate mobility is more likely if

[3]We use the *Book of the States*, published annually by *The Council of State Governments*, to compile data on individual income taxation and sales taxation at the state level over the period 2002-2009. See Appendix A.3 for more details.

[4]The corporate income, individual income, and sales taxes represent the three primary sources of tax revenue for most states.

agents are located near the state border, where the mobility cost is lower. The incentive to move to another state is expected to decrease with the distance from the border. To test for spillover effects, we follow the strategy adopted by Huang (2008). Instead of comparing two counties sharing a state border, we now compare a county located at the state border (the treated county) with a *co-contiguous interior county* located in the hinterland (the control county). This *co-contiguous interior county* satisfies the following criteria:

1. Not immediately adjacent to the treated county
2. Separated from the treated county by only one county
3. Not located at the state border
4. Situated on the opposite side of the state border

The first two conditions state that the control county is *co-contiguous* to the treated county, with the contiguous border county between them. The last two requirements ensure that the control county is *interior*, i.e., located in the hinterland. For clarity, an example is illustrated in Figure 3. Let's consider two states: New York and Pennsylvania. Suppose the state of New York decides to raise the corporate income tax rate. The county highlighted in red, situated at the border of New York state, can be designated as a treated county. The county highlighted in orange, adjacent to the treated county but located on the other side of the state border, represents a contiguous border county. The county highlighted in green, also located in the state of Pennsylvania and adjacent to the contiguous border county but not to the treated county, serves as a co-contiguous interior county. In this robustness test, this co-contiguous interior county plays the role of the control. If there are spillover effects, they are expected to diminish with distance from the border. Estimates using co-contiguous interior counties as controls (see Figure 4), which are farther from state borders, should be less affected by spillover effects.

We re-estimate the model presented in Equation 1 using this alternative control group and report the results in Table 3. The estimates from this robustness test are very similar to the benchmark, regardless of the subsample considered.[5]

[5]Our conclusions remain unchanged if we include individual income taxation, sales taxation, or the minimum wage as additional control variables.

Figure 3: Example of a co-contiguous county pair

Figure 4: Sample of co-contiguous county pairs

3 Toy model: Understanding the differentiated effects of a corporate tax increase

We first propose to illustrate the differentiated effects of a corporate tax increase, by firm size and age, using a toy model. We consider a two-period model where firms can be either young or old. Firms face convex hiring costs and operating costs that are proportional to the firm's productivity. We assume that operating costs decrease with firm age due to continuous improvements in the production process. Firms are heterogeneous according to their productivity level, which can be low or high. Low-productivity firms have lower operating costs, employ fewer workers, and remain small throughout their lifespan. High-productivity firms incur higher operating costs, start small, and grow significantly with age, becoming large when old. Wages are proportional to the firm's productivity. To save space, the model and the proofs of the results stated below are provided in Appendix D. The key mechanisms and results of the model are presented below and can be illustrated in Figure 5, where n_A (n_B) represents the size of a young (old) high-productivity firm, and where n_C (n_D) represents the size of a young (old) low-productivity firm.

Result 1 For a given technology, young firms are smaller than older firms ($n_A \leq n_B$ and $n_C \leq n_D$). Firms grow over their life cycle for two reasons. First, convex hiring costs encourage firms to spread hiring over both periods. Second, operating costs decrease with age. Profits are higher in the second period, which encourages older firms to hire.

Corollary 1 Firms may grow faster when young than when old ($n_A - 0 \geq n_B - n_A$ and $n_C - 0 \geq n_D - n_C$). Due to convex hiring costs and a finite horizon, firms may opt for a rapid hiring phase to reach their desired production level. This can lead to a concave life-cycle profile, as illustrated in Figure 5. However, this corollary is contingent upon the profile of operating costs throughout the firm's life cycle. If there is a sharp decline in operating costs between the two periods, firms may be more inclined to delay hiring until the second stage of their life cycle, resulting in higher employment growth in old age than in young age.

Result 2 For a given age, less productive firms are smaller than more productive firms ($n_A \geq n_C$ and $n_B \geq n_D$). High-productivity firms face higher operating costs, but their profit is higher due to significantly greater productivity. Consequently, they hire more.

Corollary 2 High-productivity firms grow faster than low-productivity firms between the young and old periods ($n_B - n_A \geq n_D - n_C$). Corollary 2 is a direct consequence of Corollary 1 and Result 2. Low-productivity firms hire few workers and remain small over the life cycle. In contrast, high-productivity firms hire more early on to smooth hiring costs and to reach the desired employment level. Consequently, they experience a stronger and more rapid growth than low-productivity firms.

Results 3 and 4 For a given technology, young firms are more sensitive to a tax increase than older ones $\left(\frac{n_A - n_{A'}}{n_A} \geq \frac{n_B - n_{B'}}{n_B} \text{ and } \frac{n_C - n_{C'}}{n_C} \geq \frac{n_D - n_{D'}}{n_D}\right)$. This results hold in the short-run, *i.e.* when focusing solely on the impact on current hiring decisions, and in the long-run, *i.e.* when also taking into account the impact on employment decisions in previous periods (and therefore the employment inherited from the previous periods). Young firms face higher operating costs which reduces their profits and involves a lower value. Older firms generate larger surplus, which absorb variations in the tax.

Result 5 Small firms are more sensitive to a tax changes than large firms. From Figure 5 it means that total employment variation of small firms is greater (in absolute terms) than that of large firms,

$$\frac{(n_A - n_{A'} + n_C - n_{C'} + n_D - n_{D'})}{n_A + n_C + n_D} \geq \frac{(n_B - n_{B'})}{n_B}$$

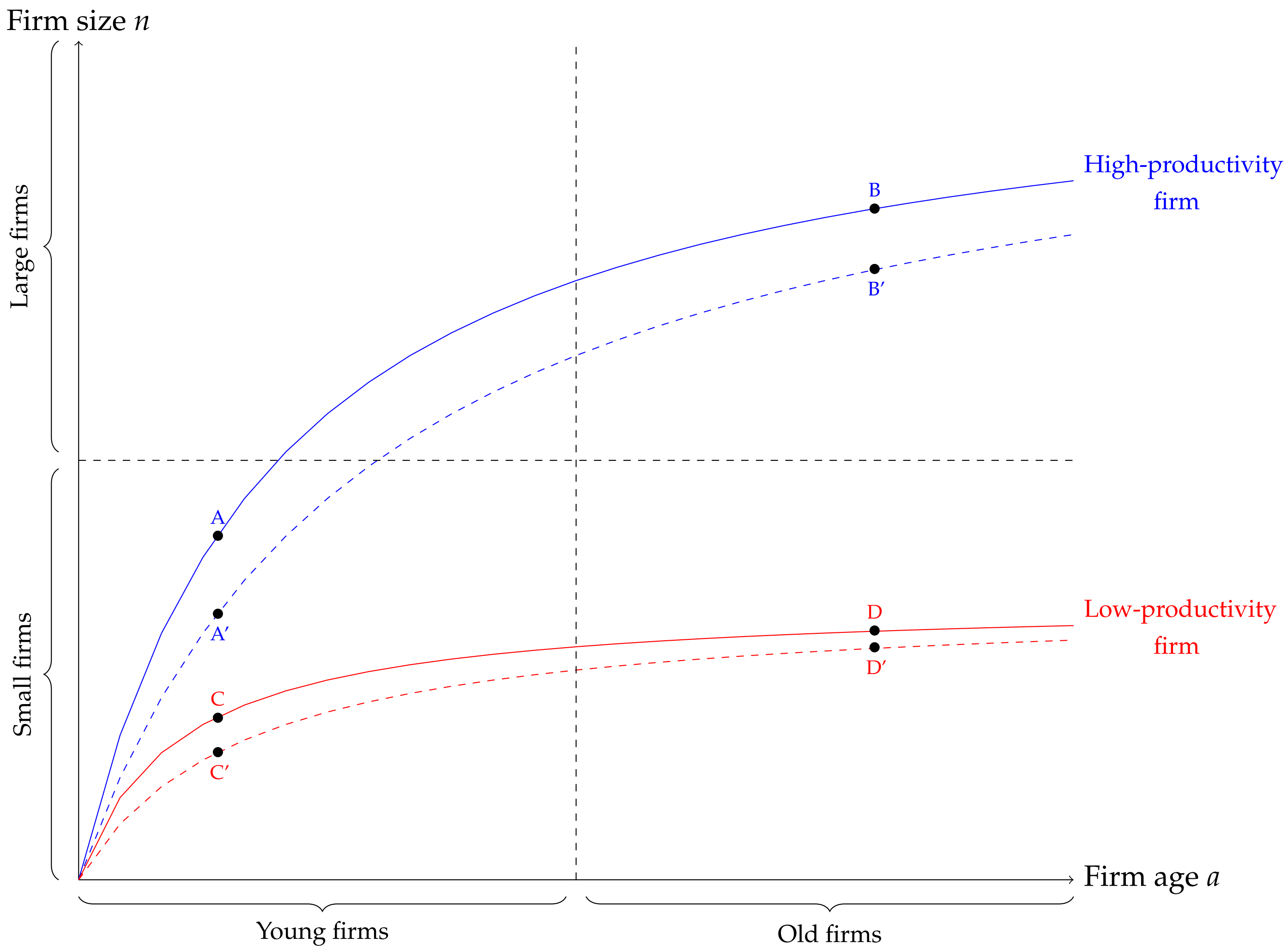
Result 6 Young firms are more sensitive to a tax variation than small firms. In our illustrative example, this involves the comparison of employment $A + C$ to $A + C + D$, that is:

$$\frac{(n_A - n_{A'} + n_C - n_{C'})}{n_A + n_C} \geq \frac{(n_A - n_{A'} + n_C - n_{C'} + n_D - n_{D'})}{n_A + n_C + n_D}$$

.

This toy model helps to rationalize the empirical facts presented in Section 2. However, it is too simple to allow for a quantitative assessment of a corporate tax reform. In particular, this toy model ignores the existence of search frictions and assumes that firm entry and exit are exogenous. In the following section, we develop a more sophisticated model that includes search frictions and where firm entry and exit are endogenous, allowing for an analysis of the impact of corporate taxation on the distribution of firms. We also extend the firms' life cycle and add heterogeneity to replicate the dynamics and distribution of firms observed in the data. We then use this model to analyze the effects of different corporate tax schedules.

Figure 5: An illustrative example



Note: Solid line: before corporate tax increase; Dashed line: after corporate tax increase

4 Model

We develop a life-cycle model of firm dynamics to analyze firms' responses to an increase in corporate tax, taking into account their size and age. In the spirit of Elsby and Michaels (2013), firms may employ multiple workers and have different sizes. We follow Sedláček and Sterk (2017) and Sedláček (2020) in incorporating the concept of firm age. Firms can differ in size due to differences in technology, age, or idiosyncratic productivity shocks. Entry and exit of firms are endogenous, as in Hopenhayn (1992). Finally, we consider a frictional labor market with directed search, in the spirit of Menzio et al. (2016).

4.1 Heterogeneity

Size. Firms can have multiple employees. We denote the number of employees within a firm as n. For simplicity, we will subsequently use "firm size" to refer to the number of employees in a firm. Let $\mathcal{N} = \{n_1, \ldots, n_N\}$ represent the set of possible employment levels ($n_i = i$, with $i = 1, \ldots, N$), where n_N is the upper bound. Due to idiosyncratic productivity shocks, each firm adjusts its employment level differently, resulting in heterogeneous firms and an endogenous distribution of firm sizes.

Age. Firms are characterized by an age denoted by a. Let $\mathcal{A} = \{a_1, \ldots, a_A\}$ represent the set of possible ages ($a_i = i$, with $i = 1, \ldots, A$). Firms enter the market at age $a = a_1$. Throughout the firm's life cycle, the transition from age $a = a_i$ to the next age $a' = a_{i+1}$ occurs every period. When the firm reaches the maximum age $a = a_A$, it remains there until its closure or destruction.

Idiosyncratic productivity shock. Firms are subject to idiosyncratic productivity shocks. Let $\varepsilon \in \mathcal{E} =]0, \infty[$ denote the idiosyncratic component of firm productivity. Each period, existing firms draw a new productivity ε' from the conditional distribution $G(\varepsilon'|\varepsilon)$. Entering firms, in contrast, draw a productivity ε' from the unconditional distribution $G_0(\varepsilon')$.

Technology. Firms are characterized by a technology level $x \in \mathcal{X} = \{x_1, x_2, \ldots, x_X\}$, where the productivity of a firm with technology x_i is lower than the productivity of a firm with technology x_{i+1}. More precisely, x represents the willingness or ability to adopt an efficient production technology. A higher value of x indicates greater productivity and a higher potential for growth. The technology level x is permanent and does not change throughout the firm's life cycle. The entry decision and the choice of x are endogenous and will be described later.

Labor force by skill level. We assume that the distribution of the labor force by skill level is exogenous. Let $L(x)$ denote the size of the labor force with skill level x. The total labor force, L, satisfies $L = \sum_x L(x)$.

4.2 Search and matching

The economy is characterized by matching frictions. Following Menzio et al. (2016), we assume that the market is segmented by technology x. Workers in sector x apply only to jobs with technology x. Hirings depend on the number of unemployed workers, $u(x)$, and the total number of vacancies, $\boldsymbol{v}(x)$, posted by all firms in sector x. These inputs are combined in the following matching function:

$$m(x) = m(u(x), \boldsymbol{\nu}(x)) \tag{2}$$

The matching function is increasing and concave in its two arguments and exhibits constant returns to scale. The job-finding rate and the job-filling rate are given, respectively, by:

$$f(x) = \frac{m(x)}{u(x)} = m(1, \theta(x)) \tag{3}$$

$$q(x) = \frac{m(x)}{\boldsymbol{\nu}(x)} = m(\theta(x)^{-1}, 1) \tag{4}$$

where $\theta(x)$ represents the labor market tightness, defined as follows:

$$\theta(x) = \frac{\boldsymbol{\nu}(x)}{u(x)} \tag{5}$$

4.3 Value functions

We denote by $\Pi(n,\varepsilon,a,x)$ the value of a firm in state (n,ε,a,x). Let $W(n,\varepsilon,a,x)$ and $U(x)$ denote the values of employed workers and unemployed workers, respectively. It is useful to define the conditions for worker acceptance and firm existence.

Workers accept to work if $W(n,\varepsilon,a,x) \geq U(x)$, which can be summarized by the following indicator function:

$$\mathcal{I}_w(n,\varepsilon,a,x) = \begin{cases} 1 & \text{if} \quad W(n,\varepsilon,a,x) \geq U(x) \\ 0 & \text{otherwise} \end{cases}$$

Let c_x denote the exit cost for a firm. Firms continue to exist if $\Pi(n,\varepsilon,a,x) \geq -c_x$, which can be summarized by the following indicator function:

$$\mathcal{I}_f(n,\varepsilon,a,x) = \begin{cases} 1 & \text{if} \quad \Pi(n,\varepsilon,a,x) \geq -c_x \\ 0 & \text{otherwise} \end{cases}$$

The joint decision determines the existence of a positive level of employment:

$$\mathcal{I}(n,\varepsilon,a,x) = \mathcal{I}_w(n,\varepsilon,a,x) \times \mathcal{I}_f(n,\varepsilon,a,x) \tag{6}$$

The expected values of the firm and the worker are, respectively:

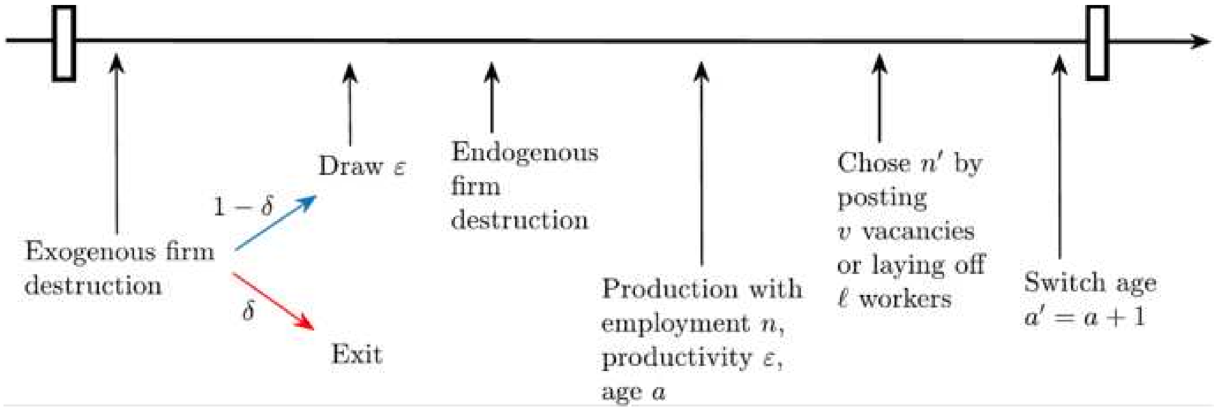
$$\Lambda(n,\varepsilon,a,x) = \mathcal{I}(n,\varepsilon,a,x)\Pi(n,\varepsilon,a,x) + (1-\mathcal{I}(n,\varepsilon,a,x))(-c_x) \tag{7}$$

$$\Omega(n,\varepsilon,a,x) = \mathcal{I}(n,\varepsilon,a,x)W(n,\varepsilon,a,x) + (1-\mathcal{I}(n,\varepsilon,a,x))U(x) \tag{8}$$

4.3.1 Incumbent firms

The timing of events is illustrated in Figure 6. Consider a firm characterized by the tuple (n,ε,a,x). At the beginning of the period, the firm may be destroyed exogenously at rate δ. If it survives, the firm draws a new productivity ε' from the conditional distribution $G(\varepsilon'|\varepsilon)$. The firm may then be destroyed endogenously or continue operating, depending on the productivity shock. The firm produces and decides how many vacancies to open or how many employees to lay off. At the end of the period, the firm advances to the next age, unless it is already at the last age.

Figure 6: Timing of events



Firm program. A homogeneous consumption good, traded in a competitive market, is produced by firms using only labor.[6] The production of a firm with employment level n is given by:

$$y(n, \varepsilon, x) \quad = \quad (n \, \varepsilon \, \bar{y}(x))^{\alpha} \tag{9}$$

where $\alpha \in]0, 1[$ is the elasticity of production with respect to the effective labor $n\varepsilon\bar{y}(x)$. Here, $\bar{y}(x)$ represents a technology-augmenting productivity parameter. In the spirit of Bagger et al. (2014) and Blundell et al. (2016), wages are determined by the following rule:

$$w(\varepsilon, x) \quad = \quad \xi(\varepsilon \, \bar{y}(x))^{\gamma} \tag{10}$$

where $\gamma \geq 0$ governs the curvature of the wage with respect to the productivity component $\varepsilon\bar{y}(x)$. Here, $\xi > 0$ is a scaling parameter. Instantaneous profits are given by:

$$\pi(n, \varepsilon, x) \quad = \quad y(n, \varepsilon, x) - w(\varepsilon, x)n \tag{11}$$

The firm adjusts its workforce by determining the optimal level of employment. The firm has four exclusive options. Firstly, it can hire workers by posting vacancies. Let v denote the number of posted vacancies.[7] The firm incurs a convex cost of $c_v(x)v^{\phi}$,

[6] For the sake of clarity, we do not consider labor adjustments along the intensive margin.

[7] Note that v represents the number of vacancies posted by a single firm, while $\boldsymbol{v}$ represents the total

where $\phi \geq 1$. Given a vacancy filling rate $q(x)$, the number of hired workers is $q(x)v$, and n' is such that $n' = n + q(x)v$. Secondly, the firm may decide to lay off workers, incurring a cost of c_d per laid-off worker. Let ℓ denote the number of laid-off workers. Employment then evolves according to $n' = n - \ell$. Thirdly, the firm may decide to remain inactive regarding labor adjustment and maintain its current labor force, such that $n' = n$. Fourthly, the firm may close. This occurs either at the exogenous rate δ, or if the firm's value is less than or equal to the exit cost c_x.

The firm determines the optimal employment level n', given a state (n, ε, a, x), by solving the following program:

$$\begin{aligned}
\Pi(n,\varepsilon,a,x) &= \max_{n',v,\ell} \Big\{ [\pi(n,\varepsilon,a,x) - c_v(x)v^{\phi} - c_d\ell](1-\tau) - c_o(a,x)n \\
&+ \beta(1-\delta)\int_{\varepsilon'} \Lambda(n',\varepsilon',a',x)dG(\varepsilon'|\varepsilon) \Big\} \qquad (12) \\
\text{with } n' &= n + q(x)v - \ell \qquad (13) \\
v &\geq 0 \qquad (14) \\
\ell &\geq 0 \qquad (15)
\end{aligned}$$

Each period, existing firms must pay an operating cost denoted by $c_o(a,x)$, which depends on age and technology. We assume that the operating cost increases with the technology used by the firm, i.e. $\partial c_o(a,x)/\partial x \geq 0$. We also assume that older firms optimize their production process, leading to lower operating costs, i.e. $\partial c_o(a,x)/\partial a \leq 0$. τ is the corporate tax rate. Equation (13) describes the evolution of the workforce. Vacancies and layoffs are constrained to be non-negative (Equations (14) and (15)). Note that vacancies and layoffs cannot be simultaneously strictly positive. The firm's decision rules are expressed as follows:

$$\begin{aligned}
v &= \mathcal{D}_v(n,\varepsilon,a,x) \\
\ell &= \mathcal{D}_\ell(n,\varepsilon,a,x) \\
n' &= n + q(x)v - \ell \equiv \mathcal{D}(n,\varepsilon,a,x) \qquad (16)
\end{aligned}$$

number of vacancies posted by all firms.

