Labor Market Responses to Payroll Tax Reductions

Thomas Breda* Luke Haywood† Haomin Wang‡

January 23, 2018

Abstract

Payroll tax reductions are a popular tool to lower the minimum labor cost and encourage employment and job creation. Effects of these tax reductions go beyond the directly affected. A particular concern about such policies is that more productive jobs may be replaced with less productive ones. We examine payroll tax reductions using an equilibrium search-and-matching model estimated from the French administrative data. We find that lowering taxes on low-paid work induces low-productivity workers to enter the labor market and low-productivity firms to post more vacancies. These behaviors congest the labor market, resulting in lower employment among high-productivity workers and negative impacts on aggregate production. We find that, rather than reducing taxes for a wide range of jobs, restricting payroll tax reductions to minimum wage jobs helps low-wage workers, but the resulting congestion effect is also stronger. Taking this trade-off into account, we determine who should benefit from payroll reductions.

1 Introduction

Low-wage payroll tax reductions are a popular tool used in many countries; these policies lower labor costs for low-wage jobs and relax the constraint that statutory minimum wages impose on labor costs.[1] The primary goal of these policies is to boost job creation and expand the employment opportunities for low-skill workers. Despite a voluminous literature, evaluations of low-wage tax reductions have mainly focused on workers who are directly targeted.[2] Nevertheless, these policies

*Paris School of Economics, CNRS, and Institut des Politiques Publiques. Email: thomas.breda@ens.fr.
†DIW Berlin. Email: lhaywood@diw.de.
‡DIW Berlin. Email: hwang@diw.de.

[1] Labor cost is defined as the total expenditure borne by employers in order to employ workers.

[2] Most studies in the empirical literature of tax reduction evaluations ignore the tax incidence on firms by assuming a perfectly elastic labor demand. This has led Hotz and Scholz (2003) and Nichols and Rothstein (2015) to conclude that the incidence of the Earned Income Tax credit (the US payroll tax reduction program) is not well-studied. This also applies to studies of payroll tax reduction policies in other countries.

have potential spill-over effects on the employment opportunities of workers with higher skill levels. Moreover, a change in the skill composition among employed workers can affect aggregate production of an economy. In this paper, we ask two questions. First, to what extent do spill-over effects impact employment and production? Second, what are the implications on the design of a tax reduction policy?

To answer these questions, we consider an equilibrium model based on Cahuc et al. (2006) and incorporate a non-linear payroll tax schedule and a statutory minimum net wage into our model. We assume that the labor market is frictional and workers and firms each have differing levels of productivity. We consider bilateral matches between workers and firms, and assume that the wage of a match is determined via bargaining over the net-of-tax match surplus, which is the present after-tax value of the match net of the outside options of the two parties. A match is viable if the net surplus is viable at a wage level that is at least equal to the minimum wage; otherwise, the worker and the firm remain unmatched.[3]

The key feature of our model that generates spill-over effects is that all workers and firms are in competition for the opportunities to meet a potential employer or employee, but not all of the matches are viable. A low-wage payroll tax reduction may expand the set of viable matches for low-productivity workers and firms, drawing them into the labor market. They may create congestions in the frictional labor market. In particular, if the rate of meeting a potential employer decreases with the number of job seekers, the increased participation of low-productivity workers would negatively affect the rate that more productive workers meet potential employers. Moreover, if high-productivity workers have good outside options, matching with low-productivity firms may be unviable. Therefore, the likelihood for them to find an acceptable employer decreases when low-productivity firms post more vacancies. The congestion effects have consequences for aggregate production. We allow for worker and firm productivities to be complementary in raising the level of production output. As a result, a lower number of matches between high-productivity workers and firms can disproportionally harm the overall level of production.

We apply our model to studying payroll tax reductions in France. France has one of the highest minimum wages among developed countries, which has been around 50-60% of the median wage since the 1990s.[4] The country also features a heavy payroll tax (known as Social Securities Contributions). In 2015, payroll taxes account for around 40-50% the labor cost for an average wage earner.[5] More notably, payroll taxes on low-paid work are substantial. In the early 1990s, payroll taxes account

[3]Our assumption regarding the minimum wage follows Flinn (2006) and Mabli and Flinn (2007). While Cahuc et al. (2006) also consider a minimum wage in their appendix, they assume that workers may use the minimum wage as their outside option in their bargaining with the employer, which may be a problematic assumption given high unemployment rates.

[4]See https://stats.oecd.org/Index.aspx?DataSetCode=RMW#.

[5]See OECD (2016).

for over 40% of the labor cost for workers in the bottom labor-cost decile.[6] The high minimum labor cost has been blamed for low employment levels in France, and it has motivated a series of reforms to reduce taxes for low-wage jobs.