4.3.2 New firms

We consider a specific skill level x (the determination of which will be explained later). Upon entry, the new firm decides how many vacancies to post. The firm determines the optimal employment level n', given the technology x, by solving the following program:

$$\begin{aligned}
\Pi_0(x) &= \max_{n',v} \left\{ -c_v(x)v^{\phi} + \beta \int_{\varepsilon'} \Lambda(n', \varepsilon', a_1, x) dG_0(\varepsilon') \right\} && (17)\\
\text{with } n' &= q(x)v && (18)\\
v &\geq 0 && (19)
\end{aligned}$$

The hiring process is the same as that of an incumbent firm. Matches occur at a rate of $q(x)$ and the firm begins operating with a size of $n' = q(x)v$ and age $a = a_1$. The firm's decision rules are expressed as follows:

$$\begin{aligned}
v &= D_{0,v}(x)\\
n' &= q(x)v = \mathcal{D}_0(x) && (20)
\end{aligned}$$

4.3.3 Workers

Employed. Recall that a worker is characterized by a type x, which is permanent. The employed worker value function is given by:

$$\begin{aligned}
W(n,\varepsilon,a,x) &= \underbrace{w(\varepsilon,x)}_{\text{wage}} + \underbrace{\beta\delta U(x)}_{\text{Destruction}} + \underbrace{\beta(1-\delta)\mathbb{1}\{n' \geq n\} \int_{\varepsilon'} \Omega(n',\varepsilon',a',x) dG(\varepsilon'|\varepsilon)}_{\text{Firm maintains or increases employment}}\\
&+ \underbrace{\beta(1-\delta)\mathbb{1}\{n' < n\} \int_{\varepsilon'} \left(\frac{n-n'}{n} U(x) + \frac{n'}{n} \Omega(n',\varepsilon',a',x) \right) dG(\varepsilon'|\varepsilon)}_{\text{Firm lowers employment}}
\end{aligned}$$

An employed worker receives a wage $w(\varepsilon, x)$. In the next period, the job status of an employed worker may change due to either exogenous separation or decisions made by agents. The firm can be destroyed at rate δ, in which case the worker becomes unemployed. The firm may also decide to reduce its workforce. Note that every employed worker within the firm has the same probability of being laid off, that is $\frac{n-n'}{n}$ where $n' < n$. With a probability of $\frac{n'}{n}$ the worker is not laid off and must decide whether to remain with the firm or leave and become unemployed, as indicated by the optimal decision in Equation (8). If the firm decides to increase employment, the employed worker then decides whether to continue the match with the (larger) firm or to terminate the relationship.

Unemployed. Unemployed workers receive unemployment benefits $b(x)$. Each unemployed worker receives a contact with a firm at rate $f(x)$. The value function of an unemployed worker is:

$$U(x) = b(x) + \beta \left(f(x)\overline{W}(x) + (1 - f(x))U(x) \right) \tag{21}$$

where $\overline{W}(x)$ is the expected value of the worker upon contact with a firm. The worker does not know the age, size, and productivity of the firm that may hire her. However, she knows the firm distribution $\lambda(n, \varepsilon, a, x)$, and the expected value $\bar{W}(x)$, which depends on the distribution of vacant jobs, as follows:

$$\begin{aligned} \overline{W}(x) &= \sum_a \int_n \int_\varepsilon \left[\omega(n, \varepsilon, a, x) \int_{\varepsilon'} \Omega(n', \varepsilon', a', x) dG(\varepsilon'|\varepsilon) \right] d\varepsilon dn \\ &+ \omega_0(x) \int_{\varepsilon'} \Omega(n', \varepsilon', a_1, x) dG_0(\varepsilon') \end{aligned}$$

where $\omega(n, \varepsilon, a, x)$ represents the probability of making contact with an existing firm in the state (n, ε, a, x), and $\omega_0(x)$ represents the probability of making contact with a new firm of type x. Using Equation (57), these probabilities can be easily determined:

$$\begin{aligned} \omega(n, \varepsilon, a, x) &= \frac{\lambda(n, \varepsilon, a, x) D_v(n, \varepsilon, a, x)}{\boldsymbol{\nu}(x)} \\ \omega_0(x) &= \frac{e(x) D_{0,v}(x)}{\boldsymbol{\nu}(x)} \end{aligned}$$

4.4 Entry and exit

Entry. As in Hopenhayn (1992), firm entry is endogenous. Let $c_e(x)$ be the entry cost of a firm of type x. Without entry friction, new firms enter the market until profit opportunities are exhausted, that is, as long as:

$$\Pi_0(x) \geq c_e(x) \tag{22}$$

New firms enter the market until the equilibrium condition $\Pi_0(x) = c_e(x)$ is satisfied. This condition determines the mass of entering firms of type x, denoted by $e(x)$. We follow Sedláček and Sterk (2017) and Sedláček (2020) by considering that an additional friction alters the entry process. This entry friction may refer to a coordination problem among agents or may indicate that a bit of luck is needed to create a firm, resulting in the effective number of entrants being lower than the number of entry attempts. This entry friction also aligns with the work of Saint-Paul (2002) and Klette and Kortum (2004), and more broadly, with the literature on innovation and investments in R&D, and firm dynamics like Acemoglu et al. (2018). In the spirit of Moll (2017), we consider that the mass of entering firms $e(x)$ satisfies the following condition:

$$e(x) = \frac{\eta_0(x)}{1+\eta_1(x)e^{-\eta_2(x)(\Pi_0(x)-c_e(x))}} \tag{23}$$

where $\eta_i(x),\ i=0,1,2$ are parameters depending on the technology x. Note that, as mentioned in Moll (2017), when $\eta_1 \longrightarrow +\infty$, or when $\eta_2 \longrightarrow +\infty$ we have the Hopenhayn model where the supply of entrants is infinitely elastic. More precisely, as the elasticity grows, the number of entries increases, and asymptotically, it converges toward the value obtained in Hopenhayn model.

Exit. Firms exit the market if $\Pi(.) < c_x$. Let $\mathcal{D}_e$ denote the decision rule governing firm exit. $\mathcal{D}_e$ takes the value of 0 in the case of an exit and 1 otherwise. $\mathcal{D}_e$ is simply equal to the firm's existence condition:

$$\mathcal{D}_e(n,\varepsilon,a,x) = \mathcal{I}_f(n,\varepsilon,a,x) \tag{24}$$

4.5 Stationary distribution

Let $\lambda(n,\varepsilon,a,x)$ be the endogenous distribution of firms. Due to entry and exit of firms, $\lambda(.)$ is a state variable for the aggregate economy. The law of motion is given by:

$$\lambda(n',\varepsilon',a',x) = \mathcal{I}(n',\varepsilon',a',x)\left\{\begin{array}{l}(1-\delta)\int_n\int_\varepsilon \mathbb{1}\{n'=\mathcal{D}(n,\varepsilon,a,x)\}\lambda(n,\varepsilon,a,x)g(\varepsilon'|\varepsilon)d\varepsilon dn \\ \\ +e(x)g_0(\varepsilon')\mathbb{1}\{n'=\mathcal{D}_0(x)\}\mathbb{1}\{a'=a_1\}\end{array}\right\} \tag{25}$$

The stationary distribution (66), along with the definition of the mass of entrants (65), imply that the number of entering firms equals the number of exiting firms at steady state. However, the number of firms in the market can vary over time, for example, in response to public policies.

4.6 Government budget

The government collects corporate taxes and makes transfers to unemployed workers. The fiscal surplus, FS, is given by:

$$FS = \underbrace{\sum_x \sum_a \int_n \int_\varepsilon \lambda(n,e,a,x)\left[\pi(n,\varepsilon,a,x) - c_v(x)v^\phi - c_d l\right]\tau \, d\varepsilon dn}_{\text{Corporate income tax revenues}} \underbrace{- \sum_x b(x)u(x)}_{\text{UI benefits}} \quad (26)$$

where $v = \mathcal{D}_v(n,\varepsilon,a,x)$ and $l = \mathcal{D}_l(n,\varepsilon,a,x)$.

The government budget is balanced with lump-sum transfers T:

$$T = FS \quad (27)$$

4.7 Equilibrium

As in Elsby and Michaels (2013), the condition for the aggregate steady-state equilibrium is obtained by determining the mass of workers and the job-finding probabilities such that the number of matches equals the number of separations. Let $N(x)$ be the aggregate employment level determined using the distribution of firms:

$$N(x) = \sum_a \int_n \int_\varepsilon \lambda(n,\varepsilon,a,x) \times n \, d\varepsilon dn \quad (28)$$

The number of unemployed workers and the number of vacancies are given, respectively, by:

$$u(x) = L(x) - N(x), \quad (29)$$

$$\boldsymbol{\nu}(x) = \sum_a \int_n \int_\varepsilon \lambda(n,\varepsilon,a,x)\mathcal{D}_v(n,\varepsilon,a,x) \, d\varepsilon dn + e(x)\mathcal{D}_{0,v}(x), \quad (30)$$

where $L(x)$ is the labor force, which is assumed to be constant. We then derive the values of labor market tightness, the job-finding rate, and the job-filling rate from Equations (59)-(61).

Let $s(x)$ be the number of separations, defined as follows:

$$\begin{aligned} s(x) &= \sum_a \int_n \int_\varepsilon \lambda(n,\varepsilon,a,x) n \Bigg\{ \Big(\delta n + (1-\delta) \Big[\mathbb{1}\{n' < n\} \\ &+ \mathbb{1}\{n' \geq n\} \int_{\varepsilon'} (1 - \mathcal{I}(n',\varepsilon',a',x)) dG(\varepsilon'|\varepsilon) \Big] \Bigg\} d\varepsilon dn. \end{aligned} \tag{31}$$

Note that at steady state, the inflow and outflow of unemployment are equal, meaning that $f(x)u(x) = s(x)$.

DEFINITION 1. *Given exogenous processes for age a and idiosyncratic productivity* ε*, the equilibrium is a list of (i) quantities* $m(x)$, $u(x)$, $\boldsymbol{v}(\boldsymbol{x})$, $N(x)$, $f(x)$, $q(x)$, $\theta(x)$ *and* $s(x)$*; (ii) optimal decisions* $D(n,\varepsilon,a,x)$, $D_0(x)$*; (iii) entry mass* $e(x)$*; (iv) stationary distribution of firms* $\lambda(n,\varepsilon,a,x)$*; and (v) fiscal surplus FS and lump-sum transfer T, satisfying the following conditions:*

(i) $m(x)$, $u(x)$, $\boldsymbol{v}(\boldsymbol{x})$, $N(x)$, $f(x)$, $q(x)$, $\theta(x)$ *and* $s(x)$ *are the solutions of the matching function (55), the number of unemployed workers (56), the number of vacancies (57), the aggregate level of employment (58), the job-finding rate (59), the vacancy-filling rate (60), the tightness (61), and the separation flows (62), respectively;*

(ii) The firms' decision rules $D(n,\varepsilon,a,x)$ *(incumbent firms) and* $D_0(x)$ *(new firms), are solution to the problems (63), and (64),respectively;*

(iii) $e(x)$ *satisfies (65);*

(iv) The stationary distribution $\lambda(n,\varepsilon,a,x)$ *solves (66);*

(v) FS and T satisfy the government budget defined by (67) and (68).

5 Quantitative analysis

5.1 Functional forms

- We assume that unemployed workers and vacancies are matched according to the following matching function:

$$m(u(x), \boldsymbol{v}(x)) \quad = \quad \chi(x)u(x)^{\rho}\boldsymbol{v}(x)^{1-\rho},$$

  where $\chi(x)$ is the matching efficiency and ρ is the elasticity of the matching function with respect to unemployment.

- The operating cost function is age- and sector-dependent, and is defined by the following logistic function:

$$c_o(a, x) \quad = \quad \psi_0(x) + \frac{\psi_0(x)(\Delta_o - 1)}{1 + \psi_1 e^{-\psi_2 a}},$$

  where $\psi_0(x)$ governs the initial cost level, and $\Delta_o \in [0, 1]$ is the decline in operating costs over the life cycle. ψ_1 controls the firm age at which the operating cost begins to decline, and ψ_2 corresponds to the pace at which this decline occurs.

5.2 Calibration and estimation

Our strategy is to first calibrate some parameters using external information, and then estimate the remaining parameters using the simulated method of moments. Additional information on the procedure is provided in the supplementary appendix.

5.2.1 Parameters set externally

State space. We consider a life-cycle of 11 years. With quarterly frequencies, this gives us $n_a = 45$ age periods. The discount factor β is set to 0.99, implying an annual interest rate of around 4%. The employment grid includes $n_n = 201$ points ranging from 0 to 1. Since there are more small firms than large ones, we assume a non-linear grid with more points for small employment levels and fewer points for larger employment levels. We consider three levels of technology ($n_x = 3$). Finally, we assume that idiosyncratic productivity evolves following an AR(1) process, which we discretize using the

Rouwenhorst method with $n_\varepsilon = 10$ points.

Labor market. We set the firing cost c_d to 0.1, which corresponds to 6% of the average wage. As is standard in the search and matching literature, the elasticity of the matching function ρ is set to 0.5. The exogenous rate of destruction δ is calibrated to 1%, providing an exit rate of firms in line with the data on a quarterly basis.[8] We set the vacancy filling rate $q(x)$ to 71%, as in den Haan et al. (2000), and the non-employment rate $ur(x) = u(x)/L(x) = 28.7\%$ to match the observed employment rate in the data of 71.3%. To match the targets, $\chi(x)$ is adjusted to balance the matching function. The entry cost $c_e(x)$ is such that the entry condition (22) is binding at equilibrium.

Labor market institutions. The federal corporate tax is 21%, while the state corporate tax varies between 0 and 12%, with an average of 6.85%. We set non-employment benefits to 60% of the average wage, a value slightly higher than the weekly benefits amount for unemployment insurance in order to account for other sources of transfers and benefits. The calibrated parameters are reported in Table 4.

Table 4: CALIBRATED PARAMETERS

Parameter	**Symbol**	**Value**
Discount factor	β	0.99
Firing cost	c_d	0.10
Exogenous destruction rate	δ	0.010
Matching function elasticity	ρ	0.50
Federal corporate tax rate	τ_f	0.21
State corporate tax rate	τ_s	0.07
Unemployment benefits	$b(x)$	[0.24, 0.66, 1.82]
Entry cost	$c_e(x)$	[0.08, 1.56, 31.52]
Matching efficiency	$\chi(x)$	[0.20, 0.21, 0.21]

5.2.2 Parameter set internally

We estimate the remaining parameters using the simulated method of moments. We have five parameters that depend on the technology level x and eight parameters that are common across technology groups. In total, we have 23 parameters to estimate. The set of structural parameters is given by:

$$\Theta = \{c_v(x), \bar{y}(x), \psi_0(x), \bar{e}(x), \eta(x), \phi_v, \gamma, \rho_\varepsilon, \sigma_\varepsilon, \alpha, \Delta_o, \psi_1, \psi_2\}$$

[8] Note that part of the exit rate is endogenous.

The objective is to replicate the following moments:

(A) The share of firms by firm size and age

(B) The share of employment by firm size and age

(C) The job creation rate, job destruction rate, and exit rate by firm size and age

(D) The elasticity of employment with respect to the corporate tax by firm size and age

For moments (A) to (C), we have 4 size groups (1-19 , 20-99, 100-499, 500+ employees) and 5 age groups (0-1, 2-3, 4-5, 6-10, 11+ years old). For moments in (D), we consider only two firm size groups: fewer than 500 employees (small), and 500 employees or more (large). We also consider only two age groups: 5 years or less (young), and 6 years or more (old). This restriction allows for a direct comparison with the micro-econometric estimation from Section 2.5. In total, we have 49 moments.

The optimization procedure consists of finding the vector of parameters Θ that minimizes the distance between the vector of moments simulated from the model $\mathcal{M}^m(\Theta)$ and the vector of moments from the data $\mathcal{M}^d$. The SMM estimator $\hat{\Theta}$ solves:

$$\hat{\Theta} \;=\; \left[\mathcal{M}^m(\Theta) - \mathcal{M}^d\right]' \boldsymbol{W} \left[\mathcal{M}^m(\Theta) - \mathcal{M}^d\right],$$

where $\boldsymbol{W}$ is a weighting matrix.

Table 5 displays the values of the parameters. The level of hiring cost c_v is roughly similar between low and middle technology firms, but higher for firms with higher technology. There is a significant discrepancy in TFP levels across firms with different technologies, with a factor of around 100 between the least and most productive firms. This strong heterogeneity in $\bar{y}$ seems necessary to replicate the dispersion of firms across sizes. The levels of operating costs are found to be consistent with the heterogeneity in TFP, with the cost for new firms with the highest technology being around 7 times larger than that for new firms with the lowest technology level. The entry mass level η_0 is significantly larger for firms with a low level of technology. This result is not surprising since the majority of firms are small. In our model, low technology adoption results in a low growth rate and smaller size. The elasticity of entry, governed by η_2, reveals that turnover is high among firms with low technology adoption. Since small and less productive firms are more sensitive to shocks, their entry and exit rates are higher. For parameters common across technology levels, the estimation indicates a strong curvature in hiring costs, with $\phi_v = 3.25$. However, the wage curvature $\gamma = 0.44$ is found to be rather small, potentially leading to large surplus (which is also a condition for the existence of large firms). The estimated process for the idiosyncratic productivity shocks involves a persistence of $\rho_\varepsilon = 0.90$ and a standard deviation of $\sigma_\varepsilon = 0.05$, in line with several studies estimating an AR(1) process for individual productivity. Lastly, the estimation involves a decline in operating cost with age of around two-thirds, with the rate of this decline (ψ_2) being rather fast.

Table 5: ESTIMATED PARAMETERS

PARAMETER	SYMBOL	TECHNOLOGY		
		Low	**Mid**	**High**
Hiring cost level	c_v	0.50	0.49	0.59
TFP	$\bar{y}$	0.10	0.99	9.97
Operating cost level	ψ_0	0.30	0.80	2.00
Entry param. (×100)	η_0	0.949	0.087	0.025
Entry param.	η_1	0.07	0.07	0.005
Entry param.	η_2	75.91	3.18	0.14
		COMMON		
Hiring cost curvature	ϕ_v	3.25		
Wage curvature	γ	0.44		
Persistence - productivity shock	ρ_ε	0.90		
Std - productivity shock	σ_ε	0.05		
Output elasticity	α	0.69		
Operating cost decline	Δ_o	0.33		
Operating cost decay	ψ_1	1.70		
Operating cost growth rate	ψ_2	1.02		

5.2.3 Model versus Data

Figure 7 shows how the model performs in replicating key moments from the data. The model perform extremely well in matching the distribution of firms and employment by size groups (Panel a). The vast majority of firms are small. Firms with 500 employees or more (large firms) represent less than 1% of all firms but account for 30% of total employment. The model accurately captures the share of firms and the share of employment by firm size, as well as by firm age (Panel b). Most firms are old, and they represent a significant portion of employment. The model also performs reasonably well in capturing the firm exit rate by size and age, which tends to be slightly higher for small and young firms. The model predicts a strong decline in job turnover for large firms (Panel c), whereas the data exhibit higher rates. Additionally, the model matches the job creation and job destruction rates by firm age quite well (Panel d). It reproduces the strong job creation rate observed in the early stages of the life cycle, followed by stabilization. Furthermore, it reasonably captures the evolution of the job destruction rate, which peaks around ages 3-4 and declines gradually thereafter.

Figure 7: Distribution of firms and employment

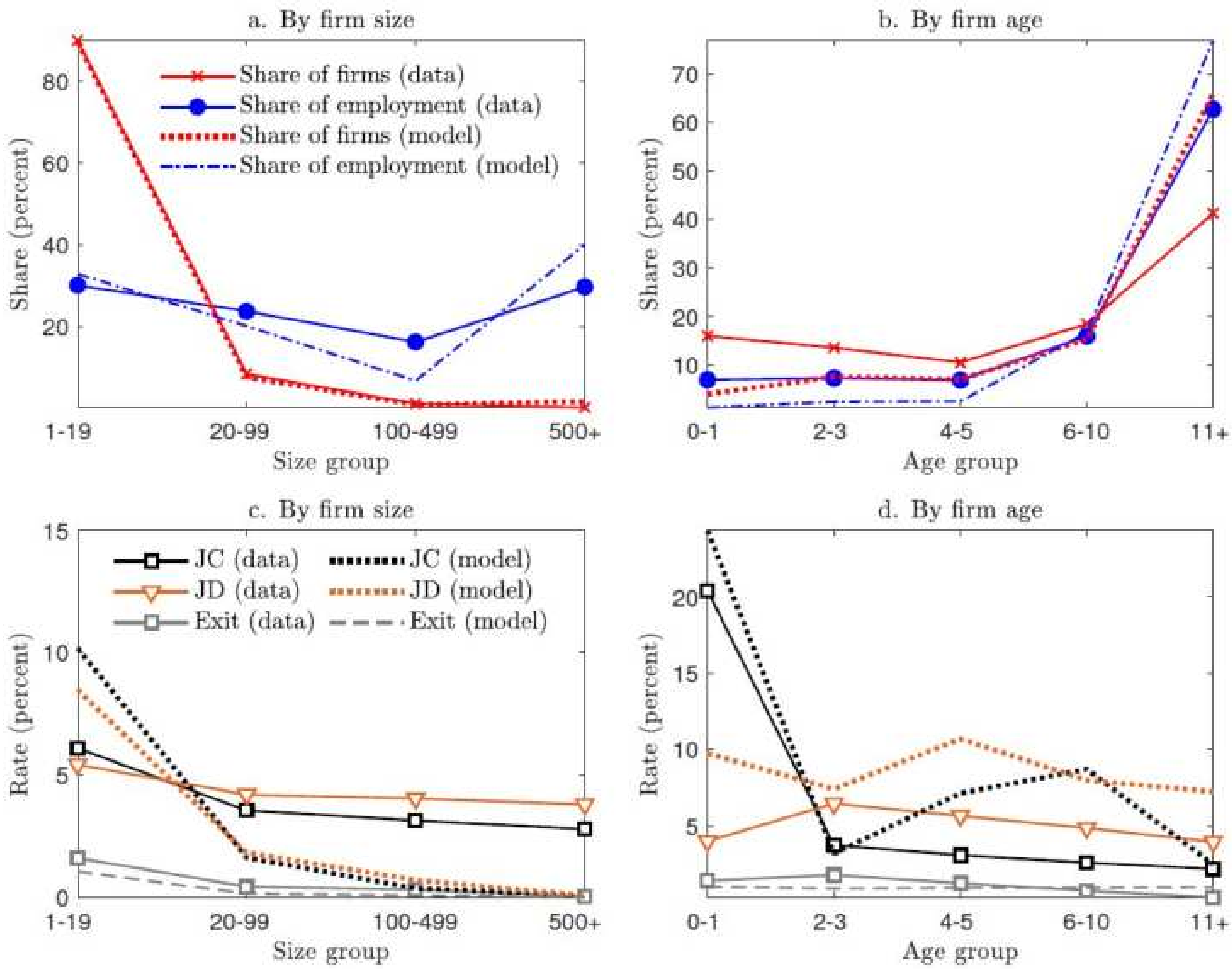


Table 6 provides a set of targeted and non-targeted moments. The targeted moments correspond to the elasticity of employment with respect to the corporate income tax. The model accurately captures the observed strong differences in elasticities. In particular, employment in young firms, and to some extent in small firms, is significantly more sensitive to variations in the corporate tax rate compared to large or old firms. For the non-targeted moments, it is shown that the model matches the long-run values of standard aggregate variables like the employment rate and the separation rate. However, it generates a slightly lower vacancy rate compared to that observed in the JOLTS data. Lastly, the model produces sufficient wage disparities to replicate inter-decile ratios D5/D1 and D9/D5, but overstates D9/D1.

Table 6: MOMENTS

Variable	Model	Data	Data source
LABOR MARKET VARIABLES			
Employment rate	71.30	71.30	BLS
Vacancy rate	1.37	3.12	JOLTS
Separation rate	2.59	3.75	JOLTS
INTERDECILE WAGE RATIOS			
D9/D1	8.04	4.98	CPS
D5/D1	3.00	2.11	CPS
D9/D5	2.68	2.36	CPS
Gini	0.37	0.37	CPS
ELASTICITY OF EMPLOYMENT WITH RESPECT TO THE CORPORATE TAX			
Total	-0.75	-0.78	Own calculation
Small firms	-1.19	-1.35	Own calculation
Large firms	-0.12	-0.10	Own calculation
Young firms	-3.23	-4.42	Own calculation
Old firms	-0.60	-0.13	Own calculation

% change in employment from a 1-percentage-point increase in the corporate income tax rate

5.3 Simulation

We now employ our model to analyze the impact of the corporate income tax schedule on employment and welfare. We begin by varying the tax rate over a wide range of values while maintaining a single tax rate for all firms. Next, we modulate the tax based on the size and age of firms to determine the optimal design.

5.4 Change in the corporate tax rate

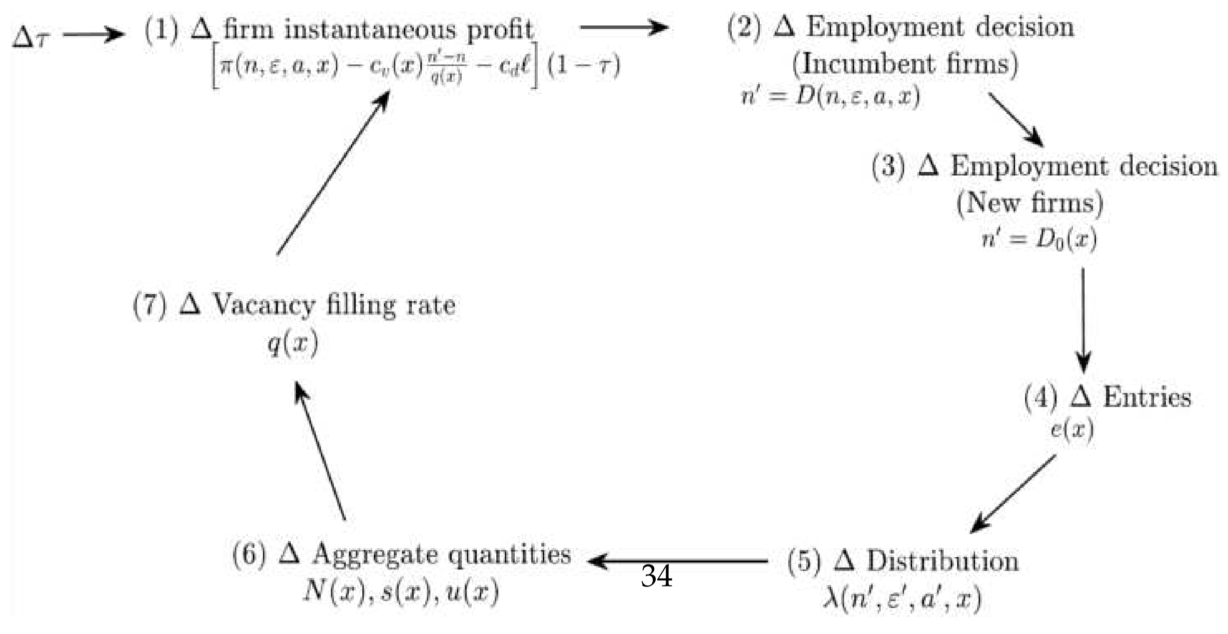
We first consider variations in the corporate tax rate ranging from a decrease of 6 percentage points (p.p.) to an increase of 6 p.p. relative to the benchmark value, which corresponds to the range of values observed in the data. Figure 8 shows the long-run level of employment as a function of the tax rate (Panel a; solid black line). We observe that the relationship is fairly linear. As expected, aggregate employment (Panel a) decreases as the corporate tax rate increases. The slope is steeper for small firms (Panel c) than for large firms (Panel d), except for a 1 p.p. decline in the tax rate.[9] This indicates that small firms are more responsive to changes in the tax rate compared to large firms. Similarly, employment is more responsive to changes in the tax rate in young firms (Panel e) compared to old firms (Panel f). When the tax rate is 6 p.p. higher than the benchmark, employment in young firms declines by around 20%, compared to 10% for old firms. Firm turnover, measured by entries and exits (Panel h), declines substantially when the tax rate increases. Worker turnover (Panel g) also shows a negative relationship with the tax rate. These results are consistent with our micro-econometric estimation. Lastly, we observe that the fiscal surplus increases when the tax rate increases, suggesting a peak for the Laffer curve on the right.

Inspecting the mechanisms. Several mechanisms can explain the aforementioned results. Figure 9 illustrates how the change in the tax rate propagates in the model. Consider an increase in the corporate tax rate. (1) It directly reduces firm profits, leading to (2) adjustments in decision rules such as job creation and job destruction. This reduction in the value of incumbent firms propagates to new firms, causing a decrease in job creation (3), and consequently fewer firm entries (4). The total number of firms decreases (5). This reduction in aggregate employment and increase in unemployment rates (6) raises the vacancy filling rate (7). These general equilibrium effects feed back into firm profits by lowering hiring costs, thereby mitigating the initial negative impacts seen in (1) and in the subsequent stages. We observe that the shocks propagates through five different channels: job creation, job destruction, firm entry, firm exit and general equilibrium effects. Following a change in the corporate income tax rate, what is the contribution of each of these channels to the employment response?

[9] This result occurs because a significant number of small firms, located close to the size limit of 500 employees, switch to an employment level of more than 500 employees.

Figure 8: Effects of a non-contingent change in corporate tax

Figure 9: Inspecting the mechanisms

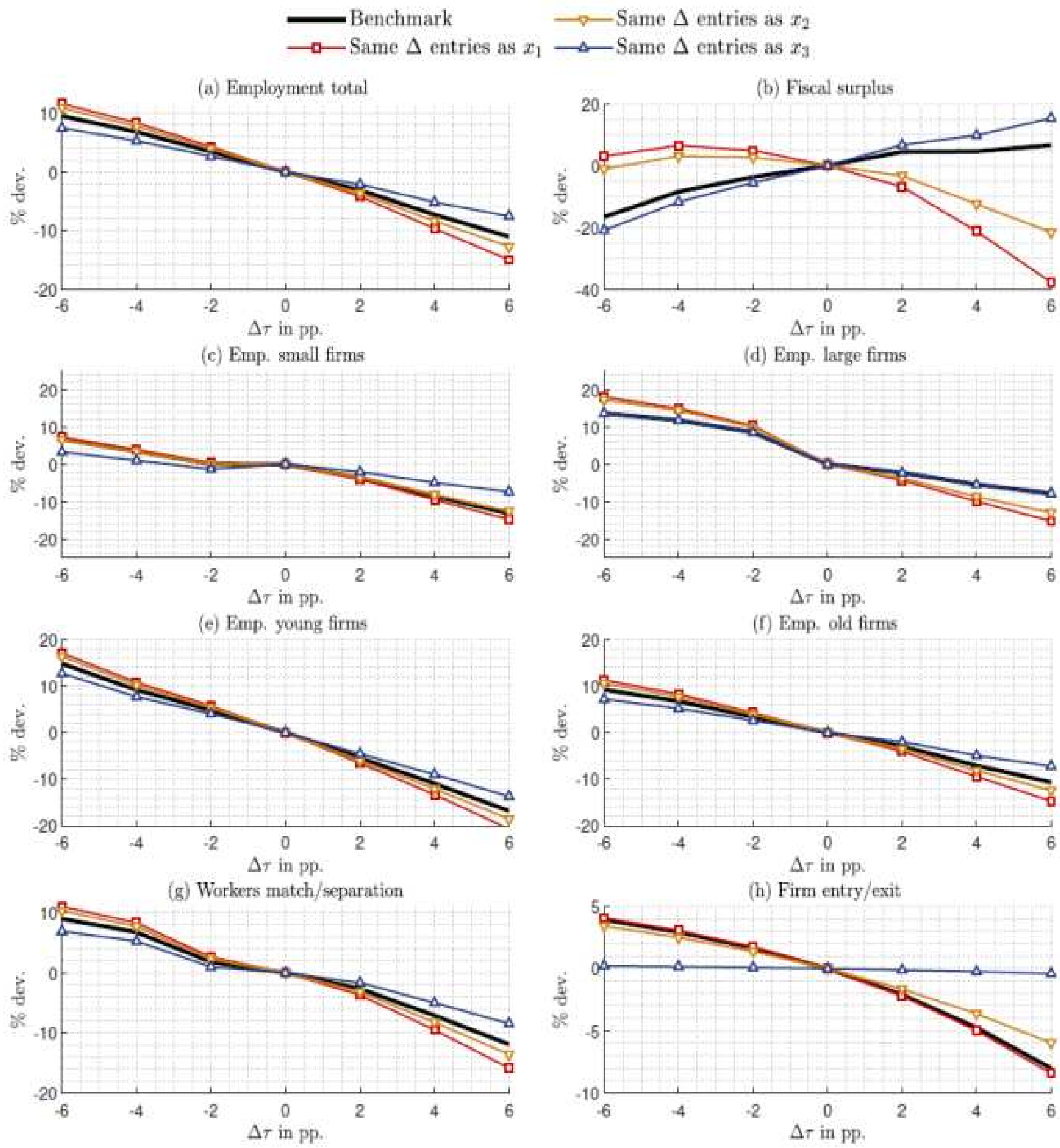


To address this question, we propose a decomposition exercise where we sequentially mute each channel to identify their relative contribution.[10] The green dashed-dotted line in Figure 8 shows that when the job creation channel is muted, the decrease in the tax causes a decline in total employment (Panel a) by around half of that found when all channels are are active (black solid line). This means that the job creation channel explains a significant part of the increase in employment resulting from corporate tax cuts. This is particularly relevant for young firms, where most of the increase in employment is accounted for by the job creation channel. The decline in employment resulting from the increase in the corporate tax rate is mainly driven by the job destruction channel, while the job creation channel plays a minor role. The entry channel also plays a significant role in explaining variation in employment for both tax cuts and tax increases, especially for small firms. Exits and the general equilibrium effect do not seem to contribute significantly to employment variations.

Lastly, we focus on the entry effects. Recall that, from Equation (65), the entry of firms reacts differently depending on their technology level. We now examine how the economy would be affected if the entry of firms were similar across sectors. Figure 10 shows the long-run value of the variables resulting from a change in the corporate tax rate, assuming the entry of firms moves identically across sectors. Since firms with lower technology are more sensitive to tax rate variations, applying the entry variation of the least productive firms to all other firms results in stronger reactions of the variables. If all firms experience the same variation in entry as the least productive firms, the employment gains from tax cuts would be larger. Conversely, if all firms experience the same variation in entry as the most productive firms, the employment gains from tax cuts would be smaller. While these results may not be surprising, they illustrate the heterogeneous impact of variation in entries on labor market outcomes.

[10] See Appendix E for details on this decomposition exercise.

Figure 10: Change in corporate tax: decomposing entry effects



5.5 State-dependent optimal taxation

The previous results show that firms do not react in the same way to a tax variation depending on their size and age. Several countries such as Canada, France, and the United Kingdom have implemented a reduced corporate tax rate for small businesses (based on the number of employees or profit). This raises several questions. Is it relevant to adjust corporate tax based on the size of firms? Should we go even further by adjusting it according to other criteria, and if so, which ones?

The policymaker's objective is to find the tax schedule that maximizes employment or welfare. However, there are two limitations. Firstly, the variables used to adjust corporate tax must be perfectly observable. Intuitively, one would want to target tax relief towards firms with the highest potential for growth and job creation. However, it is difficult for the policymaker to distinguish such firms, especially when they are young. A result emerges from our model: old firms of small size are necessarily low-productivity firms with very limited growth and job creation potential. It thus seems possible to avoid subsidizing such firms and to concentrate tax relief on firms with greater growth and job creation potential. Secondly, a modification of the corporate tax schedule can impact the fiscal surplus. On the one hand, a reduction in the corporate tax increases the number and size of businesses, which increases the tax base. On the other hand, the tax rate is lower. Tax cuts can therefore increase or reduce the fiscal surplus, depending on whether the economy is initially situated at the left or right of the peak of the Laffer curve. As shown in Figure 8, our simulations suggest that the benchmark economy is located on the left of the peak of the Laffer curve. This means that lowering the corporate tax rate would reduce the tax surplus. However, it is difficult to know *a priori* what the effect of modulating taxes according to size and age would be on the tax surplus.

We propose, in this section, to determine the optimal tax schedule (based on firm size and age) that maximizes welfare, without any fiscal constraint, and then under the constraint of no reduction in the tax surplus. To implement the optimal policy, we consider the following functional form for the tax rate:[11]

$$\tau(n,a) \quad = \quad \frac{\xi_1}{1+\xi_2 e^{-\xi_3 n-\xi_4 a}}$$

The above function is flexible enough to generate nonlinear taxation based on size and age, potentially incorporating kinks. The goal is to find the coefficients $\xi_i \in \boldsymbol{\xi}$, $i = 1, ..., 4$ that maximize welfare (see supplementary appendix for details of the algorithm). To discipline our experiment, we then impose the constraint that the tax surplus must not fall below the benchmark level.[12] The optimization program is de-

[11] We also consider alternative forms such as standard polynomial functions or logistic functions (see supplementary appendix for more details).

[12] Note that the constraint on the fiscal surplus applies *ex-post*, meaning after all adjustments in the labor market, including general equilibrium effects.

fined as follows:

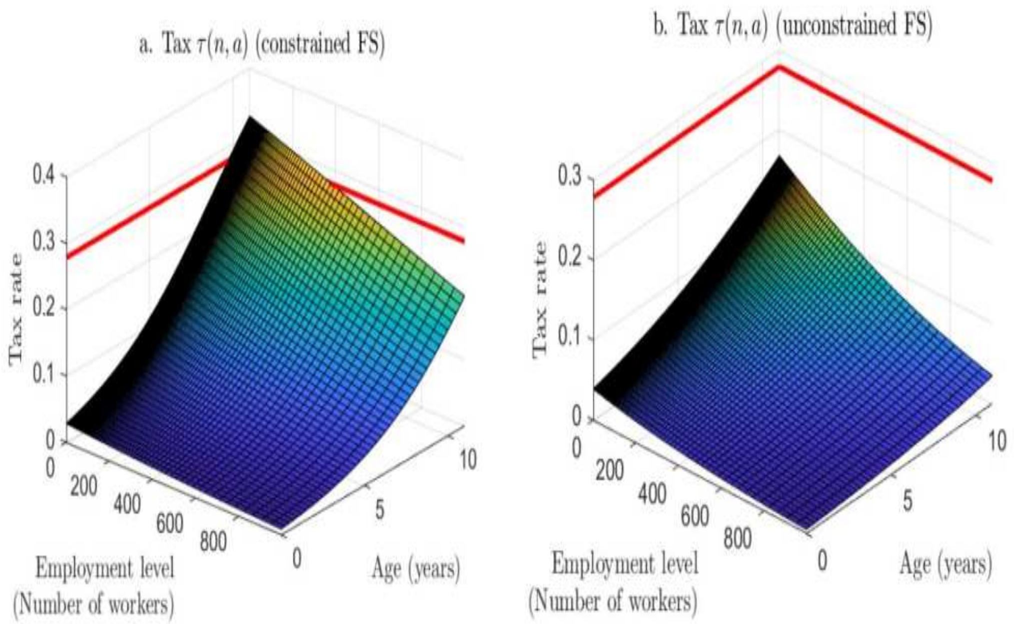
$$\xi = \arg\max \mathcal{W}(\boldsymbol{\xi})$$
$$s.t.\ \ FS(\boldsymbol{\xi}) \geq FS(\boldsymbol{\xi}_{bench.})$$

where $\boldsymbol{\xi}_{bench.}$ is the vector of coefficients in the benchmark economy, with $\xi_1 = 0.27$, and $\xi_2 = \xi_3 = \xi_4 = 0$.

We investigate the properties of the optimal tax schedule. Figure 11 shows the optimal corporate tax rate as a function of the firm size and age. In Panel a, the optimal tax schedule is calculated under the constraint that the tax surplus is not lower than that of the benchmark (constrained). Panel b relaxes the constraint on the fiscal surplus (unconstrained) to assess whether this constraint limits potential welfare gains.

It is shown that in both cases considered, the optimal tax rate exhibits a similar pattern: (i) it increases with firm age, (ii) it decreases with firm size, and (iii) it is significantly lower than the benchmark rate for most firm sizes and ages. Result (i) is logical given the high sensitivity of young firms to tax variations. Lowering the tax rate for young firms yields substantial employment gains, which translates into higher welfare. In contrast, older firms are less responsive to tax changes, thus contributing more to the fiscal surplus. Result (ii) might seem surprising at first, given that small firms are more sensitive to tax changes than large firms. However, recall that the pool of small firms contains both old, unproductive firms and young, more productive ones. By reducing taxes for young firms and increasing them for small firms, we thereby avoid subsidizing unproductive old firms. This approach supports the entry and survival of high-potential startups while encouraging the exit of less productive, older firms. Result (iii) indicates that lowering taxes can enhance employment and welfare, suggesting that a tax policy dependent on firm size and age could yield different outcomes with respect to the Laffer curve.

Figure 11: Optimal corporate tax rate



FS stands for fiscal surplus. Panels (a-b): optimal size and age-dependent tax. Red line: benchmark value.

Table 7 presents the welfare gains from implementing the optimal tax policy. Compared to the benchmark, the constrained optimal policy increases welfare by approximately 23% and employment by about 15 percentage points. The tax schedule, which decreases with firm size and increases with firm age, results in most of the employment gains being concentrated in large and young firms. Additionally, the optimal policy increases the number of firms by 3.5%. By relaxing the constraint on the fiscal surplus, we can achieve slightly higher welfare gains. However, this comes at the cost of a significant deterioration in the fiscal surplus due to the substantial tax reductions. Note that the optimal policy results in a greater increase in welfare compared to a laissez-faire economy ($\tau = 0$). This indicates that the tax helps to correct some of the externalities present in the economy. While it is challenging to pinpoint the specific role of each externality,[13] we can still identify three main sources. The first source is matching friction. Firms do not internalize the impact of their hiring decisions on the vacancy-filling rate (negative intra-group externality) and the job-finding rate (positive inter-group externality). In some cases, the wage negotiation process can

[13]Solving for the Pareto allocation and comparing it to the decentralized economy to isolate each externality is extremely difficult in this environment due to the high level of heterogeneity.

help bring the decentralized equilibrium closer to the social optimum. However, given the wage-setting mechanism used in our model, there is no reason to assume that the decentralized equilibrium coincides with the social optimum. The second source is entry friction. These frictions prevent firms from entering the market despite profit opportunities, thereby reducing the number of entries and the total number of firms in equilibrium. The last source is the distortive effect of a corporate tax rate proportional to firm profits. Setting the tax rate to zero (laissez-faire) can improve welfare but does not address the remaining externalities. Implementing the optimal tax rate further increases welfare gains. By lowering the tax rate for young firms, the optimal policy mitigates, to some extent, the entry friction.

Table 7: OPTIMAL POLICY I

Variable	Benchmark	Laissez-faire	Optimal	
			Constrained	Unconstrained
Welfare	100.0	120.9	122.7	125.8
Fiscal surplus/output	0.00	-10.54	0.14	-6.50
Employment	71.3	23.6	15.5	23.3
Share employment (small)	59.6	-0.9	-9.7	-5.5
Share employment (young)	5.8	1.2	1.4	1.0
Number of firms	100.0	6.7	3.5	6.2

"Constrained" refers to the scenario where the fiscal surplus constraint is applied ($FS(\Psi) \geq FS(\Psi_{bench.})$). "Unconstrained" relaxes this assumption. "Benchmark" column displays the variables in levels. Columns for optimal policies show the variations relative to the benchmark: percentage deviation for welfare (with a base of 100) and the number of firms, and percentage points for other variables.

The results raise one final question: Do the welfare gains primarily come from the tax dependence on firm size, age, or the combination of both? To address this question, we compare the economy with an optimal tax that depends on both size and age to an economy where the optimal tax is determined solely based on size ($\xi_4 = 0$), and then solely based on age ($\xi_3 = 0$). When we impose the constraint on the fiscal surplus (upper part), the welfare gains are significantly smaller when the optimal tax is based only on size or age. This indicates that the interaction between size and age is crucial for achieving substantial welfare and employment gains. When the constraint is relaxed (lower part), we observe that the size-dependent optimal tax allows us to achieve the maximum level of welfare (around +25%), which is not attainable with the age-dependent optimal tax (+6%). Figure 12 shows that the size-dependent optimal tax decreases with firm size, while the age-dependent optimal tax increases with firm age, irrespective of the fiscal surplus constraint. Overall, these findings suggest that a corporate tax policy based on both firm size and age is most effective.