We estimate our model using the Simulated Method of Moments based on data from the Annual Declaration of Social Data (DADS). We focus on men of age 30-55 who are primarily employed in full-time, public sector, non-executive jobs. The greatest challenge is to construct statistics that inform us about the structural parameters that govern the productivity distributions of workers and firms and the degree of production complementarity. The difficulty comes from the fact that we do not observe the productivity levels or ranks as there is not a monotone relationship between wage and productivity. Nevertheless, our model informs us about statistics that are consistent with productivities, and our data allow us to use these statistics to estimate the ranks of individuals and firms. We then use the estimated ranks to construct moment statistics to recover the structural productivity distribution and production parameters.

Using the estimated model, we first consider payroll tax reductions in France between 1995 and 1997. Our simulation indicates that, in equilibrium, employment increases by 2.5% and aggregate production increases by 1.2%. These results are in line with the findings of Crépon and Desplatz (2003) and Chéron et al. (2008). By examining the effects on workers of different levels of productivity, we find that, despite an overall positive employment effect, the employment rate declines by 0.6% among workers in the top productivity quartile. This is mainly driven by the congestion effects: there are a 2.1% increase in labor force participation from low-productivity workers and a 4.8% increase in vacancies from low-productivity firms. If high-productivity individuals were not impacted, aggregate production would have increased by 1.4% instead of 1.2%.

We then study whether an optimal payroll tax reduction program reduces taxes only for minimum wage jobs or whether individuals with higher wages should also pay lower taxes. In particular, we focus on a family of tax reduction programs that grants the highest tax reduction to minimum wage jobs. We make different programs budget neutral by redistributing any surplus as a lump-sum transfer, and restrict all programs to raise employment to a fixed level. The trade-off between aggregate production and redistribution is evident in our simulations. While a relatively narrow tax reduction is more effective at making low-productivity individuals better-off in terms of consumption and employment, it results in lower aggregate production. A tax reduction that is too narrowly focused can even result in a lower level of aggregate production than the baseline despite the employment increase. Adopting a social welfare criterion that allows the policy maker to care more about the less well-off than others, we find that the optimal tax reduction to raise employment by

[6] Authors' calculation based on data from the Annual Declaration of Social Data (DADS) and computations of payroll taxes by the tax simulator TAXIPP. The calculation account for jobs that are full-time, non-executive and private-sector.

5% should have a broad coverage - jobs with a wage less than twice the minimum wage should benefit from the tax reduction.

Our main contribution is to the literature that studies payroll tax reductions by quantifying spill-over effects due to congestion. Based on reduced-form estimates, Rothstein (2008) and Leigh (2010) find that the Earned Income Tax Credit (EITC), a tax reduction program that targets low-income working individuals in the U.S., negatively affects wages for low-productivity workers who do not participate in the program. Azmat (2006) finds similar spill-over effects from the Working Family Tax Credit (the UK payroll tax reduction program). The reduced-form evidence suggests that taking an equilibrium approach to study payroll tax reductions can be fruitful.

There is a small but growing literature that studies labor market policies using equilibrium frameworks (Chéron et al., 2008; Shephard, 2016; Moser and Engbom, 2016). Chéron et al. (2008) and Shephard (2016) examine payroll-tax reduction policies in France and in the U.K., respectively. However, both of their papers assume that individuals are homogenous in the productivity level. We complement their papers by considering heterogeneity in worker productivity and complementarity in the production technology. These are essential features of our model that lead to substantial congestion effects. In addition, their papers do not consider the redistributive effects of a tax policy, which is another result of homogeneity in worker productivity. Our paper also differs from Shephard (2016) in that he studies tax policies in the U.K. during a period with no minimum wage constraint.

Methodologically, although our model is based on Cahuc et al. (2006) in that wages are determined through a combination of employer-employee bargaining and competition between incumbent and poaching firms, there are significant distinctions since we consider a non-linear payroll tax and a statutory minimum wage.[7] The presence of payroll taxes implies that utility is not directly transferrable between workers and firms, and therefore the net match surplus is not known until it is shared. Together with the minimum wage, the inclusion of the two labor market institutions poses challenges as the equilibrium must be solved numerically.[8]

Our finding that a low-wage payroll tax reduction substitutes employment opportunities for high-productivity workers with those for low-productivity workers is supported by empirical evidence. For example, Crépon and Desplatz (2003) examine the effects of French payroll tax reduction in the mid-1990s on firm-level decisions, and find that it leads to lower average labor cost, higher shares of unskilled workers, and a decrease in the productivity of capital and labor in most sectors. Since changing payroll taxes on low-wage work changes the minimum labor cost constraint, a change in

[7] Similar wage determination mechanism (without taxes and a minimum wage) is first considered in Dey and Flinn (2005), and has been a common assumption in the search-and-matching literature (Bagger and Lentz, 2014; Lise et al., 2013).