Table 8: OPTIMAL POLICY II

	Constrained			Unconstrained		
Variable	**Size and age**	**Size only**	**Age only**	**Size and age**	**Size only**	**Age only**
Welfare	122.7	108.5	101.6	125.8	125.8	106.2
Fiscal surplus/output	0.14	0.11	0.01	-6.50	-6.34	-1.70
Employment	15.5	5.9	1.9	23.3	23.1	6.8
Share employment (small)	-9.7	-5.7	1.0	-5.5	-5.9	-0.9
Share employment (young)	1.4	-0.2	1.5	1.0	0.4	1.6
Number of firms	3.5	0.9	2.1	6.2	5.9	4.6

"Constrained" refers to the scenario where the fiscal surplus constraint is applied ($FS(\Psi) \geq FS(\Psi_{bench.})$). "Unconstrained" relaxes this assumption. "Benchmark" column displays the variables in levels. Columns for optimal policies show the variations relative to the benchmark: percentage deviation for welfare (with a base of 100) and the number of firms, and percentage points for other variables.

Figure 12: Optimal corporate tax rate

FS stands for fiscal surplus. Panels (a-b): size-dependent optimal tax. Panels (c-d): age-dependent optimal tax.

6 Conclusion

What is the effect of corporate income tax on employment, and does this effect vary according to firm age and size? Answering this question is empirically challenging. In this paper, we identify the effect of corporate tax on employment by comparing contiguous counties in neighboring states with different corporate tax rates. We show that corporate tax significantly and negatively affects employment in small (young) firms, while there is no significant effect observed for large (old) firms. When comparing the effect on small vs young firms, our estimates suggest that a one percentage point increase in state corporate income tax rate reduces employment in small (young) firms by 1.35% (4.42%), suggesting that tax reductions implemented in some countries could be more effective if targeted to young firms (instead of small firms). We then develop a large-firm model with different cohorts of firms to analyze firm dynamics over their life cycle under various tax schedules. Our simulations demonstrate that it is possible to substantially increase both employment and welfare while maintaining the fiscal surplus by adjusting corporate tax rates based on both firm size and age. These findings provide new insights into firm responses to taxation and are crucial in helping policymakers design effective tax policies.

References

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., and Kerr, W. (2018). Innovation, reallocation, and growth. *American Economic Review*, 108(11):3450–91.

Adelman, M. A. (1957). The corporate income tax in the long run. *Journal of Political Economy*, 65(2):151–157.

Albertini, J., Poirier, A., and Terriau, A. (2020). The impact of EITC on education, labor market trajectories, and inequalities. Working Papers halshs-03082382, HAL.

Auerbach, A. J., Aaron, H. J., and Hall, R. E. (1983). Corporate taxation in the United States. *Brookings Papers on Economic Activity*, 1983(2):451–513.

Bagger, J., Fontaine, F., Postel-Vinay, F., and Robin, J.-M. (2014). Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics. *American Economic Review*, 104(6):1551–96.

Blundell, R., Costa Dias, M., Meghir, C., and Shaw, J. (2016). Female labor supply, human capital, and welfare reform. *Econometrica*, 84(5):1705–1753.

Cameron, A. C. and Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372.

Cebula, R. J. and Alexander, G. M. (2006). Determinants of net interstate migration, 2000-2004. *Journal of Regional Analysis and Policy*, 36(1100-2016-89796).

Davis, S. J., Haltiwanger, J., and Schuh, S. (1996). Small business and job creation: Dissecting the myth and reassessing the facts. *Small Business Economics*, 8(4):297–315.

den Haan, W. J., Ramey, G., and Watson, J. (2000). Job destruction and propagation of shocks. *American Economic Review*, 90(3):482–498.

Djankov, S., Ganser, T., McLiesh, C., Ramalho, R., and Shleifer, A. (2010). The effect of corporate taxes on investment and entrepreneurship. *American Economic Journal: Macroeconomics*, 2(3):31–64.

Dube, A., Lester, T. W., and Reich, M. (2010). Minimum wage effects across state borders: Estimates using contiguous counties. *The Review of Economics and Statistics*, 92(4):945–964.

Due, J. F. (1961). Studies of state-local tax influences on location of industry. *National Tax Journal*, 14(2):163–173.

Edgerton, J. (2010). Investment incentives and corporate tax asymmetries. *Journal of Public Economics*, 94(11-12):936–952.

Elsby, M. W. and Michaels, R. (2013). Marginal jobs, heterogeneous firms, and unemployment flows. *American Economic Journal: Macroeconomics*, 5(1):1–48.

Giroud, X. and Rauh, J. (2019). State taxation and the reallocation of business activity: Evidence from establishment-level data. *Journal of Political Economy*, 127(3):1262–1316.

Hagedorn, M., Manovskii, I., and Mitman, K. (2015). The impact of unemployment benefit extensions on employment: the 2014 employment miracle? Technical report, National Bureau of Economic Research.

Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2013). Who creates jobs? Small versus large versus young. *Review of Economics and Statistics*, 95(2):347–361.

Harden, J. W. and Hoyt, W. H. (2003). Do states choose their mix of taxes to minimize employment losses? *National Tax Journal*, 56(1):7–26.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, 60(5):1127–1150.

Huang, R. R. (2008). Evaluating the real effect of bank branching deregulation: Comparing contiguous counties across us state borders. *Journal of Financial Economics*, 87(3):678–705.

Jacob, M. (2021). Real effects of corporate taxation: A review. *European Accounting Review*, pages 1–28.

Klette, T. J. and Kortum, S. (2004). Innovating firms and aggregate innovation. *Journal of Political Economy*, 112(5):986–1018.

Ljungqvist, A. and Smolyansky, M. (2014). To cut or not to cut? On the impact of corporate taxes on employment and income. Technical report, National Bureau of Economic Research.

Menzio, G., Telyukova, I. A., and Visschers, L. (2016). Directed search over the life cycle. *Review of Economic Dynamics*, 19:38–62.

Mertens, K. and Ravn, M. O. (2013). The dynamic effects of personal and corporate income tax changes in the united states. *American Economic Review*, 103(4):1212–47.

Moll, B. (2017). Hopenhayn model in continuous time. *Mimeo*.

Ohrn, E. (2018). The effect of corporate taxation on investment and financial policy: Evidence from the dpad. *American Economic Journal: Economic Policy*, 10(2):272–301.

Saint-Paul, G. (2002). Employment protection, international specialization, and innovation. *European Economic Review*, 46(2):375–395.

Sandmo, A. (1974). Investment incentives and the corporate income tax. *Journal of Political Economy*, 82(2, Part 1):287–302.

Sedláček, P. (2020). Lost generations of firms and aggregate labor market dynamics. *Journal of Monetary Economics*, 111:16–31.

Sedláček, P. and Sterk, V. (2017). The growth potential of startups over the business cycle. *American Economic Review*, 107(10):3182–3210.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, 64(5):416–424.

Weinzierl, M. (2011). The surprising power of age-dependent taxes. *The Review of Economic Studies*, 78(4):1490–1518.

Appendix

A Data description

In our study, we make use of three datasets:

i) The *Business Dynamics Statistics (BDS)*, provided by the U.S. Census Bureau

ii) The *Book of the States*, published by the Council of State Governments

iii) The intercensal estimates of the *Population Estimates Program (PEP)*, provided by the U.S. Census Bureau

The following appendix provides a detailed description of each of these datasets.

A.1 Sample

Our empirical strategy requires restricting the sample to states that have a corporate tax system similar to the federal system and share a border with at least one other state. Consequently,

i) We exclude Michigan, Ohio, and Texas, which have a corporate tax system that is not comparable to the federal system.

ii) We exclude Alaska and Hawaii, which do not share a border with any other state.

Our final sample contains 46 states, including the District of Columbia.

A.2 Business Dynamics Statistics (BDS)

A.2.1 Description

The *Business Dynamics Statistics* (BDS) is a longitudinal business database, provided by the U.S. Census Bureau, that tracks firms over time and provides annual measures of business dynamics at the county level. Interestingly, the BDS is one of the few U.S. databases that provide information about firm size (allowing us to distinguish between small and large firms) and firm age (allowing us to distinguish between young and old firms). In our study, we use the BDS for the period from 2002 to 2009 to compile data on firm dynamics at the county level.

A.2.2 Variables

- *Year*: This variable corresponds to the observation year.
- *State*: This variable identifies each state using the state-level FIPS codes.
- *County*: This variable identifies each county using the county-level FIPS codes.
- *Pair*: This variable identifies each contiguous border-county pair using the *Boundary files* and the *Adjacency files* described below.

  Firm size: This variable identifies the firm size. It is a categorical variable (1-4; 5-9; 10-19; 20-99; 100-499; 500-999; 1000-2499; 2500-4999; 5000-9999; 10000+), defined as the average of the firm's employment in year $t-1$ and year t. As noted by Davis et al. (1996) and Haltiwanger et al. (2013), this measure of firm size is more robust to regression-to-the-mean effects. We refer readers to their paper for a discussion of this issue. The Small Business Administration (SBA) defines small firms as those with fewer than 500 employees. We use this cutoff to distinguish between small and large firms. In Section 2.6, we split our sample into two sub-samples: the "large firms" sample, which contains all firms with 500 employees or more, and the "small firms" sample, which contains all firms with fewer than 500 employees.

- *Firm age*: This variable indicates the firm age. It is a categorical variable (0; 1; 2; 3; 4; 5; 6-10; 11-15; 16-20; 21-25; 26+). Note that firms founded before 1977 are of unknown age (left-censored) and are therefore excluded from our analysis. In Section 2.6, we split our sample into two sub-samples: the "old firms" sample, which contains all firms aged 6 years or more, and the "young firms" sample, which contains all firms aged 5 years or less.

A.2.3 Supplementary material

Boundary files: We use IPUMS GIS boundary files, available at `https://www.nhgis.org/`, to define county and state boundaries.

Adjacency files: We use NBER adjacency files, available at `https://data.nber.org/data/county-adjacency.html/`, to compute county and state contiguity matrices.

A.3 Book of the States

A.3.1 Description

The *Book of the States*, published by the Council of State Governments, provides accurate and comparable information on state policies over time, including taxation. In our study, we use the Book of the States for the period from 2002 to 2009 to compile data on corporate income, personal income, and sales tax rates at the state level.

A.3.2 Variables

- *Year*: This variable corresponds to the observation year.
- *State*: This variable identifies each state based on the state-level FIPS codes.
- *Corporate Income Tax*: This variable corresponds to the top statutory marginal state corporate income tax rate. This approach is appropriate for two reasons:
  i) Most states have a flat rate. In 2009, only 12 of the 46 states of our sample did not have a flat rate.
  ii) For states without a flat rate, the highest corporate income tax bracket is generally very low, affecting a small number of corporations. In 2009, 8 states had tax brackets ranging from \$0 to \$100 000. The highest tax bracket exceeded \$100 000 in only 4 states.
- *Individual Income Tax*: This variable corresponds to the top statutory marginal state individual income tax rate.
- *Sales Tax*: This variable corresponds to the tax rate for general sales and gross receipts.

A.4 Population Estimates Program (PEP)

A.4.1 Description

The *PEP*, provided by the U.S. Census Bureau, produces county-level estimates of the population by age, sex, and race. In our study, we use the PEP for the period from 2002 to 2009 to compile data on population dynamics at the county level.

A.4.2 Variables

- *Year*: This variable corresponds to the observation year.
- *State*: This variable identifies each state using the state-level FIPS codes.
- *County*: This variable identifies each county using the county-level FIPS codes.
- *Total population*: This variable corresponds to the total population at the county level.
- *Total male population*: This variable corresponds to the total male population at the county level.
- *Total female population*: This variable corresponds to the total female population at the county level.
- *Total Hispanic white male population*: This variable corresponds to the total Hispanic white male population at the county level.
- *Total Hispanic white female population*: This variable corresponds to the total Hispanic white female population at the county level.
- *Total non-Hispanic white male population*: This variable corresponds to the total non-Hispanic white male population at the county level.
- *Total non-Hispanic white female population*: This variable corresponds to the total non-Hispanic white female population at the county level.

B Number of firms and employment by firm size and age

a) Number of firms	Firm size									
Firm age	1-4	5-9	10-19	20-99	100-499	500-999	1,000-2,499	2,500-4,999	5,000-9,999	10,000+	All
0	345,891	25,618	10,975	6,031	445	36	12	0	0	0	389,008
1	252,373	49,536	25,705	16,366	1,352	77	26	10	5	7	345,457
2	213,284	50,997	26,799	17,365	1,474	92	37	15	3	3	310,069
3	203,948	56,882	29,918	19,464	1,798	97	38	9	0	5	312,159
4	163,872	48,169	26,042	17,859	1,848	121	54	17	5	9	257,996
5	140,171	42,260	23,501	16,895	1,821	118	63	15	4	7	224,855
6-10	498,241	164,916	96,646	74,202	10,177	854	447	136	55	40	845,714
11-15	361,787	136,294	83,406	66,674	9,749	952	517	203	93	70	659,745
16-20	241,820	103,667	64,596	53,452	8,330	808	447	169	83	74	473,446
21-25	199,387	93,020	57,239	48,993	8,531	1,033	677	238	131	112	409,361
26+	155,961	80,197	52,154	48,221	9,486	1,079	667	255	133	102	348,255
All	2,776,735	851,556	496,981	385,522	55,011	5,267	2,985	1,067	512	429	4,576,065

b) Employment	Firm size									
Firm age	1-4	5-9	10-19	20-99	100-499	500-999	1,000-2,499	2,500-4,999	5,000-9,999	10,000+	All
0	846,811	339,442	292,877	430,738	154,158	44,333	32,617	0	0	0	2,140,976
1	525,788	350,695	364,478	624,701	233,458	53,465	26,020	23,042	87	356	2,202,090
2	442,117	342,774	361,775	635,496	251,113	60,040	48,429	22,475	24,583	500	2,189,302
3	430,385	379,643	400,799	710,121	318,985	59,046	51,024	17,579	0	8,556	2,376,138
4	346,350	319,866	347,245	658,463	323,653	74,288	74,543	37,134	23,343	33,678	2,238,563
5	295,980	279,618	312,565	631,906	330,156	75,291	80,154	37,128	300	13,898	2,056,996
6-10	1,065,129	1,091,116	1,286,162	2,806,383	1,837,350	503,463	562,744	288,994	258,236	222,168	9,921,745
11-15	792,265	897,626	1,106,754	2,523,602	1,792,749	573,590	649,890	466,903	482,700	801,131	10,087,210
16-20	544,736	685,142	858,780	2,022,540	1,531,894	492,493	590,238	427,546	419,170	896,505	8,469,044
21-25	462,378	613,846	759,431	1,876,847	1,574,100	609,839	891,015	696,663	690,924	2,211,122	10,386,165
26+	364,125	532,575	694,885	1,881,840	1,774,223	638,709	869,370	727,279	651,418	1,927,575	10,061,999
All	6,116,064	5,832,343	6,785,751	14,802,637	10,121,839	3,184,557	3,876,044	2,744,743	2,550,761	6,115,489	62,130,228

Source: BDS, 2009

C Descriptive statistics by state

State	Total population	Total employment	Number of firms	Corporate income tax	Individual income tax	Sales tax
Alabama	4,549,668	1,650,260	81,784	6.50	5.00	4.00
Arizona	5,891,964	2,139,820	93,851	6.97	4.86	5.60
Arkansas	2,770,691	999,232	53,009	6.50	6.94	5.67
California	35,594,036	13,147,319	646,603	8.84	9.30	7.38
Colorado	4,659,636	1,968,251	115,849	4.63	4.63	2.90
Connecticut	3,471,655	1,525,743	71,526	7.50	4.88	6.00
Delaware	834,291	384,756	18,537	8.70	5.95	0.00
District of Columbia	582,214	427,033	15,035	9.98	8.70	5.75
Florida	17,495,596	6,732,030	373,443	5.50	0.00	6.00
Georgia	9,039,147	3,477,358	171,472	6.00	6.00	4.00
Idaho	1,417,022	501,179	34,762	7.60	7.80	5.63
Illinois	12,665,522	5,254,133	240,687	7.30	3.00	6.25
Indiana	6,244,933	2,574,855	118,112	8.43	3.40	5.88
Iowa	2,952,424	1,263,803	65,640	12.00	8.98	5.13
Kansas	2,742,151	1,132,615	60,902	4.39	6.45	5.25
Kentucky	4,171,624	1,498,500	74,198	7.50	6.00	6.00
Louisiana	4,432,074	1,589,842	82,923	8.00	6.00	4.00
Maine	1,306,710	495,645	32,022	8.93	8.50	5.00
Maryland	5,542,741	2,152,010	105,914	7.32	4.94	5.25
Massachusetts	6,464,741	3,011,783	133,739	9.50	5.19	5.00
Minnesota	5,100,520	2,415,064	112,964	9.80	7.85	6.55
Mississippi	2,890,605	925,984	50,260	5.00	5.00	7.00
Missouri	5,791,563	2,394,314	119,252	6.25	6.00	4.22
Montana	933,084	326,665	28,794	6.75	8.44	0.00
Nebraska	1,747,864	768,075	41,243	7.81	6.82	5.38
Nevada	2,364,014	1,059,234	41,840	0.00	0.00	6.50
New Hampshire	1,294,400	559,829	30,812	8.50	0.00	0.00
New Jersey	8,596,699	3,588,340	189,722	9.00	7.67	6.38
New Mexico	1,906,875	595,547	35,715	7.60	6.25	5.00
New York	19,294,738	7,349,757	399,334	7.51	7.17	4.06
North Carolina	8,664,430	3,420,518	172,021	6.90	8.09	4.47
North Dakota	636,344	273,455	17,275	8.19	5.54	5.00
Oklahoma	3,540,742	1,226,958	71,075	6.00	6.24	4.50
Oregon	3,615,318	1,394,739	84,781	6.60	9.00	0.00
Pennsylvania	12,418,741	5,079,604	234,398	9.99	2.97	6.00
Rhode Island	1,062,587	428,737	23,504	9.00	8.85	7.00
South Carolina	4,254,691	1,595,042	82,203	5.00	7.00	5.25
South Dakota	765,793	315,982	20,746	0.00	0.00	4.00
Tennessee	5,978,819	2,375,406	104,945	6.44	0.00	6.88
Utah	2,489,885	980,941	49,726	5.00	6.50	4.73
Vermont	617,764	257,160	17,939	9.28	9.41	5.75
Virginia	7,499,162	3,050,295	152,399	6.00	5.75	4.69
Washington	6,250,763	2,329,711	134,913	0.00	0.00	6.50
West Virginia	1,805,095	570,983	33,019	8.88	6.50	6.00
Wisconsin	5,523,253	2,427,201	113,974	7.90	6.75	5.00
Wyoming	508,463	191,907	16,383	0.00	0.00	4.00

Source: BDS, Book of the States, and PEP (2002-2009)

D A simple two-periods model

To investigate the mechanisms working in the model presented in Section 4, we constructed a simple two-period model of labor demand. The general structure of the simplified model is the same as the main model; however, we consider a partial equilibrium environment (wages are proportional to the firm's productivity), and the firm faces a quadratic adjustment cost. There is also an operating cost that is linear in employment and a corporate tax. We make the following assumptions:

A ion 1 *$x = x_1, x_2$ is the technology level, which is fixed and constant over the life cycle. x_1 stands for low technology, and x_2 stands for high technology. We denote by $y(x)$ the production level associated with a technology level x. It holds that $y(x_2) > y(x_1)$*

A ion 2 *The wage rate $w(x)$ depends on the technology level. We assume it is proportional to productivity, meaning: $w(x) = \alpha y(x)$. It is constant over the life cycle.*

A ion 3 *There are two life cycle periods, $a = a_1, a_2$, which define the young and old periods. After age a_2, all firms exit.*

We also make assumptions concerning the operating cost and the marginal gains.

A ion 4 *There is an operating cost that depends on age and the technology level. The operating cost when young is higher, i.e., $c_0(a_1, x) > c_0(a_2, x), \forall x$.*

To make easier the obtention of analytical results, we make two simplifying assumptions.

A ion 5 *The decline in operating cost with age is proportional to the initial operating cost, that is:*

$$c_o(a_2, x) = \phi c_o(a_1, x),$$

with $\phi \in [0, 1]$.

A ion *The operating cost when young is proportional to productivity, that is:*

$$c_o(a_1, x) = \gamma y(x),$$

with $\gamma > 0$.

A ion 7 *The production level, wage, operating cost, and corporate taxes* (τ_1, τ_2) *are such that the marginal gains from employment are positive, that is:*

$$\begin{aligned}
\pi(a_1, x) &= (1-\tau_1)(y(x) - w(x)) - c_0(a_1, x) = y(x)[(1-\tau_1)(1-\alpha) - \gamma] > 0, \\
\pi(a_2, x) &= (1-\tau_2)(y(x) - w(x)) - c_0(a_2, x) = y(x)[(1-\tau_1)(1-\alpha) - \gamma\phi] > 0.
\end{aligned}$$

D.1 The firm program

The firm maximizes its intertemporal profit subject to the evolution of employment. It holds that:

$$\begin{aligned}
\max_{h_1, n_1, h_2, n_2} \quad & (1-\tau_1)\left(y(x)n_1 - w(x)n_1 - \frac{b}{2}h_1^2\right) - c_0(a_1, x)n_1 \\
& + (1-\delta)\beta\left\{(1-\tau_2)\left(y(x)n_2 - w(x)n_2 - \frac{b}{2}h_2^2\right) - c_0(a_2, x)n_2\right\} \\
\text{s.t.} \quad & \begin{cases} -n_1 + h_1 = 0 & (q_1) \\ -n_2 + (1-\omega)n_1 + h_2 = 0 & (q_2) \end{cases}
\end{aligned}$$

Let $b > 0$ be the adjustment cost parameter, $\beta \in]0,1[$ the discount rate, $\delta \in [0,1[$ an exogenous firm destruction rate, and $\omega \in]0,1[$ the quit rate. n_1, h_1, n_2, h_2 are the employment stocks and the hiring flows at age a_1 and a_2, respectively. We suppose that $n_0 = 0$.