[8] This is in contrast to Cahuc et al. (2006), in which wage equations can be derived using the match surplus function.

the minimum wage should have similar effects as a change in taxes. Recent empirical studies have confirmed that higher minimum wages lead to increases in labor productivity in firms that employ low-wage workers (Mayneris et al., 2016; Riley and Rosazza Bondibene, 2017).

Our study of a policy that reduces the minimum labor cost has implications beyond France. A binding minimum labor cost has become increasingly prevalent worldwide. The main contributor for such a constraint is a statutory minimum wage. Germany introduced its first Federal minimum wage in 2015. In 2016, the U.K. has started to implement a gradual increase of its minimum wage by 40% over 5 years. In the U.S., states and municipalities have imposed local minimum wages that are substantially higher than the federal one. Seattle is one example. The city has raised its minimum wage from $9.47 to $13 in 2016, and will continue to raise it to $15 by 2021.[9] In addition to the minimum wage, an increasing taste for non-working time may also lead to higher minimum labor costs. For example, Aguiar et al. (2017) find that improvements in leisure technologies such as video games lead to an increased demand for leisure among young men, which have caused a decline in their labor force participation.

Section 2 characterizes the equilibrium search and matching model. We describe the data in Section 3. Section 4 presents our estimation strategy and results. In Section 5, we apply our estimated model to evaluate the equilibrium effects of the payroll tax reductions implemented between 1995 and 1997. In Section 6, we compare the equilibrium effects of different tax reduction coverages and simulate the welfare implication. Finally, we conclude in Section 7.

2 Model

In this section, we present an equilibrium search-and-matching model with a wage-determination mechanism akin to Cahuc et al. (2006) but including a statutory minimum wage and a nonlinear tax schedule.

2.1 Environment

Time is continuous and the planning horizon is infinite. All agents discount the future at rate r. There is a unit measure of risk-neutral workers who are heterogenous in productivity. The productivity of a worker is drawn from an exogenous distribution. Let x be the type, or rank, of a worker according to her productivity, which is uniformly distributed in the interval $[0, 1]$, and let $\mathfrak{h}(x)$ be her productivity with $\mathfrak{h}'(x) > 0$. Non-employed individuals can choose to participate in job

[9] For an analysis of the Seattle minimum wage, see Jardim et al. (2017).

search by paying a flow cost of q. If they do so, they are considered as being unemployed; those who do not search are non-participants. The search cost captures the difference between the discomfort of search and the stigma of not looking for jobs. An employed worker receives a wage, which is determined endogenously. A non-employed individual of type x receives flow benefit $B(x)$, which benefit is increasing in worker productivity.[10]

We do not assume a search cost for employed workers; the difference between on- and off-the-job search is captured by the difference in the search efficiency. We normalize search efficiency in unemployment to 1, and denote search efficiency on-the-job by s_1.[11] Let E, U, and N denote the measures of the workers in the three states respectively, with $E + U + N = 1$. Let $u(x)$ denote the measure of unemployed workers of type x such that $\int_0^1 u(x)dx = U$. The sum of search intensity is denoted by $\xi = U + s_1 E$.

There is a continuum of firms whose productivity levels are also drawn from an exogenous distribution. We define y as the type, or rank, of a firm according to its productivity, and $\mathfrak{p}(y)$ as its productivity level. y is uniformly distributed over some interval $[0, y_h]$, where y_h is an exogenous parameter. Each firm is a collection of one-worker jobs of the same productivity level. Total production of a firm is the sum of production of its jobs. Each job determines on its own whether to purchase the opportunity to post a vacancy from a competitive market for type-y vacancies. The measure of type-y vacancies, $v(y)$, is determined by the free entry condition such that the price per unit of vacancy $c(v(y))$ is equal to the expected return of a vacancy in equilibrium. We take $c(v(y))$ to be exogenous, positive and strictly increasing, i.e. $c(\cdot) > 0$ and $c'(\cdot) > 0$. The total measure of vacancies is $V = \int_0^{y_h} v(y)dy$. Note that it is possible that some firm y has no vacancies; the firm is referred to as being inactive. Let y_l denote the rank of the least productive firm, which is endogenously determined in equilibrium.

Search is random. The measure of realized meetings depends on the sum of search intensity ξ and the sum of vacancies V. Let $M(\xi, V)$ be a constant-return to scale meeting technology that maps ξ and V to the measure of meetings. For convenience, we define $\kappa \equiv \frac{M(\xi,V)}{\xi V}$ and omit its dependence on ξ and V. κV is the flow measures of matches per unit of search efficiency, and $\kappa\xi$ is the flow measure of matches per vacancy.