We consider two different tax rates that depend on age. This allows us to analyze the effect of an instantaneous tax variation. However, we will further consider a variation around a constant tax profile, $\tau = \tau_1 = \tau_2$

D.1.1 Optimality conditions

We now derive the optimality conditions. The Lagrangian is written as:

$$\begin{aligned}
\mathcal{L} &= (1-\tau_1)\left(y(x)n_1 - w(x)n_1 - \frac{b}{2}h_1^2\right) - c_0(a_1, x)n_1 \\
&+ (1-\delta)\beta\left\{(1-\tau_2)\left(y(x)n_2 - w(x)n_2 - \frac{b}{2}h_2^2\right) - c_0(a_2, x)n_2\right\} \\
&+ q_1(-n_1 + h_1) + q_2(-n_2 + (1-\omega)n_1 + h_2).
\end{aligned}$$

We derive the following set of optimality conditions:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial h_1} &= -(1-\tau_1) b h_1 + q_1 = 0, \\
\frac{\partial \mathcal{L}}{\partial n_1} &= (1-\tau_1)(y(x) - w(x)) - c_0(a_1, x) - q_1 + q_2(1-\omega) = 0, \\
\frac{\partial \mathcal{L}}{\partial h_2} &= -\beta(1-\delta)(1-\tau_2) b h_2 + q_2 = 0, \\
\frac{\partial \mathcal{L}}{\partial n_2} &= \beta(1-\delta)\left\{(1-\tau_2)(y(x) - w(x)) - c_0(a_2, x)\right\} - q_2 = 0.
\end{aligned}
$$

Knowing that the firm's decisions depend on the state (a_t, x) (which represents the age of the firm and the technology level), the solution of the above system (the firm's optimal decisions) is written as:

$$
\begin{aligned}
h(a_1, x) &= \frac{1}{(1-\tau_1)b}\left\{(1-\tau_1)(y(x) - w(x)) - c_0(a_1, x)\right. \\
&+ \left.\overline{\beta}\left[(1-\tau_2)(y(x) - w(x)) - c_0(a_2, x)\right]\right\}, \qquad (32) \\
h(a_2, x) &= \frac{1}{(1-\tau_2)b}\left\{(1-\tau_2)(y(x) - w(x)) - c_0(a_2, x)\right\}, \qquad (33) \\
n(a_1, x) &= h(a_1, x), \qquad (34) \\
n(a_2, x) &= (1-\omega) n(a_1, x) + h(a_2, x), \qquad (35)
\end{aligned}
$$

with $\overline{\beta} = \beta(1-\omega)(1-\delta)$.

Finally, the implicit prices are written as:

$$
\begin{aligned}
q(a_1, x) &= (1-\tau_1)(y(x) - w(x)) - c_0(a_1, x) + (1-\omega) q(a_2, x) \qquad (36) \\
q(a_2, x) &= \beta(1-\delta)\left\{(1-\tau_2)(y(x) - w(x)) - c_0(a_2, x)\right\} \qquad (37)
\end{aligned}
$$

R ark 1 *It follows from equations (32) and (33) that, in the absence of operating costs, firm hiring (and thus firm employment) is independent of the tax rate.*

D.1.2 Analysis of the young period hiring

Consider the optimal hiring decision during the young period. The effect of a variation in the young-age tax τ_1 is complex, with several effects at work. Let us write:

$$h(a_1, x) = h^1(a_1, x) + h^2(a_1, x),$$

with

$$h^1(a_1, x) = \frac{1}{(1-\tau_1)b} \{(1-\tau_1)(y(x) - w(x)) - c_0(a_1, x)\}, \quad (38)$$

$$h^2(a_1, x) = \frac{\overline{\beta}}{(1-\tau_1)b} \{[(1-\tau_2)(y(x) - w(x)) - c_0(a_2, x)]\}. \quad (39)$$

Hiring during the young period has two terms because it has an intertemporal dimension. Indeed, the hiring decision depends on the discounted marginal gain of an additional worker:

1. The first term, $h^1(a_1, x)$, is the instantaneous component of hiring and depends on the period 1 marginal gain;

2. The second term, $h^2(a_1, x)$, is the intertemporal component of hiring, which depends on the period 2 marginal gain. Hiring today generates gains in the future.

We now discuss the impact of an increase in the tax rate τ_1. To begin, consider the first component (equation (38)). An increase in τ_1 has an ambiguous effect. First, it reduces the marginal gain (the numerator of expression (38)) and thus reduces hiring. Second, it acts as a reduction in the adjustment cost (the denominator of expression (38)) and thus increases hiring. This occurs because the adjustment cost reduces the base on which the tax is levied. The overall effect seems ambiguous. However, it is easy to show that:

$$\frac{\partial h^1(a_1, x)}{\partial \tau_1} = -\frac{c_0(a_1, x)}{(1-\tau_1)^2 b} < 0.$$

Now consider the second component (expression (39)). We obtain:

$$\frac{\partial h^2(a_1, x)}{\partial \tau_1} = \frac{\overline{\beta}}{(1-\tau_1)^2 b} \{[(1-\tau_2)(y(x) - w(x)) - c_0(a_2, x)]\} > 0.$$

An increase in the tax rate τ_1, via the intertemporal effect, induces, other things being equal, an increase in hiring. This effect comes from the deductibility of the adjustment cost. The firm is incentivized to hire with consideration for the second period. We also observe that an increase in τ_2 reduces the marginal gain from employment and thus has a negative effect on hiring. That is:

$$\frac{\partial h^2(a_1, x)}{\partial \tau_2} = -\frac{\overline{\beta}}{(1-\tau_1)b}\{y(x) - w\} < 0.$$

Finally, the effect of a permanent increase in τ on hiring can be expressed as

$$\frac{\partial h(a_1, x)}{\partial \tau} = -\frac{c_0(a_1, x)}{(1-\tau)^2 b} - \overline{\beta}\frac{c_0(a_2, x)}{(1-\tau)^2 b} < 0$$

It leads to a reduction in hiring. This occurs if the operating costs are strictly positive and are not tax-deductible.

D.1.3 Analysis of the old period hiring

Consider the optimal hiring decision during the old period (equation (33)). There is only an instantaneous component. Since the firm is at the end of its life cycle, there is no intertemporal effect considering the future of the firm. The effect of a variation in the old-age tax τ_2 on the hiring flow is straightforward and is expressed as:

$$\frac{\partial h(a_2, x)}{\partial \tau_2} = -\frac{c_0(a_2, x)}{(1-\tau_2)^2 b} < 0.$$

It results in a reduction in hiring for old firms. As for young firms, this result stems from the fact that the operating cost is not tax-deductible.

D.2 Employment level over the life cycle

We now provide some results and analysis of the model solution. We define the size of a firm by the number of workers employed in the firm.

D.2.1 Young vs old firms

R l 1 *For any given level x, and if $\tau_1 = \tau_2 = \tau$, young firms are smaller than old firms, that is, $n(a_1, x) < n(a_2, x)$.*

Proof Suppose that $\tau_1 = \tau_2 = \tau$. Equations (34) and (35) are expressed as follows:

$$n(a_1, x) = \frac{1}{(1-\tau)b}\left(\pi(a_1, x) + \overline{\beta}\pi(a_2, x)\right), \quad (40)$$

$$n(a_2, x) = (1-\omega)n(a_1, x) + \frac{1}{(1-\tau)b}\pi(a_2, x). \quad (41)$$

We deduce:

$$\begin{aligned} n(a_2, x) - n(a_1, x) &= \frac{1}{(1-\tau)b}\left\{\pi(a_2, x) - \omega[\pi(a_1, x) + \overline{\beta}\pi(a_2, x)]\right\}, \\ &= \frac{1}{(1-\tau)b}\left\{(1-\omega\overline{\beta})\pi(a_2, x) - \omega\pi(a_1, x)\right\}. \quad (42) \end{aligned}$$

Given that $c_0(a_1, x) > c_0(a_2, x)$, it follows that $\pi(a_2, x) > \pi(a_1, x)$.

It follows that the RHS of equation (42) satisfies:

$$\begin{aligned} (1-\omega\overline{\beta})\pi(a_2, x) - \omega\pi(a_1, x) &> (1-\omega\overline{\beta})\pi(a_1, x) - \omega\pi(a_1, x) \\ &= \left(1-\omega\overline{\beta}-\omega\right)\pi(a_1, x) \\ &= \left(1-\omega\beta(1-\delta)(1-\omega)-\omega\right)\pi(a_1, x) \\ &= (1-\omega)(1-\beta\omega(1-\delta))\pi(a_1, x) > 0. \end{aligned}$$

We conclude that $n(a_2, x) - n(a_1, x) > 0$. □

This result would also hold if $c_0(a_1, x) = c_0(a_2, x)$. However, the reduction in the operating cost in the old period amplifies the difference $n(a_2, x) - n(a_1, x) > 0$.

Corollary 1 *Suppose that the operating costs satisfy assumptions 5 and 6. If ϕ is sufficiently close to 1, then firms grow faster when young than when old.*

Proof We must show that:

$$n(a_1, x) - 0 > n(a_2, x) - n(a_1, x)$$

or equivalently,

$$\pi(a_1,x)+\overline{\beta}\pi(a_2,x) > (1-\omega\overline{\beta})\pi(a_2,x)-\omega\pi(a_1,x).$$

Using assumptions 2, 5, and 6, the above condition can be expressed as follows:

$$\begin{aligned}&(1+\omega)\left\{(1-\tau)(1-\alpha)-\gamma\right\}\\&+\left\{(1+\omega)\overline{\beta}-1\right\}\left\{(1-\tau)(1-\alpha)-\gamma\phi\right\}>0.\end{aligned} \tag{43}$$

Suppose now that $\phi = 1$. In this case, inequality (43) becomes:

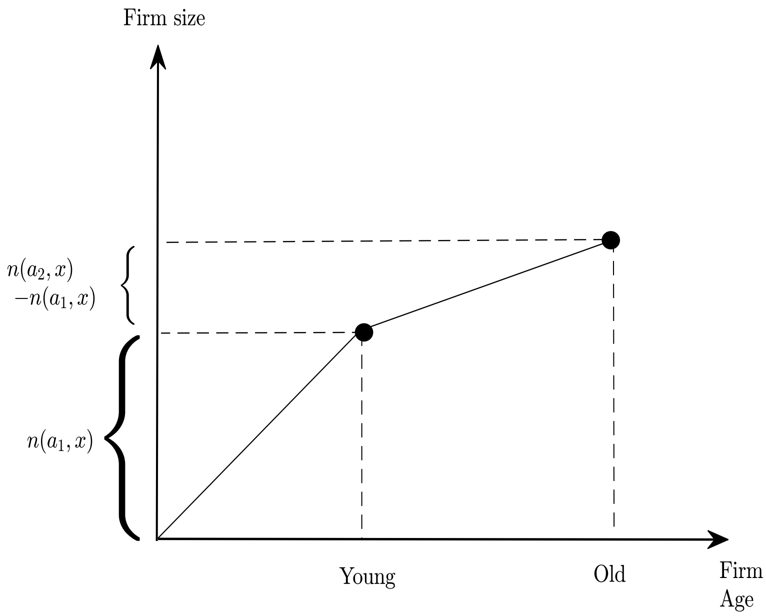
$$\left\{(1-\tau)(1-\alpha)-\gamma\right\}\left\{(1+\omega)(1+\overline{\beta})-1\right\}>0. \tag{44}$$

It is obvious that $(1+\omega)(1+\overline{\beta})-1>0$. Furthermore, assumption 4 ensures that $(1-\tau)(1-\alpha)-\gamma > 0$. We conclude that inequality (43) is satisfied if $\phi = 1$. By continuity, it is also satisfied if $\phi < 1$ and sufficiently close to 1. □

Young firms grow faster only if the operating costs differ little. On the other hand, if the operating cost is significantly lower when old, the firm will hire massively when old, and the reverse of corollary 1 will hold.

Figure 13 illustrates corollary 1. The employment profile over the life cycle is increasing and concave.

Figure 13: Firm size over the life cycle



D.2.2 Small vs large firms

We now investigate the impact of the productivity level on the size of the firm. Recall that we consider two level of productivity: $x = \{x_1, x_2\}$, with $x_2 > x_1$ and $y(x_2) > y(x_1)$. We know that 99% of young firms are small, so young firms $n(a_1, x)$ are all small $\forall x$. However, not all old firms are small or large. Our objective is to show that highly productive firms are larger when old, that is:

$$n(a_2, x_2) \;\geq\; n(a_2, x_1).$$

Note that, from result 1, we have $n(a_2, x_2) \geq n(a_1, x_2)$ and $n(a_2, x_1) \geq n(a_1, x_1)$. Consequently, showing that $n(a_2, x_2) \geq n(a_2, x_1)$ implies that $n(a_2, x_2) \geq n(a_1, x_1)$: larger firms are those with the highest productivity and are old. We have the following result:

R l 2 *Suppose that assumptions 1-7 are satisfied and that* $\tau = \tau_1 = \tau_2$. *Then,*

$$n(a_2, x_2) \;\geq\; n(a_2, x_1).$$

Proof To begin, we will show that $n(a_2, x_2) \geq n(a_2, x_1)$. Let us consider:

$$n(a_2, x_2) - n(a_2, x_1) \;=\; (1-\omega)(n(a_1, x_2) - n(a_1, x_1)) + \frac{1}{(1-\tau)b}\left(\pi(a_2, x_2) - \pi(a_2, x_1)\right),$$

which can be rearranged as follows:

$$\begin{aligned} n(a_2, x_2) - n(a_2, x_1) \;&=\; \frac{1-\omega}{(1-\tau)b}\Big(\pi(a_1, x_2) - \pi(a_1, x_1) + \overline{\beta}(\pi(a_2, x_2) - \pi(a_2, x_1))\Big) \\ &+\; \frac{1}{(1-\tau)b}\left(\pi(a_2, x_2) - \pi(a_2, x_1)\right). \end{aligned}$$

The objective is to show that $\pi(a_1, x_2) - \pi(a_1, x_1) > 0$ and $\pi(a_2, x_2) - \pi(a_2, x_1) > 0$. It is straightforward to show that:

$$\begin{aligned} \pi(a_1, x_2) - \pi(a_1, x_1) \;&=\; [(1-\tau)(1-\alpha) - \gamma](y(x_2) - y(x_1)) \\ \pi(a_2, x_2) - \pi(a_2, x_1) \;&=\; [(1-\tau)(1-\alpha) - \gamma\phi](y(x_2) - y(x_1)) \end{aligned}$$

Assumption 7 ensures that $(1-\tau)(1-\alpha) - \gamma > 0$ and $(1-\tau)(1-\alpha) - \gamma\phi > 0$. Furthermore, according to assumption 1, one has $y(x_2) > y(x_1)$. We conclude that

$\pi(a_1, x_2) - \pi(a_1, x_1) > 0$ and $\pi(a_2, x_2) - \pi(a_2, x_1) > 0$ and then $n(a_2, x_2) > n(a_2, x_1)$. $\square$

R ark 2 *Furthermore, it is immediate that* $n(a_1, x_2) \geq n(a_1, x_1)$. *Employment in* x_2 *is higher than employment in* x_1*, at every age* a.

Corollary 2 *Firms with higher productivity grow faster than firms with lower productivity between the young and old periods.*

Proof For a young firm, the employment variation is simply $n(a_1, x)$ as the initial level of employment is zero. We already demonstrated that $n(a_1, x_2) > n(a_1, x_1)$, meaning that a young productive firm grows faster than a young low productive firm. We now show this is true when the firm becomes old. Let us define:

$$
\begin{aligned}
g(x_1) &= n(a_2, x_1) - n(a_1, x_1) = \frac{1}{(1-\tau)b}\Big[(1-\omega\overline{\beta})\pi(a_2, x_1) - \omega\pi(a_1, x_1)\Big] \\
g(x_2) &= n(a_2, x_2) - n(a_1, x_2) = \frac{1}{(1-\tau)b}\Big[(1-\omega\overline{\beta})\pi(a_2, x_2) - \omega\pi(a_1, x_2)\Big]
\end{aligned}
$$

By calculating the difference $g(x_2) - g(x_1)$, one obtains:

$$
g(x_2) - g(x_1) = \frac{1}{(1-\tau)b}\Big[(1-\omega\overline{\beta})(\pi(a_2, x_2) - \pi(a_2, x_1)) - \omega(\pi(a_1, x_2) - \pi(a_1, x_1))\Big].
$$

Using assumptions 5-6, the above expression writes as follows:

$$
g(x_2) - g(x_1) = \frac{y(x_2) - y(x_1)}{(1-\tau)b}\left\{(1-\omega\overline{\beta})[(1-\tau)(1-\alpha) - \gamma\phi] - \omega[(1-\tau)(1-\alpha) - \gamma]\right\}.
$$

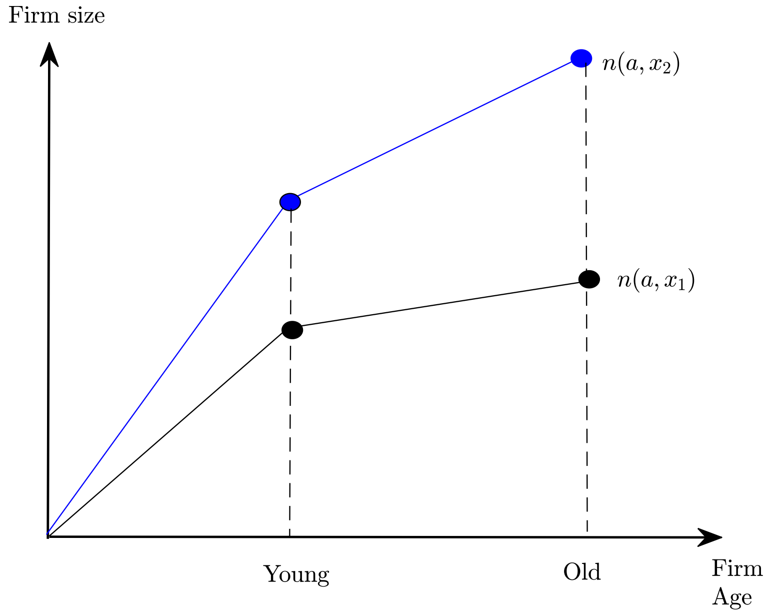
After some manipulation, and noting that $\overline{\beta} = \beta(1-\delta)(1-\omega) < 1$, the term in brackets can be rewritten as follows:

$$
\begin{aligned}
\Lambda &\equiv (1-\omega\overline{\beta})[(1-\tau)(1-\alpha) - \gamma\phi] - \omega[(1-\tau)(1-\alpha) - \gamma], \\
&= (1-\omega)[(1-\tau)(1-\alpha) - \gamma](1-\omega\beta(1-\delta)) \\
&\quad + \gamma(1-\phi)(1-\omega\beta(1-\omega)(1-\delta)).
\end{aligned}
$$

Assumption 7 ensures that $(1-\tau)(1-\alpha) - \gamma > 0$. Furthermore, one has $\omega < 1$, $\beta < 1$, $\delta < 1$ and $\phi \leq 1$. It follows that $\Lambda \geq 0$ and $g(x_2) - g(x_1) \geq 0$. $\square$

The results are summarized by figure 14, which illustrates the firm size over the life cycle with heterogenous productivities. Note that since the employment profile is concave, the parameter ϕ necessarily satisfies the conditions of Corollary 1.

Figure 14: Firm size heterogeneity over the life cycle



D.3 Elasticity with respect to corporate tax

We now investigate the impact of a tax variation on employment, distinguishing between long-term and short-term effects.

To define the long-term effect of a tax variation, we consider a constant tax profile. We analyze the impact of a permanent change in the tax profile and derive the firm's employment level with respect to the tax rate τ.

D fini ion 1 *The long-term semi-elasticity of firm employment with respect to the tax rate is given by:*

$$\mathcal{E}_{n(a,x),\tau} = \frac{1}{n(a,x)} \frac{\partial n(a,x)}{\partial \tau},$$

with,

$$\frac{\partial n(a_1,x)}{\partial \tau} = -\frac{[c_o(a_1,x) + \beta(1-\delta)(1-\omega)c_o(a_2,x)]}{(1-\tau)^2 b} \leq 0, \tag{45}$$

$$\frac{\partial n(a_2,x)}{\partial \tau} = \frac{\partial n(a_1,x)}{\partial \tau}(1-\omega) - \frac{c_o(a_2,x)}{(1-\tau)^2 b} \leq 0. \tag{46}$$

R ark 3 *The above elasticity is labeled long-term and refers to the impact of a permanent change in the tax profile ($d\tau = d\tau_1 = d\tau_2$) on employment through all channels. Note that in an infinite-horizon problem, the long-run impact would be the effect of the tax on the stationary level of employment.*

The hiring flow of a firm is impacted through different channels throughout its life cycle. This is discussed in detail in sections D.1.2 and D.1.3. In the case of a permanent tax increase, the hiring flow is reduced at each age because the operating cost is not tax deductible.

Let us now discuss the impact on employment. When the firm is young, there is no employment stock inherited from the past, so the employment stock is equal to the hiring flow. When the firm is old, the tax impacts new hires $h(a_2, x)$ as well as the employment inherited from the previous period $n(a_1, x)(1-\omega)$. Thus, the tax impacts the firm's employment when old through both the current hiring decisions and the inherited employment stock.

While our estimations cannot perfectly distinguish this aspect (past and current effects), we propose an alternative measure which simply eliminates the inheritance effects and corresponds to the short-term response of employment, or equivalently, the contemporaneous effects. We believe that the observed elasticity is somewhere between the two bounds.

D fini ion 2 *We consider a permanent change in the tax profile. The short-term semi-elasticity of firm employment with respect to the tax is given by:*

$$\Sigma_{n(a,x),\tau} = \frac{1}{n(a,x)} \left(\frac{\partial n(a,x)}{\partial \tau} \right)_{n(a_{-1},x)}$$

with,

$$\left(\frac{\partial n(a_1,x)}{\partial \tau} \right)_{n(a_0,x)} = -\frac{[c_o(a_1,x) + \beta(1-\delta)(1-\omega)c_o(a_2,x)]}{(1-\tau)^2 b} \leq 0, \tag{47}$$

$$\left(\frac{\partial n(a_2,x)}{\partial \tau} \right)_{n(a_1,x)} = -\frac{c_o(a_2,x)}{(1-\tau)^2 b} \leq 0. \tag{48}$$

R ark 4 *To compute the short-term elasticity, we derive employment given the employment level of the previous period. For an old firm, we discard the response of employment inherited from the young period.*

Finally, note that when young, the short-term and long-run elasticities are identical.

D.3.1 Young vs Old

R l 3 *For any level x and given **Assumption 4 and Result 1**, young firms are more sensitive to variation in the tax in the short-term.*

Proof We must verify that:

$$\Sigma_{n(a_1,x),\tau} \leq \Sigma_{n(a_2,x),\tau} \iff \frac{1}{n(a_1,x)}\left(\frac{\partial n(a_1,x)}{\partial \tau}\right)_{n(a_0,x)} \leq \frac{1}{n(a_2,x)}\left(\frac{\partial n(a_2,x)}{\partial \tau}\right)_{n(a_1,x)}$$

Substituting (47) and (48) into the above expression, one gets:

$$-\frac{1}{n(a_1,x)}\frac{c_o(a_1,x)+\overline{\beta}c_o(a_2,x)}{(1-\tau)^2 b} \leq -\frac{1}{n(a_2,x)}\frac{c_o(a_2,x)}{(1-\tau)^2 b}$$

or equivalently,

$$c_0(a_1,x) - \frac{n(a_1,x)}{n(a_2,x)}c_0(a_2,x) + \overline{\beta}c_0(a_2,x) \geq 0$$

Under assumption 3 and result 1, $c_o(a_1,x) > c_o(a_2,x)$ and $n(a_2,x) \geq n(a_1,x)$. It follows that $c_0(a_1,x) - \frac{n(a_1,x)}{n(a_2,x)}c_0(a_2,x) > 0$ and the above inequality is satisfied. □

R l 4 *For any level x and given assumption 3 and result 1, young firms are more sensitive to variation in the tax in the long-term.*

Proof. We need to verify that:

$$\mathcal{E}_{n(a_1,x),\tau} < \mathcal{E}_{n(a_2,x),\tau} \iff \frac{1}{n(a_1,x)}\frac{\partial n(a_1,x)}{\partial \tau} \leq \frac{1}{n(a_2,x)}\frac{\partial n(a_2,x)}{\partial \tau}.$$

Substituting (45) and (46) into the above expression, we get:

$$-\frac{1}{n(a_1,x)}\frac{c_0(a_1,x)+\overline{\beta}c_0(a_2,x)}{(1-\tau)^2 b} \leq -\frac{1}{n(a_2,x)}\left\{(1-\omega)\frac{c_0(a_1,x)+\overline{\beta}c_0(a_2,x)}{(1-\tau)^2 b} + \frac{c_0(a_2,x)}{(1-\tau)^2 b}\right\}$$

After some manipulations, we obtain:

$$\{c_0(a_1,x)+\overline{\beta}c_0(a_2,x)\}\left\{\frac{n(a_2,x)}{n(a_1,x)} - (1-\omega)\right\} \geq c_0(a_2,x).$$

Using equations (40) and (41), and after some algebraic manipulations, one obtains:

$$\frac{n(a_2,x)}{n(a_1,x)} - (1-\omega) \quad = \quad \frac{\pi(a_2,x)}{\pi(a_1,x) + \overline{\beta}\pi(a_2,x)}$$

Substituting into the previous inequality yields:

$$\left\{c_0(a_1,x) + \overline{\beta}c_0(a_2,x)\right\} \frac{\pi(a_2,x)}{\pi(a_1,x) + \overline{\beta}\pi(a_2,x)} \geq c_0(a_2,x),$$

which simplifies to:

$$c_0(a_1,x)\pi(a_2,x) \geq c_0(a_2,x)\pi(a_1,x).$$

Knowing that $c_0(a_1,x) > c_0(a_2,x)$ and $\pi(a_2,x) > \pi(a_1,x)$, the above inequality holds. □

R ark 5 *Recall that the elasticities are negative; one has, in absolute value:*

$$\left|\mathcal{E}_{n(a_1,x),\tau}\right| > \left|\mathcal{E}_{n(a_2,x),\tau}\right|$$

and

$$\left|\Sigma_{n(a_1,x),\tau}\right)| \geq \left|\Sigma_{n(a_2,x),\tau}\right|$$

D.3.2 Small vs Large

We know from result 2 that the largest firm in the economy is a firm of age a_2 and productivity level x_2 with $n(a_2,x_2)$ workers. More generally, we have $n(a_2,x_2) > n(a_2,x_1)$ (result 2), $n(a_1,x_2) > n(a_1,x_1)$ (see section D.2.2) and $n(a_2,x) > n(a_1,x), \forall x$ (result 1). The aim of this section is to study the impact of a tax variation according to the size of the firm. More precisely, we show that small firms are more sensitive than large ones to a tax variation, if the following inequalities are satisfied:

$$\begin{aligned}
(i) \quad \mathcal{E}_{n(a_1,x_1),\tau} &\leq \mathcal{E}_{n(a_2,x_2),\tau} \\
(ii) \quad \mathcal{E}_{n(a_1,x_2),\tau} &\leq \mathcal{E}_{n(a_2,x_2),\tau} \\
(iii) \quad \mathcal{E}_{n(a_2,x_1),\tau} &\leq \mathcal{E}_{n(a_2,x_2),\tau}
\end{aligned}$$

or in absolute value:

$$\begin{aligned}
\left|\mathcal{E}_{n(a_1,x_1),\tau}\right| &\geq \left|\mathcal{E}_{n(a_2,x_2),\tau}\right| \\
\left|\mathcal{E}_{n(a_1,x_2),\tau}\right| &\geq \left|\mathcal{E}_{n(a_2,x_2),\tau}\right| \\
\left|\mathcal{E}_{n(a_2,x_1),\tau}\right| &\geq \left|\mathcal{E}_{n(a_2,x_2),\tau}\right|
\end{aligned}$$

Note that to define the sensitivity of firms to a tax variation, we used the long-term elasticities.

R ark *Result 4 has several implications. It implies that condition (ii) is satisfied and that if (iii) holds, then (i) also holds.*

We have the following result.

R l 5 *Small firms are more sensitive to a tax variation than large ones in the sense of conditions (i), (ii) and (iii).*

Following Remark 6, we only need to demonstrate that (iii) is satisfied. If this is the case, we will have a set of sufficient conditions to show that small firms are necessarily more sensitive to tax variation than large firms, regardless of the characteristics of small firms (whether they are young with $x = x_1$ or $x = x_2$, or old with $x = x_1$).

Proof We need to show that:

$$\mathcal{E}_{n(a_2,x_1),\tau} \leq \mathcal{E}_{n(a_2,x_2),\tau} \tag{49}$$

with,

$$\begin{aligned}
\mathcal{E}_{n(a_2,x_1),\tau} &= -\frac{1}{n(a_2,x_1)}\left\{\frac{(1-\omega)(c_0(a_1,x_1)+\overline{\beta}c_0(a_2,x_1))}{(1-\tau)^2 b} + \frac{c_0(a_2,x_1)}{(1-\tau)^2 b}\right\}, \\
\mathcal{E}_{n(a_2,x_2),\tau} &= -\frac{1}{n(a_2,x_2)}\left\{\frac{(1-\omega)(c_0(a_1,x_2)+\overline{\beta}c_0(a_2,x_2))}{(1-\tau)^2 b} + \frac{c_0(a_2,x_2)}{(1-\tau)^2 b}\right\}.
\end{aligned}$$

Substituting into inequality (49), one gets:

$$\begin{aligned}
&\frac{1}{n(a_2,x_1)}\left((1-\omega)(c_0(a_1,x_1)+\overline{\beta}c_0(a_2,x_1))+c_0(a_2,x_1)\right) \\
&\geq \frac{1}{n(a_2,x_2)}\left((1-\omega)(c_0(a_1,x_2)+\overline{\beta}c_0(a_2,x_2))+c_0(a_2,x_2)\right),
\end{aligned}$$

or,

$$\frac{n(a_2,x_2)}{n(a_2,x_1)} \geq \frac{(1-\omega)(c_0(a_1,x_2)+\overline{\beta}c_0(a_2,x_2))+c_0(a_2,x_2)}{(1-\omega)(c_0(a_1,x_1)+\overline{\beta}c_0(a_2,x_1))+c_0(a_2,x_1)}. \tag{50}$$

Assumptions (5)-(6) imply that:

$$\begin{aligned} c_0(a_1,x_1) &= \gamma y(x_1) \quad c_0(a_2,x_1) = \gamma\phi y(x_1), \\ c_0(a_1,x_2) &= \gamma y(x_2) \quad c_0(a_2,x_2) = \gamma\phi y(x_2), \end{aligned}$$

with $y(x_2) > y(x_1)$. After some calculations, (40) and (41) write as follows:

$$\begin{aligned} n(a_2,x_1) &= \frac{1}{(1-\tau)b}\{(1-\omega)[(1-\tau)(1-\alpha)-\gamma] \\ &+ (1+\overline{\beta}(1-\omega))[(1-\tau)(1-\alpha)-\gamma\phi]\}\, y(x_1), \quad (51) \\ n(a_2,x_2) &= \frac{1}{(1-\tau)b}\{(1-\omega)[(1-\tau)(1-\alpha)-\gamma] \\ &+ (1+\overline{\beta}(1-\omega))[(1-\tau)(1-\alpha)-\gamma\phi]\}\, y(x_2). \quad (52) \end{aligned}$$

Furthermore, the terms in the RHS of expression (50) write:

$$(1-\omega)(c_0(a_1,x_1)+\overline{\beta}c_0(a_2,x_1))+c_0(a_2,x_1) = \{(1-\omega)(\gamma+\overline{\beta}\gamma\phi)+\gamma\phi\}\, y(x_1) \tag{53}$$

$$(1-\omega)(c_0(a_1,x_2)+\overline{\beta}c_0(a_2,x_2))+c_0(a_2,x_2) = \{(1-\omega)(\gamma+\overline{\beta}\gamma\phi)+\gamma\phi\}\, y(x_2) \tag{54}$$

Substituting (51), (52), (53) and (54) in (50), one gets:

$$\frac{(1-\omega)[(1-\tau)(1-\alpha)-\gamma]+(1+\overline{\beta}(1-\omega))[(1-\tau)(1-\alpha)-\gamma\phi]\, y(x_2)}{(1-\omega)[(1-\tau)(1-\alpha)-\gamma]+(1+\overline{\beta}(1-\omega))[(1-\tau)(1-\alpha)-\gamma\phi]\, y(x_1)} \geq \frac{y(x_2)}{y(x_1)},$$

which simply binds. □

D.3.3 Young vs small

Lastly, our objective is to show how young firms, including both low and high productive firms, are sensitive to tax variations compared to small firms. The latter group encompasses both the young firms and old firms with low productivity (x_1). Contrary to previous analytical results, we are now dealing with different groups of firms. The comparison in terms of elasticity requires knowledge of the distribution of firms before and after the tax change. For simplicity, we only investigate the case where the distribution is fixed. The modeling of the endogenous distribution of firms and the general equilibrium will be addressed in the quantitative model presented in the next section.

There is a fixed mass of entering firms equal to 1. Suppose that the proportion of young firms with productivity x_1 is p, and the proportion of young firms with productivity x_2 is $(1-p)$. After period 1, firms are destroyed at a rate δ, and the life cycle ends after period 2. The mass of old firms with productivity x_1 is $p(1-\delta)$, and the mass of old firms with productivity x_2 is $(1-p)(1-\delta)$.

R I *Young firms are more sensitive to a tax variation than small firms.*

Proof We need to show that:

$$\frac{p\frac{\partial n(a_1,x_1)}{\partial \tau}+(1-p)\frac{\partial n(a_1,x_2)}{\partial \tau}}{pn(a_1,x_1)+(1-p)n(a_1,x_2)} \leq \frac{p\frac{\partial n(a_1,x_1)}{\partial \tau}+(1-p)\frac{\partial n(a_1,x_2)}{\partial \tau}+p(1-\delta)\frac{\partial n(a_2,x_1)}{\partial \tau}}{pn(a_1,x_1)+(1-p)n(a_1,x_2)+p(1-\delta)n(a_2,x_1)}.$$

Note that the above expression can be written as:

$$\frac{pn(a_1,x_1)+(1-p)n(a_1,x_2)+p(1-\delta)n(a_2,x_1)}{pn(a_1,x_1)+(1-p)n(a_1,x_2)} \leq \frac{p\frac{\partial n(a_1,x_1)}{\partial \tau}+(1-p)\frac{\partial n(a_1,x_2)}{\partial \tau}+p(1-\delta)\frac{\partial n(a_2,x_1)}{\partial \tau}}{p\frac{\partial n(a_1,x_1)}{\partial \tau}+(1-p)\frac{\partial n(a_1,x_2)}{\partial \tau}},$$

which simplifies to:

$$1+\frac{p(1-\delta)n(a_2,x_1)}{pn(a_1,x_1)+(1-p)n(a_1,x_2)} \leq 1+\frac{p(1-\delta)\frac{\partial n(a_2,x_1)}{\partial \tau}}{p\frac{\partial n(a_1,x_1)}{\partial \tau}+(1-p)\frac{\partial n(a_1,x_2)}{\partial \tau}},$$

and,

$$\frac{n(a_2,x_1)}{pn(a_1,x_1)+(1-p)n(a_1,x_2)} \leq \frac{\frac{\partial n(a_2,x_1)}{\partial \tau}}{p\frac{\partial n(a_1,x_1)}{\partial \tau}+(1-p)\frac{\partial n(a_1,x_2)}{\partial \tau}}.$$

To further simplify the demonstration, recall that $\overline{\beta}=\beta(1-\delta)(1-\omega)$ and let us define $\Omega=(1-\tau)(1-\alpha)<1$. The employment level can be written as:

$$n(a_1,x) = \frac{y(x)}{(1-\tau)b}\left\{\Omega-\gamma+\overline{\beta}(\Omega-\gamma\phi)\right\},$$

and,

$$
\begin{aligned}
n(a_2, x) &= (1-\omega)n(a_1, x) + \frac{y(x)}{(1-\tau)b}\{\Omega - \gamma\phi)\} \\
&= \frac{y(x)}{(1-\tau)b}\left\{(1-\omega)(\Omega-\gamma) + ((1-\omega)\overline{\beta}+1)(\Omega-\gamma\phi)\right\}.
\end{aligned}
$$

Differentiating with respect to τ, one gets:

$$
\begin{aligned}
\frac{\partial n(a_1, x)}{\partial \tau} &= -\frac{y(x)\gamma(1+\overline{\beta}\phi)}{(1-\tau)^2 b} \\
\frac{\partial n(a_2, x)}{\partial \tau} &= -\frac{y(x)\gamma\left\{(1-\omega)(1+\overline{\beta}\phi)+\phi\right\}}{(1-\tau)^2 b}
\end{aligned}
$$

We need to verify that the following condition holds:

$$
\frac{n(a_2, x_1)}{pn(a_1, x_1) + (1-p)n(a_1, x_2)} > \frac{\frac{\partial n(a_2, x_1)}{\partial \tau}}{p\frac{\partial n(a_1, x_1)}{\partial \tau} + (1-p)\frac{\partial n(a_1, x_2)}{\partial \tau}}.
$$

Let us define:

$$
\begin{aligned}
B &= \frac{n(a_2, x_1)}{pn(a_1, x_1) + (1-p)n(a_1, x_2)}, \\
A &= \frac{\frac{\partial n(a_2, x_1)}{\partial \tau}}{p\frac{\partial n(a_1, x_1)}{\partial \tau} + (1-p)\frac{\partial n(a_1, x_2)}{\partial \tau}}.
\end{aligned}
$$

A and B can be written as follows:

$$
\begin{aligned}
A &= \frac{y(x_1)\left\{(1-\omega)(1+\overline{\beta}\phi)+\phi\right\}}{(1+\overline{\beta}\phi)\left\{py(x_1)+(1-p)y(x_2)\right\}}, \\
B &= \frac{y(x_1)\left\{(1-\omega)(\Omega-\gamma) + ((1-\omega)\overline{\beta}+1)(\Omega-\gamma\phi)\right\}}{\left\{\Omega-\gamma+\overline{\beta}(\Omega-\gamma\phi)\right\}\left\{py(x_1)+(1-p)y(x_2)\right\}}.
\end{aligned}
$$

The condition $B > A$ becomes:

$$
\frac{y(x_1)\left\{(1-\omega)(\Omega-\gamma) + ((1-\omega)\overline{\beta}+1)(\Omega-\gamma\phi)\right\}}{\left\{\Omega-\gamma+\overline{\beta}(\Omega-\gamma\phi)\right\}\left\{py(x_1)+(1-p)y(x_2)\right\}} > \frac{y(x_1)\left\{(1-\omega)(1+\overline{\beta}\phi)+\phi\right\}}{(1+\overline{\beta}\phi)\left\{py(x_1)+(1-p)y(x_2)\right\}},
$$

which reduces to:

$$\frac{\{(1-\omega)(\Omega-\gamma)+((1-\omega)\overline{\beta}+1)(\Omega-\gamma\phi)\}}{\{\Omega-\gamma+\overline{\beta}(\Omega-\gamma\phi)\}} > \frac{\{(1-\omega)(1+\overline{\beta}\phi)+\phi\}}{(1+\overline{\beta}\phi)}.$$

We deduce:

$$\begin{aligned}&(1+\overline{\beta}\phi)\left\{(1-\omega)(\Omega-\gamma)+((1-\omega)\overline{\beta}+1)(\Omega-\gamma\phi)\right\}\\&>\left\{(1-\omega)(1+\overline{\beta}\phi)+\phi\right\}\left\{\Omega-\gamma+\overline{\beta}(\Omega-\gamma\phi)\right\},\end{aligned}$$

Rearranging yields:

$$\begin{aligned}&\{\Omega-\gamma\}\left\{(1+\overline{\beta}\phi)(1-\omega)-(1-\omega)(1+\overline{\beta}\phi)-\phi\right\}\\&+\{\Omega-\gamma\phi\}\left\{(1+\overline{\beta}\phi)(\overline{\beta}(1-\omega)+1)-\overline{\beta}((1-\omega)(1+\overline{\beta}\phi)+\phi)\right\}>0,\end{aligned}$$

and finally:

$$\begin{aligned}&\{\Omega-\gamma\}\left\{(1+\overline{\beta}\phi)(1-\omega)-(1-\omega)(1+\overline{\beta}\phi)-\phi\right\}\\&+\{\Omega-\gamma\phi\}\left\{(1+\overline{\beta}\phi)(1-\omega)\overline{\beta}+1+\overline{\beta}\phi-\overline{\beta}(1-\omega)(1+\overline{\beta}\phi)-\overline{\beta}\phi\right\}>0.\end{aligned}$$

Further simplifying leads to:

$$\{\Omega-\gamma\}\{-\phi\}+\{\Omega-\gamma\phi\}>0 \iff \Omega(1-\phi)>0.$$

Knowing that $\phi \leq 1$, the above inequality is satisfied. □

E Simulations

- To mute the job creation channel, we impose that the decision rule resulting from a change in the tax, labeled $D^{new}(n, \varepsilon, a, x)$, is the same as that of the benchmark

$D^{bench}(n, \varepsilon, a, x)$ for any job creation. Formally:

$$\mathcal{D}^{new}(n,\varepsilon,a,x) = \begin{cases} \mathcal{D}^{bench}(n,\varepsilon,a,x) & \text{if} \quad \mathcal{D}^{bench}(n,\varepsilon,a,x) \leq n \\ & \text{and} \quad \mathcal{D}^{new}(n,\varepsilon,a,x) > n \\ \mathcal{D}^{bench}(n,\varepsilon,a,x) & \text{if} \quad D^{bench}(n,\varepsilon,a,x) > n \\ & \text{and} \quad \mathcal{D}^{new}(n,\varepsilon,a,x) > n \\ \mathcal{D}^{new}(n,\varepsilon,a,x) & \text{otherwise} \end{cases}$$

The first set of conditions is such that if the decision rule following the change in the tax involves a job creation ($n' > n$) while this is not the case in the benchmark case, we impose the decision rule to be the same as in the benchmark case. The second set of conditions involves that if there is job creation in both the benchmark and the alternative tax rate, the job creation is of the same magnitude as in the benchmark. In addition, for new firms, the decision rule is the same as in the benchmark: $\mathcal{D}_0^{new}(x) = \mathcal{D}_0^{bench}(x)$ because it pertains only to job creation.

- We apply the same reasoning to mute job destruction. We impose the following restriction on the decision rule:

$$\mathcal{D}^{new}(n,\varepsilon,a,x) = \begin{cases} \mathcal{D}^{bench}(n,\varepsilon,a,x) & \text{if} \quad \mathcal{D}^{bench}(n,\varepsilon,a,x) \geq n \\ & \text{and} \quad \mathcal{D}^{new}(n,\varepsilon,a,x) < n \\ \mathcal{D}^{bench}(n,\varepsilon,a,x) & \text{if} \quad \mathcal{D}^{bench}(n,\varepsilon,a,x) < n \\ & \text{and} \quad \mathcal{D}^{new}(n,\varepsilon,a,x) < n \\ \mathcal{D}^{new}(n,\varepsilon,a,x) & \text{otherwise} \end{cases}$$

The job destruction channel is muted (relative to the benchmark) if the new decision rule involves job destructions while the benchmark does not. It is of the same magnitude as in the benchmark if job destruction occurs in both cases.

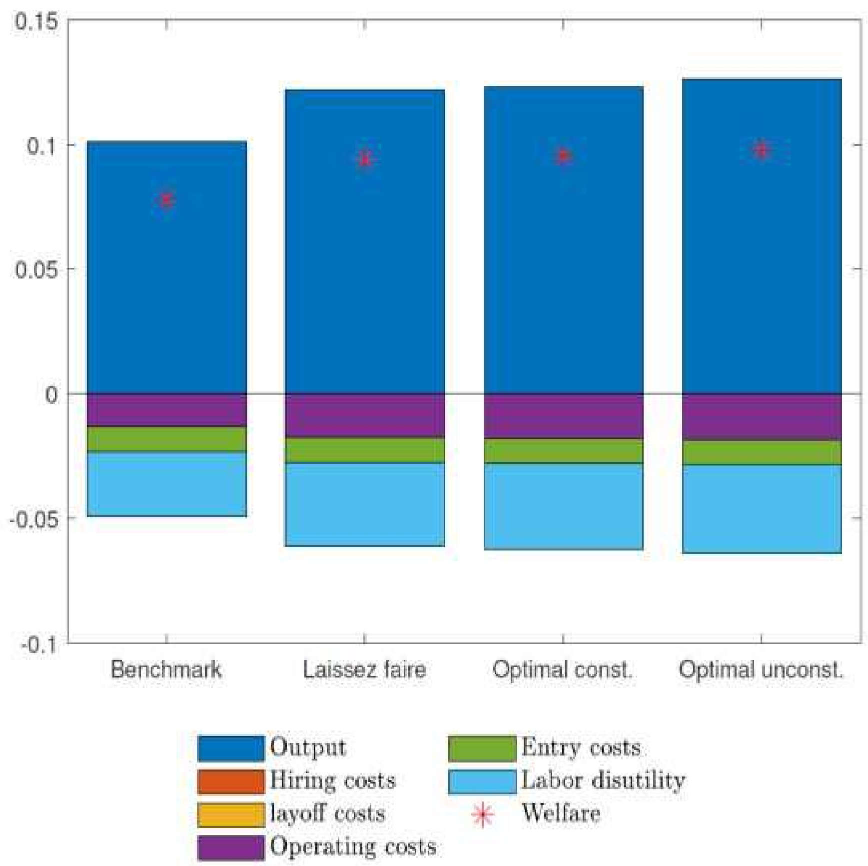
- The entry channel is muted by simply assuming that $e^{new}(x) = e^{bench}(x)$, which involves no recalculation of entries.

- We neutralize the exit channel by imposing the same exit strategy in the new decision rule as in the benchmark:

$$\mathcal{D}^{new}(n,\varepsilon,a,x) = \begin{cases} \mathcal{D}^{bench}(n,\varepsilon,a,x) & \text{if} \quad \mathcal{D}^{bench}(n,\varepsilon,a,x) \geq 0 \\ & \text{and} \quad \mathcal{D}^{new}(n,\varepsilon,a,x) = 0 \\ \mathcal{D}^{bench}(n,\varepsilon,a,x) & \text{if} \quad \mathcal{D}^{bench}(n,\varepsilon,a,x) = 0 \\ & \text{and} \quad \mathcal{D}^{new}(n,\varepsilon,a,x) > 0 \\ \mathcal{D}^{new}(n,\varepsilon,a,x) & \text{otherwise} \end{cases}$$

- The general equilibrium effect is neutralized by removing the feedback effect of the vacancy-filling rate (step 7 from Figure 9) on firm profits.

Figure 15: Welfare decomposition



A Full model

DEFINITION 1. *-Given exogenous processes for age a, and idiosyncratic productivity* ε*, the equilibrium is a list of (i) quantities* $m(x)$, $u(x)$, $\boldsymbol{v(x)}$, $N(x)$, $f(x)$, $q(x)$, $\theta(x)$ *and* $s(x)$*; (ii) Optimal decisions* $D(n,\varepsilon,a,x)$, $D_0(x)$*; (iii) Entry mass* $e(x)$*; (iv) Stationary distributions of firms* $\lambda(n,\varepsilon,a,x)$*; ; and (v) fiscal surplus FS and lump-sum transfer T, satisfying the following conditions:*

(i) $m(x)$, $u(x)$, $\boldsymbol{v(x)}$, $N(x)$, $f(x)$, $q(x)$, $\theta(x)$ *and* $s(x)$ *are the solutions of the matching function (55), the number of unemployed workers (56), the number of vacancies (57), the aggregate level of employment (58), the job-finding rate (59), the vacancy-filling rate (60), the tightness (61), and the separation flows (62), respectively:*

$$
\begin{aligned}
m(x) &= m(u(x),\boldsymbol{v}(x)), && (55)\\
u(x) &= L(x) - N(x), && (56)\\
\boldsymbol{v}(x) &= \sum_a \int_n \int_\varepsilon \lambda(n,\varepsilon,a,x)\mathcal{D}_v(n,\varepsilon,a,x)\, d\varepsilon dn + e(x)\mathcal{D}_{0,v}(x), && (57)\\
N(x) &= \sum_a \int_n \int_\varepsilon \lambda(n,\varepsilon,a,x) \times n\, d\varepsilon dn, && (58)\\
f(x) &= \frac{m(x)}{u(x)}, && (59)\\
q(x) &= \frac{m(x)}{\boldsymbol{v}(x)}, && (60)\\
\theta(x) &= \frac{\boldsymbol{v}(x)}{u(x)}, && (61)\\
S(x) &= \sum_a \int_n \int_\varepsilon \lambda(n,\varepsilon,a,x) n \Big\{ (\delta n + (1-\delta) \Big[\mathbb{1}\{n' < n\} \\
&+ \mathbb{1}\{n' \geq n\} \int_{\varepsilon'} (1 - \mathcal{I}(n',\varepsilon',a',x)) dG(\varepsilon'|\varepsilon) \Big] \Big\} d\varepsilon dn. && (62)
\end{aligned}
$$

(ii) The optimal firm employment decision $D(n,\varepsilon,a,x)$ *(incumbent firms) and* $D_0(x)$ *(new*

firms), are solution to the problem (63), and (64),respectively:

$$\begin{aligned} \Pi(n,\varepsilon,a,x) &= \max_{n',v,f}\Big\{[\pi(n,\varepsilon,a,x)-c_v(x)v^{\phi}-c_d\ell](1-\tau)-c_o(a,x)n \\ &+ \beta(1-\delta)\int_{\varepsilon'}\Lambda(n',\varepsilon',a',x)dG(\varepsilon'|\varepsilon)\Big\} \\ \text{with } n' &= n+q(x)v-\ell \\ v &\geq 0 \\ \ell &\geq 0 \end{aligned} \tag{63}$$

$$\begin{aligned} \Pi_0(x) &= \max_{n',v}\Big\{-c_v(x)v^{\phi}+\beta\int_{\varepsilon'}\Pi(n',\varepsilon',a_1,x)\mathcal{I}(n',\varepsilon',a_1,x)dG_0(\varepsilon')\Big\} \\ \text{with } n' &= q(x)v \\ v &\geq 0 \end{aligned} \tag{64}$$

(iii) $e(x)$ *satisfies (65):*

$$e(x) = \bar{e}(x)\exp\big(\eta(x)(\Pi_0(x)-c_e(x))\big). \tag{65}$$

(iv) *The stationary distribution* $\lambda(n,\varepsilon,a,x)$ *solves (66):*

$$\lambda(n',\varepsilon',a',x)=\mathcal{I}(n',\varepsilon',a',x)\left\{\begin{array}{l}(1-\delta)\int_n\int_\varepsilon \mathbb{1}\{n'=\mathcal{D}(n,\varepsilon,a,x)\}\lambda(n,\varepsilon,a,x)g(\varepsilon'|\varepsilon)d\varepsilon dn \\ \\ +e(x)g_0(\varepsilon')\mathbb{1}\{n'=\mathcal{D}_0(x)\}\mathbb{1}\{a'=a_1\}\end{array}\right\} \tag{66}$$

(v) FS *and* T *satisfy the government budget defined by (67) and (68):*

$$\begin{aligned} FS &= \sum_x\sum_a\int_n\int_\varepsilon \lambda(n,e,a,x)\big[\pi(n,\varepsilon,x)-c_v(x)v^{\phi}-c_d\ell\big]\tau\, d\varepsilon dn \\ &\quad -\sum_x b(x)u(x), \end{aligned} \tag{67}$$

$$T = FS \tag{68}$$

B Solution algorithm

B.1 Overview

The numerical solution technique consists in finding :

(a) a sequence of control variables by iterating backward from terminal conditions,

(b) a sequence of state variables by iterating forward from initial conditions.

Remark. As the variables obtained from stationary distribution (state variables) are not present in the forward looking equation (Bellman), the model is bloc recursive. Consequently, we can solve for (a) and for (b) sequentially.

B.2 State-space

The state-space (see summary in Table 9) is discretized and given by a set of:

$$\begin{aligned} \text{Employment} \quad & n \in \mathcal{N} = \{n_1, ..., n_N\}, \\ \text{idiosyncratic productivity} \quad & \varepsilon \in \mathcal{E} = \{\varepsilon_1, ..., \varepsilon_{n_E}\}, \\ \text{age} \quad & a \in \mathcal{A} = \{a_0, ..., a_A\}, \\ \text{technology level} \quad & x \in \mathcal{X} = \{x_1, ..., x_X\}. \end{aligned}$$

Employment. The employment grid includes $N = 201$ points from 0 to 1. As there are more small firms than large ones, we assume that the grid is non-linear with more points for small employment levels and fewer points on larger employment levels. To build the non linear employment grid consider the linear function:

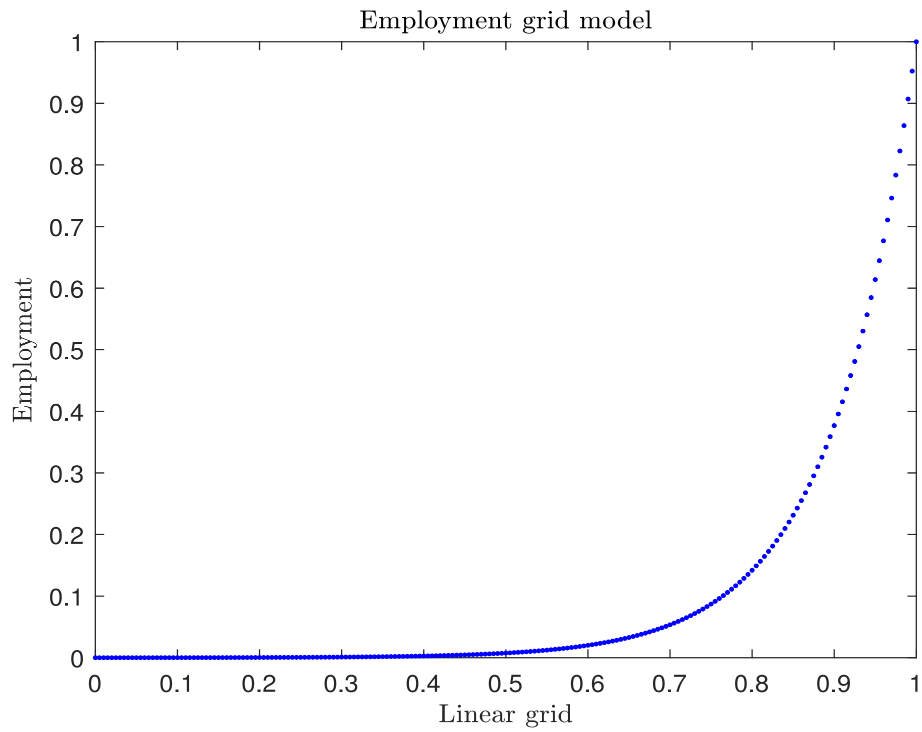
$$\begin{aligned} h_i &= \underline{h} + \frac{i\Delta_h - \underline{h}}{\overline{h} - \underline{h}} \\ \Delta_h &= \frac{\overline{h} - \underline{h}}{n_N}, \quad i = 1, ..., N, \end{aligned}$$

with $\underline{h} = 0$ and $\overline{h} = 10$ The non linear employment grid is obtained as:

$$n_j = n_1 + \frac{\exp(h_j) - n_1}{h_N - h_1},$$

with $n_1 = 0$. The employment levels are depicted in Figure 16.

Figure 16: EMPLOYMENT GRID



Age The age evolves deterministically. Every period of the life-cycle age moves from age $a = a_i$ to $a' = a_{i+1}$. Note that we consider quarterly frequencies so that $a_{i+1} - a_i$ represent one quarter.

Idiosyncratic productivity shock For the idiosyncratic productivity shock, we consider a first-order autoregressive process:

$$\log x' = \rho_x \log x + \sigma_x \varepsilon'_x$$

with $\varepsilon'_x \sim \mathcal{N}(0,1)$. The process is discretized using the method of Rouwenhorst. This discretization technique provides the grid nodes $\{x_1, ..., x_H\}$ and the transition matrix:

$$G(x'|x) = \begin{bmatrix} G_{1,1} & \cdots & G_{1,n_X} \\ \vdots & \ddots & \vdots \\ G_{n_X,1} & \cdots & G_{n_X,n_X} \end{bmatrix}$$

The probability density function for individual productivity of the new firm G_0 corresponds to the unconditional density of $G(.)$:

$$G_0 = G'G_0$$

Technology We consider three level of technology $n_x = 3$.

Table 9: STATE SPACE

Parameter	Symbol	Value
Number of firm age	n_a	45
Time frequency (quarter)	dt	0.25
Age grid	$\{a_1, ..., a_A\}$	[0, 11]
Number of firm size	n_N	201
Size grid model	$\{n_1, ..., n_N\}$	[0.0, 1.0]
Non-linear grid parameter	ζ	100
Number of technology level	n_x	3
Number of idiosyncratic productivity	n_ε	20
Distribution	$G(.)$	log-$\mathcal{N}(0, \sigma_\varepsilon^2)$

B.3 Algorithm

The algorithm consists in iterating backward the value functions, starting from the terminal condition and using the Bellman equations in a recursive manner. The terminal condition corresponds the final age a_A that is pervasive. We need first to use a fixed point algorithm to find the value function at the terminal age.

Step 0. Set big loop iteration $i = 0$. Initialize the vacancy filling rate $q^i(x)$ and the optimal exit/quit strategy $\mathcal{I}^i(n, \varepsilon, a, x) = 1$

Step 1. For each employment level $n \in \mathcal{N}$ and technology $x \in \mathcal{X}$,

1.a and next period employment $n' \in \mathcal{N}$ compute employment adjustment costs:

$$\Delta(n, n', x) = \begin{cases} \left(\frac{n'-n}{q^i(x)}\right)_v^\phi & \text{if } n' > n \\ (n - n')c_d & \text{if } n' < n \\ 0 & \text{otherwise} \end{cases}$$

1.b and for each idiosyncratic productivity level ε, age a, compute instantaneous profit:

$$
\begin{aligned}
y(n,\varepsilon,x) &= (n\,\varepsilon\,\bar{y}(x))^{\alpha} \\
w(\varepsilon,x) &= \xi(\varepsilon\,\bar{y}(x))^{\gamma} \\
\pi(n,\varepsilon,x) &= y(n,\varepsilon,x) - w(\varepsilon,x)n \\
\text{profit}(n,n',\varepsilon,a,x) &= \pi(n,\varepsilon,x) - \Delta(n,n',x) - c_o(a,x)n
\end{aligned}
$$

Step 2. Set age $a = a_A$ and an initial firm value $\Pi(n,\varepsilon,a_A,x)^k$ at iteration $k = 0$. For each employment level $n \in \mathcal{N}$, idiosyncratic productivity $\varepsilon \in \mathcal{E}$, and technology $x \in \mathcal{X}$,

2.a compute optimal employment decision n':

$$
\begin{aligned}
\Pi^{k+1}(n,\varepsilon,a_A,x) &= \max_{n'}\left\{\text{profit}(n,n',\varepsilon,a_A,x) + \beta(1-\delta)\sum_{\varepsilon'}\Lambda(n',\varepsilon',a_A,x)g(\varepsilon'|\varepsilon)\right\} \\
\Lambda(n,\varepsilon,a,x) &= \mathcal{I}^{i}(n,\varepsilon,a,x)\Pi^{k}(n,\varepsilon,a,x) + (1-\mathcal{I}^{i}(n,\varepsilon,a,x))(-c_x)
\end{aligned}
$$

2.b Compute Euclidian distance and check if it is higher than ϵ

$$
\frac{||\Pi^{k+1}(n,\varepsilon,a_A,x) - \Pi^{k}(n,\varepsilon,a_A,x)||}{||\Pi^{k}(n,\varepsilon,a_A,x)||} < \epsilon
$$

If the condition is satisfies go to **Step 3.**, otherwise set $k = k+1$ and return to step 2.a

Step 3. For each age $a = a_A - 1, a_A - 2, ..., a_1$, employment level $n \in \mathcal{N}$, idiosyncratic productivity $\varepsilon \in \mathcal{E}$, and technology $x \in \mathcal{X}$, compute optimal employment decision:

$$
\Pi(n,\varepsilon,a,x) = \max_{n'}\left\{\text{profit}(n,n',\varepsilon,a,x) + \beta(1-\delta)\sum_{\varepsilon'}\Lambda(n',\varepsilon',a+1,x)^{k}g(\varepsilon'|\varepsilon)\right\}
$$

Step 4. Given $\Pi(n,\varepsilon,a_1,x)$, compute new firm value function and optimal employment decision:

$$\Pi_0(x) = \max_{n'} \left\{ -c_v(x)\Delta(n,n',x) + \beta \sum_{\varepsilon'} \Lambda(n',\varepsilon',a_1,x) g_0(\varepsilon') \right\}$$

Step 5. Solve for workers value function using a fixed point iteration. Set iteration $k = 0$ and initialize value functions $W^k(n,\varepsilon,a,x)$, $U^k(x)$ and $\Omega^k(n,\varepsilon,a,x)$, as well as the density function $\lambda(n,\varepsilon,a,x)$ (which will be revised later).

5.a Given optimal decision $n' = \mathcal{D}(n,\varepsilon,a,x)$, for each employment level $n \in \mathcal{N}$, age $a \in \mathcal{A}$, idiosyncratic productivity $\varepsilon \in \mathcal{E}$, and technology $x \in \mathcal{X}$, compute

$$\begin{aligned} W^{k+1}(n,\varepsilon,a,x) &= w(\varepsilon,x) + \beta\delta U^k(x) + \beta(1-\delta)\mathbb{1}\{n' \geq n\} \sum_{\varepsilon'} \Omega^k(n',\varepsilon',a',x) g(\varepsilon'|\varepsilon) \\ &+ \beta(1-\delta)\mathbb{1}\{n' < n\} \sum_{\varepsilon'} \left(\frac{n-n'}{n} U^k(x) + \frac{n'}{n} \Omega^k(n',\varepsilon',a',x) \right) g(\varepsilon'|\varepsilon) \end{aligned}$$

5.b Compute $\overline{W}(x)$, the expected value of the worker in case of contact with a firm:

$$\begin{aligned} \overline{W}(x) &= \sum_a \sum_n \sum_\varepsilon \left[\omega(n,\varepsilon,a,x) \sum_{\varepsilon'} \Omega^k(n',\varepsilon',a',x) g(\varepsilon'|\varepsilon) \right] d\varepsilon dn \\ &+ \omega_0(x) \sum_{\varepsilon'} \Omega^k(n',\varepsilon',a_1,x) g_0(\varepsilon') \end{aligned}$$

5.c Compute unemployed value function:

$$U(x)^{k+1} = b(x) + \beta \left(f(x)\overline{W}(x) + (1-f(x))U^k(x) \right),$$

5.d Check if:

$$\frac{||W^{k+1}(n,\varepsilon,a_A,x) - W^k(n,\varepsilon,a_A,x)||}{||W^k(n,\varepsilon,a_A,x)||} < \epsilon$$

$$\frac{||U^{k+1}(x) - U^k(x)||}{||U^k(x)||} < \epsilon$$

if yes, go to **Step 6.**, otherwise set $k = k+1$, compute

$$\Omega^k(n,\varepsilon,a,x) = \mathcal{I}^i(n,\varepsilon,a,x)W^k(n,\varepsilon,a,x) + (1-\mathcal{I}^i(n,\varepsilon,a,x))U^k(x),$$

and return to step 5.a.

Step 6. Solve for entry

$$e(x) \quad = \quad \bar{e}(x) \exp\left(\eta(x)(\Pi_0(x) - c_e(x))\right).$$

Step 7. Given optimal decision $n' = \mathcal{D}(n, \varepsilon, a, x)$, for each employment level $n \in \mathcal{N}$, age $a \in \mathcal{A}$, idiosyncratic productivity $\varepsilon \in \mathcal{E}$, and technology $x \in \mathcal{X}$, solve for the stationary distribution by iterating until convergence of

$$\lambda(n', \varepsilon', a', x) = \mathcal{I}(n', \varepsilon', a', x) \left\{ \begin{array}{c} (1-\delta) \sum_n \sum_\varepsilon \mathbb{1}\{n' = \mathcal{D}(n, \varepsilon, a, x)\} \lambda(n, \varepsilon, a, x) g(\varepsilon' | \varepsilon) d\varepsilon dn \\ \\ +e(x) g_0(\varepsilon') \mathbb{1}\{n' = \mathcal{D}_0(x)\} \mathbb{1}\{a' = a_1\} \end{array} \right\}$$

Step 8. Compute tax and transfers using:

$$\begin{aligned} FS \quad &= \quad \sum_x \sum_a \sum_n \sum_\varepsilon \lambda(n, e, a, x) \left[\pi(n, \varepsilon, a, x) - \Delta(n, n', x)\right] \tau \, d\varepsilon dn \\ &\quad - \sum_x b(x) u(x) \\ T \quad &= \quad FS \end{aligned}$$

Step 9. Compute general equilibrium quantities as:

$$
\begin{aligned}
N(x) &= \sum_a \sum_n \sum_\varepsilon \lambda(n,\varepsilon,a,x) \times n, \\
u(x) &= L(x) - N(x), \\
\boldsymbol{\nu}(x) &= \sum_a \sum_n \sum_\varepsilon \lambda(n,\varepsilon,a,x)(\mathcal{D}(n,\varepsilon,a,x) - n)\mathbb{1}\{n' > n\} + e(x)\mathcal{D}_0(x) \\
S(x) &= \sum_a \sum_n \sum_\varepsilon \lambda(n,\varepsilon,a,x) n \Bigg\{ (\delta n + (1-\delta)\Big[\mathbb{1}\{n' < n\} \\
&+ \mathbb{1}\{n' \geq n\} \sum_{\varepsilon'} (1 - \mathcal{I}^i(n',\varepsilon',a',x)) g(\varepsilon'|\varepsilon)\Big]\Bigg\}, \\
m(x) &= m(u(x), \boldsymbol{\nu}(x)), \\
f(x) &= \frac{m(x)}{u(x)}, \\
q^{i+1}(x) &= \frac{m(x)}{\boldsymbol{\nu}(x)}, \\
\theta(x) &= \frac{\boldsymbol{\nu}(x)}{u(x)}.
\end{aligned}
$$

Step 10. Compute optimal exit/quit decision:

$$
\begin{aligned}
\mathcal{I}_f(n,\varepsilon,a,x) &= \mathbb{1}\{\Pi(n,\varepsilon,a,x) \geq -c_x\} \\
\mathcal{I}_w(n,\varepsilon,a,x) &= \mathbb{1}\{W(n,\varepsilon,a,x) \geq U(x)\} \\
\mathcal{I}^{i+1}(n,\varepsilon,a,x) &= \mathcal{I}_w(n,\varepsilon,a,x) \times \mathcal{I}_f(n,\varepsilon,a,x)
\end{aligned}
$$

Step 11. Check if

$$
\begin{aligned}
\frac{||q^{i+1}(x) - q^i(x)||}{||q^i(x)||} &< \epsilon \\
\frac{||\mathcal{I}^{i+1}(n,\varepsilon,a,x) - \mathcal{I}^i(n,\varepsilon,a,x)||}{||\mathcal{I}^i(n,\varepsilon,a,x)||} &< \epsilon
\end{aligned}
$$

if yes stop the entire algorithm, otherwise set $i = i + 1$ and return to **Step 1.**

C Estimation procedure

C.1 Targets

The objective is to replicate the following moments:

(A) The share of firms by age and size of the firm

(B) The share of employment by age and size of the firm

(C) The job creation rate, the job destruction rate, and the exit rate by age and size of the firm

(D) The elasticity of employment with respect to the corporate tax by age and size of the firm

For moments (A) to (C), we have 4 size groups (1-19 , 20-99, 100-499, 500+ employees) and 5 age groups (0-1, 2-3, 4-5, 6-10, 11+ years old). For moments in (D), we consider only two firm size groups: less than 500 employees (small), and 500 employees or more (large). For (D), we also consider only two age groups: less than 6 years (young), and 6 years or more (old). In total, we have 49 moments.

C.2 Estimated parameters

We estimate the remaining parameters using a simulated method of moments. We have five parameters that depend on the technology level x and eight parameters that are common across technology groups. We thus have 23 parameters to estimate. The set of structural parameters is given by:

$$\Theta = \{c_v(x), \bar{y}(x), \psi_0(x), \bar{e}(x), \eta_2(x), \phi_v, \gamma, \rho_\varepsilon, \sigma_\varepsilon, \alpha, \Delta_o, \psi_1, \psi_2\}$$

C.3 Minimization problem

The optimization procedure consists of finding the vector of parameters Θ that minimizes the distance between the vector of moments simulated from the model $\mathcal{M}^m(\Theta)$ and the vector of moments from the data $\mathcal{M}^d$. The SMM estimator $\hat{\Theta}$ solves:

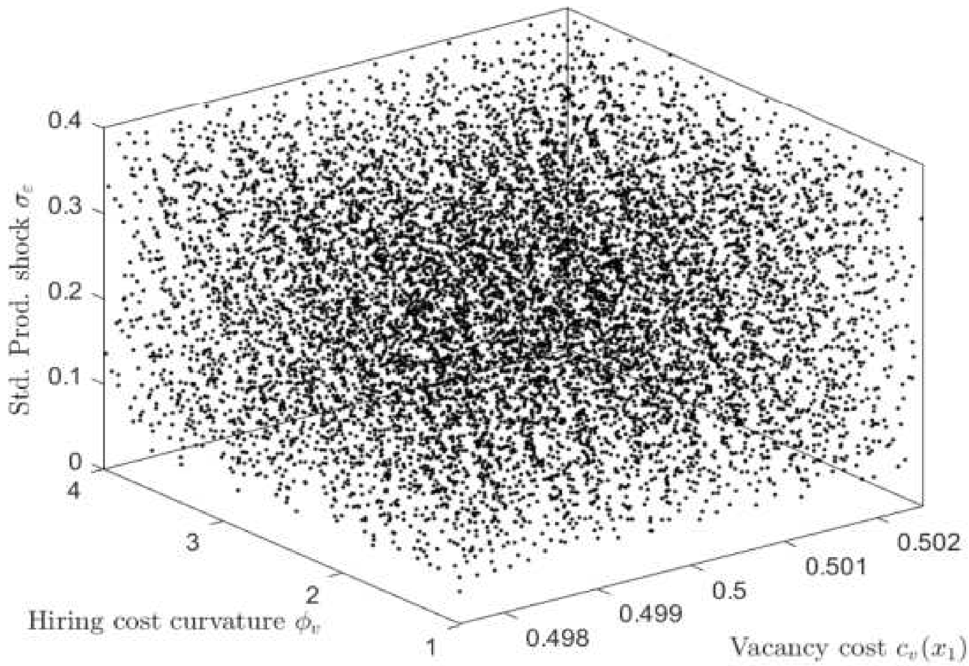
$$\hat{\Theta} = \arg\min_{\Theta} \left[\mathcal{M}^m(\Theta) - \mathcal{M}^d\right]' \boldsymbol{W} \left[\mathcal{M}^m(\Theta) - \mathcal{M}^d\right],$$

where $\boldsymbol{W}$ is a weighting matrix. It is considered as a identity matrix.

C.4 Adaptive grid

We use the adaptive grid method to find Θ similar to Albertini et al. (2020). The procedure builds a wide multidimensional grid covering the space of the parameter values. The grid is automatically refined so as to minimize the residual function. The grid is initially built using Halton low-discrepancy sequences to generate points in space in a highly uniform manner projected inside an hyper cube. A three dimensional example (out of the 23 dimensions) is provided in Figure 17.

Figure 17: SIMULATED GRID FOR ESTIMATED PARAMETERS



We use the following notation:

- Denote by $\Theta = \{\theta_1, \theta_2, ..., \theta_D\}$ the set of parameters with $D = dim(\Theta) = 23$.
- $\underline{\Theta} = \{\underline{\theta}_1, \underline{\theta}_2, ..., \underline{\theta}_D\}$ is the lower bound of the parameters.
- $\overline{\Theta} = \{\overline{\theta}_1, \overline{\theta}_2, ..., \overline{\theta}_D\}$ is the upper bound of the parameters.
- $x(n) = (x_{p_1}(n), x_{p_2}(n), ..., x_{p_D}(n))$ is the Halton sequence of length n in a D dimensional space.

- The grid over each parameter writes:

$$\mathcal{G}(n) = \{\varphi(\theta_1), \varphi(\theta_2), ..., \varphi(\theta_D)\}$$

where,

$$\varphi(\theta_d) = \underline{\theta}_d + x_{p_d}(n)(\overline{\theta}_d - \underline{\theta}_d)$$

C.5 Algorithm

Step 1: Set iteration number $k = 1$ and set initial value of the bounds $\underline{\Theta}$ and $\overline{\Theta}$ for the parameters. Compute the multidimensional grid $\mathcal{G}(n)$ using Halton sequence. Set the refinement parameter $\sigma^k < 1$ and the initial value of the residual function $\mathcal{R}^k$.

Step 2: Solve the model for each combination of parameters and compute the residual function for each parameter combination

$$\mathcal{R}^k(\mathcal{G}(n)) = [\mathcal{M}^m(\mathcal{G}(n)) - \mathcal{M}^d]'\mathbf{W}[\mathcal{M}^m(\mathcal{G}(n)) - \mathcal{M}^d]$$

Step 3: Find the node $n^\star$ from $\mathcal{G}(n)$ that minimizes the residual function.

Step 4: Compute the difference between the residual function at k and $k+1$:

$$\Delta^{k+1} = \frac{||\mathcal{R}(\mathcal{G}(n^{\star k+1}))^{k+1} - \mathcal{R}(\mathcal{G}(n^{\star k}))^k||}{||\mathcal{R}(\mathcal{G}(n^{\star k}))^k||}$$

Step 5: If $\Delta^{k+1} \leq \tilde{\varepsilon}$ stop the algorithm. Otherwise, refine the grid of the parameters using the following updating scheme:

a. Set $k = k + 1$.

b. Decrease the grid dispersion $\sigma^k = \varpi\sigma^{k-1}$, $\varpi < 1$

c. Compute the new lower band and upper band of estimated parameters : $\underline{\Theta} = (1 - \sigma^k)\mathcal{G}(n^{\star k-1})$ and $\overline{\Theta} = (1 + \sigma^k)\mathcal{G}(n^{\star k-1})$

d. Compute new Halton sequences for each parameter

$$\varphi(\theta_d) = \underline{\theta}_d + x_{p_d}(n)(\overline{\theta}_d - \underline{\theta}_d)$$

e. Compute the multidimensional grid as:

$$\mathcal{G}(n) = \{\varphi(\theta_1), \varphi(\theta_2), ..., \varphi(\theta_D)\}$$

f. Return to **Step 2.**

D Optimal policy

D.1 Optimization program

The optimal tax schedule (based on age and size of firms) maximizes the welfare under a constraint on the fiscal surplus. We consider the following functional form for the tax rate:

$$\tau(n,a) = \frac{\xi_1}{1 + \xi_2 e^{-\xi_3 n - \xi_4 a}}$$

The optimization program consists in finding the coefficients $\xi_i \in \boldsymbol{\xi}$, $i = 1, ..., 4$ maximizing the welfare:

$$\begin{aligned} \xi &= \arg\max \mathcal{W}(\boldsymbol{\xi}) \\ & \quad s.t. \;\; FS(\boldsymbol{\xi}) \geq FS(\boldsymbol{\xi}_{bench.}) \end{aligned}$$

where FS is defined by equation (67).

D.2 Welfare

In order to simplify the exposition, we use the following variables.

- Total wage:

$$\boldsymbol{w} = \sum_x \sum_a \int_n \int_\varepsilon \lambda(n, \varepsilon, a, x) \times n \times w(\varepsilon, x) d\varepsilon dn$$

- Total transfers to non-employed

$$\boldsymbol{b} = \sum_x b(x) u(x)$$

- Total profits net of hiring costs

$$\begin{aligned} \boldsymbol{\pi} &= \sum_x \sum_a \int_n \int_\varepsilon \lambda(n, e, a, x) \left[\pi(n, \varepsilon, x) - c_v(x) v^{\phi_v} - c_d \ell\right] d\varepsilon dn \\ \text{with } v &= \mathcal{D}_v(n, \varepsilon, a, x) \\ \ell &= \mathcal{D}_\ell(n, \varepsilon, a, x) \end{aligned}$$

- Total hiring costs for new firms:

$$\begin{aligned} \boldsymbol{c}_{new} &= \sum_x e(x) c_v(x) v^{\phi_v} \\ \text{with } v &= \mathcal{D}_{0,v}(x) \end{aligned}$$

- Total operating costs:

$$\boldsymbol{c_o} = \sum_x \sum_a \int_n \int_\varepsilon \lambda(n, e, a, x) c_o(a, x) n d\varepsilon dn$$

- Total entry costs:

$$\boldsymbol{c_e} = \sum_x e(x) c_e(x)$$

- Total labor disutility

$$\boldsymbol{d_l} = \sum_x \sum_a \int_n \int_\varepsilon \lambda(n, e, a, x) n d\varepsilon dn$$

We consider linear utility. The welfare writes:

$$\begin{aligned} \mathcal{W} &= \sum_{j=0}^{\infty} \beta^j (C_{t+j} - \boldsymbol{d_l}) \\ \text{with } C &= \boldsymbol{w} + \boldsymbol{b} + T + \boldsymbol{\pi}(1-\tau) - \boldsymbol{c_o} - \boldsymbol{c}_{new} - \boldsymbol{c_e} \end{aligned}$$

which simplifies to:

$$C = \boldsymbol{y} - \boldsymbol{c_v} - \boldsymbol{c}_{new} - \boldsymbol{c_d} - \boldsymbol{c_o} - \boldsymbol{c_e}$$

where

$$\begin{aligned} \boldsymbol{y} &= \sum_x \sum_a \int_n \int_\varepsilon \lambda(n,e,a,x) y(n,\varepsilon,x) \, d\varepsilon dn \\ \boldsymbol{c_v} &= \sum_x \sum_a \int_n \int_\varepsilon \lambda(n,e,a,x) c_v(x) v^{\phi_v} d\varepsilon dn \\ \boldsymbol{c_d} &= \sum_x \sum_a \int_n \int_\varepsilon \lambda(n,e,a,x) c_d \ell d\varepsilon dn \end{aligned}$$

D.3 Algorithm

The algorithm is a root finding procedure similar to that employed for estimating the model's parameters.

Step 1: Build a multidimensional grid of dimension four for the parameters $\{\xi_1, \xi_2, \xi_3, \xi_4\}$ using Halton's low discrepancy sequences.

Step 2: Solve the model for each combination of parameters and compute the welfare $\mathcal{W}$.

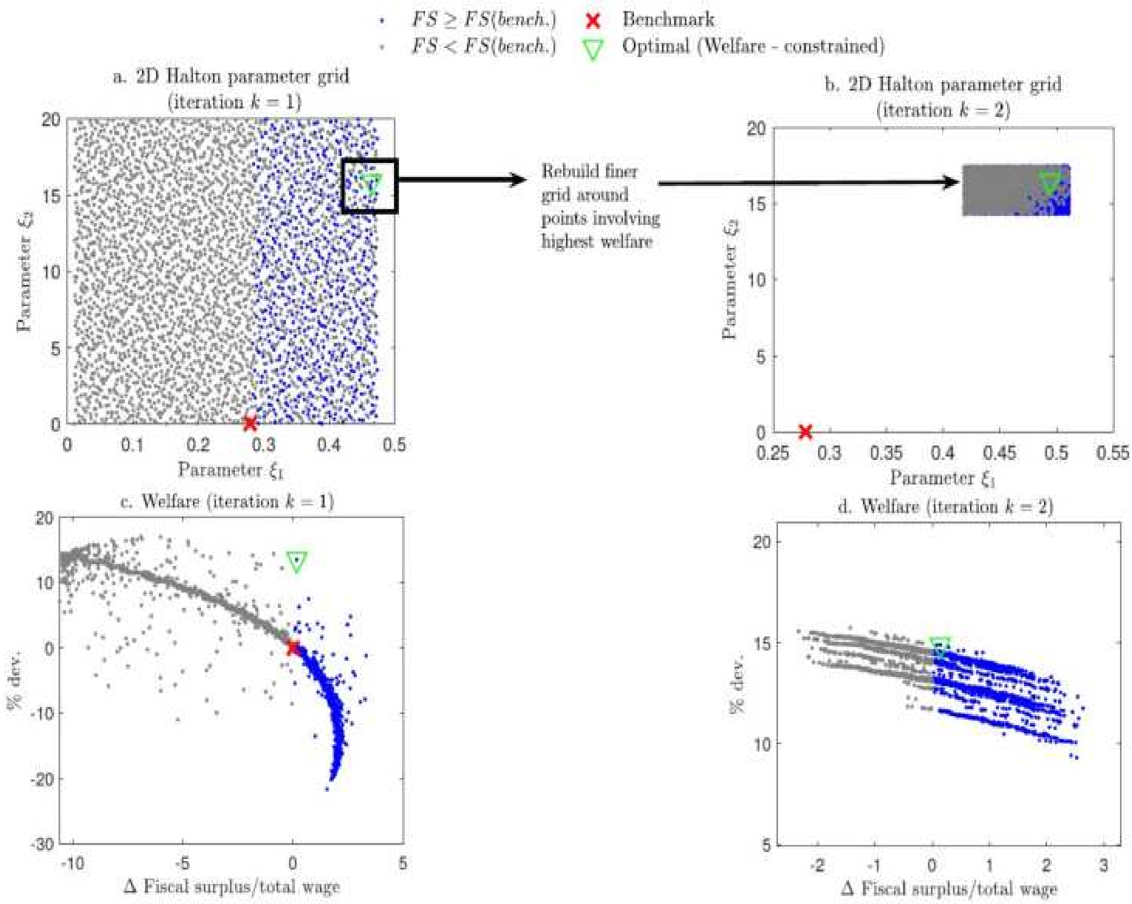
Step 3: Discard all simulations for which the constraints on the fiscal surplus is not respected, *i.e.* when the fiscal surplus is lower than that in the benchmark calibration.

Step 4: Find the combination of parameter that maximize the welfare $\mathcal{W}$.

Step 5: Recompute the grid around the ξ_i that have highest welfare level and go back to **Step 2.** Repeat this procedure until no improvement in welfare are possible, that is when the maximum welfare between iteration k and iteration $k+1$ is lower than a given threshold.

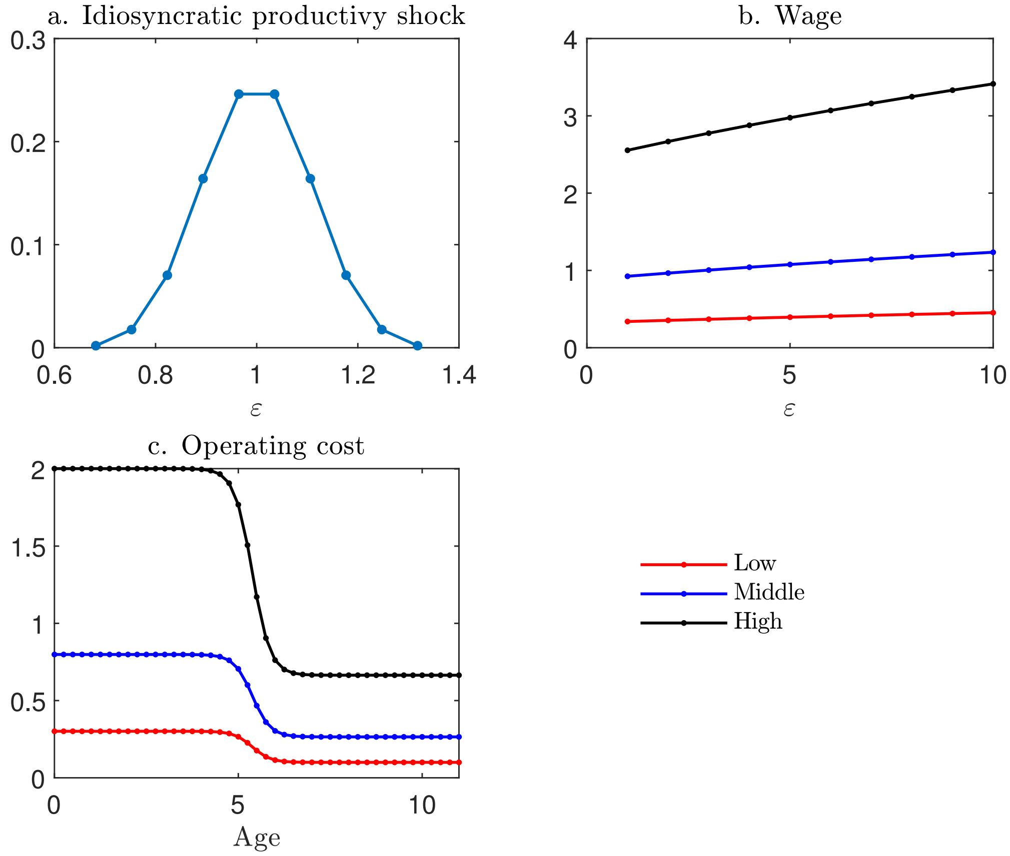
A simulation between two steps k is provided in Figure 18.

Figure 18: ALGORITHM WELFARE MAXIMIZATION



E Supplementary simulations

Figure 19: PRODUCTIVITY SHOCK, WAGE AND OPERATING COSTS



TEPP Working Papers 2024

24-7. Effects of a business support program on firm performances in France
Fabrice Gilles, Yannick L'Horty, Ferhat Mihoubi

24-6. Increased fine for repeat offenders and conglomerate dynamics
Armel Jacques

24-5. The valuation of energy efficiency labels in the French housing market
Sylvain Chareyron

24-4. A comprehensive analysis of production efficiency : a tax reform perspective
Laurence Jacquet, Etienne Lehmann

24-3. How to measure energy poverty in warm and cold climate territories? A multidimensional approach
Manitra Rakotomena, Olivia Ricci

24-2. Innovating for the good or for the bad. An EU-wide analysis of the impact of technological transformation on job polarisation and unemployment
Ylenia Curci, Nathalie Greenan, Silvia Napolitano

24-1. Is training helpful in boosting the self-confidence and professional integration of young people not in employment, education or training? Results from a randomized experiment
Nicolas Moreau, Alexis Parmentier, Mylène Lebon-Eyquem

The TEPP Institute

The CNRS **Institute for Theory and Evaluation of Public Policies** (the TEPP Institute, FR n°2024 CNRS) gathers together research centres specializing in economics and sociology:

- L'**Equipe de Recherche sur l'Utilisation des Données Individuelles en lien avec la Théorie Economique** (Research Team on Use of Individuals Data in connection with economic theory), **ERUDITE**, University of Paris-Est Créteil, University of Gustave Eiffel;
- Le **Centre d'Etudes des Politiques Economiques** (Research Centre focused on the analysis of economic policy and its foundations and implications), **EPEE**, University of Evry Paris-Saclay ;
- Le **Centre Pierre Naville** (Research on Work and Urban Policies), **CPN**, University of Evry Paris-Saclay
- Le **Groupe d'Analyse des Itinéraires et des Niveaux Salariaux** (Group on Analysis of Wage Levels and Trajectories), **GAINS**, Le Mans University
- Le **Centre de Recherches en Economie et en Management**, (Research centre in Economics and Management), **CREM**, University of Rennes 1, University of Caen Basse-Normandie ;
- Le **Groupe de Recherche ANgevin en Économie et Management** (Angevin Research Group in Economics and Management), **GRANEM**, University of Angers ;
- Le **Centre de Recherche en Economie et Droit** (Research centre in Economics and Law) **CRED**, University of Paris II Panthéon-Assas ;
- Le **Laboratoire d'Economie et de Management Nantes-Atlantique** (Laboratory of Economics and Management of Nantes-Atlantique) **LEMNA**, Nantes University ;
- Le **Laboratoire interdisciplinaire d'étude du politique Hannah Arendt – Paris-Est**, **LIPHA-PE**, University of Paris-Est Créteil and University of Gustave Eiffel ;
- Le **Centre d'Economie et de Management de l'Océan Indien**, **CEMOI**, University of La Réunion ;
- Le **Laboratoire d'économie de Poitiers**, **LéP**, University of Poitiers ;
- L'UMR **Structures et marchés agricoles, ressources et territoires**, **SMART**, INRAE, Agro Rennes-Angers Institute ;
- Le **Centre de recherche en économie et en droit sur le développement insulaire**, **CREDDI**, University of the Antilles.

TEPP brings together 230 teacher-researchers and 100 doctoral students. It is both one of the main academic operators in the evaluation of public policies in France, and the largest multidisciplinary federation of research on work and employment. It responds to the demand for impact assessment of social programs using advanced technologies combining theoretical and econometric modeling, qualitative research techniques and controlled experiences.