The rates at which a vacancy meets an unemployed and an employed worker of rank x is, respectively. The measure of filled jobs, or matches, of type (x, y) is given by $h(x, y)$, and the total measure of

[10] As we will explain later, we consider $B(x)$ as the non-employment transfers which in practice is linked to the individual's previous wage. It is thus plausible to have $B' > 0$.

[11] On-the-job search can be seen as a passive search: while unemployed workers have to search actively to meet potential employers, employed workers face a positive meeting rate without explicit efforts. Allowing employed workers to choose search intensity is an additional source of assortative matching (see Bagger and Lentz, 2014), which will affect the estimated decomposition of wage dispersion. However, since our focus is not on job-to-job transition or wage dispersion, this is beyond the scope of this paper.

employed workers is $E = \int_0^{y_h} \int_0^1 h(x,y)dxdy$.

When a vacancy is filled, the firm collects the production output $F(x,y)$ and pays the net wage w and taxes $T(w)$. We assume that the production output is increasing in both worker and firm productivities: $F_x(x,y) > 0$ and $F_y(x,y) > 0$ for all x and y. Moreover, we make the free-disposal assumption: jobs can freely dispose any level of output and mimic a less-productive job.

A match may be destroyed by nature at rate δ, or destroyed endogenously if a worker transitions to another job.

2.2 Policies

We consider two labor market policies. The first is a statutory floor on the net wage, w_{min}. The second is a payroll tax $T(w)$ collected on the net wage w. We assume that $T(\cdot)$ is differentiable and increasing in w. However, all equilibrium results would remain the same if we only assumed that $T(\cdot)$ is continuously increasing. This accommodates a wide range of tax functions, particularly ones with non-monotone marginal tax rates.

We model the policies in terms of the net wage because this translates immediately to workers' individual utility. Nominally, the payroll tax in France is shared between employers and employees. However, the tax incidence does not affect the equilibrium outcomes.

2.3 Meeting between an unemployed worker and a job

When an unemployed worker x meets a vacancy y, the pair first determines a bargained wage ϕ by splitting the net-of-tax match surplus. Let $W_n(x)$ and $J_u(y)$ denote, respectively, the present values of non-employment and an unfilled vacancy, and let $W_e(w,x,y)$ and $J_f(w,x,y)$ denote, respectively, the present values of employment and a filled position when worker x is matched with firm y at wage w. All values are net-of-tax.

The net surplus of a match is therefore defined by

$$S(w,x,y) = W_e(w,x,y) - W_n(x) + J_f(w,x,y) - J_u(y) \tag{1}$$

Because taxes depend on wage, the exact value of the net surplus cannot be determined without knowing how the surplus is split.

We assume proportional bargaining such that the net surplus is split proportional to the bargaining powers of the worker and the firm, α and $1-\alpha$.[12] The bargained wage ϕ must also make the net surplus positive. Formally, ϕ must satisfy the following system:

$$\begin{cases} W_e(\phi, x, y) - W_n(x) = \alpha S(\phi, x, y) \\ S(\phi, x, y) \geq 0 \end{cases} \tag{2}$$

If such a ϕ exists, the pair proceeds to the second step, in which they compare ϕ to w_{min}. If $\phi \geq w_{min}$, a match is viable and carries out at wage ϕ. If $\phi < w_{min}$, a match is only viable if w_{min} is mutually agreeable, and the match is carried out at wage w_{min}. We define match viability in the following definition.

Definition 1. A match is viable if ϕ that solves Eq. 2 exists and either of the following holds:

1. $\phi \geq w_{min}$, or
2. $\phi < w_{min}$, $W_e(w_{min}, x, y) - W_n(x) \geq 0$ and $J_f(w_{min}, x, y) - J_u(y) \geq 0$.

Let $\phi_u(x, y) = \max\{\phi, w_{min}\}$ be the match wage of viable matches, and let $\mathcal{A}_u(x) \subseteq [0, y_h]$ be the subset of firms with whom an unemployed worker x can form viable matches.

2.4 Meeting between an employed worker and an outside firm

Similar to Cahuc et al. (2006), when an employed worker x meets a poaching firm, the incumbent and the poaching firms compete over the worker, and a wage negoation takes place between the worker and the winner of the firm competition. In the competition step, firm y beats y' if and only if $W_e(\bar{\phi}(x, y), x, y) \geq W_e(\bar{\phi}(x, y'), x, y')$, where $\bar{\phi}(\cdot, \cdot)$ is the "best match wage". For a match (x, y), it is defined by:

$$\bar{\phi}(x, y) = \arg\max_{w} W_e(w, x, y)$$

such that

$$J_f(w, x, y) - J_u(y) \geq 0.$$

[12] We choose the proportional bargaining scheme rather than the Nash bargaining scheme because we wish to accommodate a flexible specification for the tax function that may contain kinks and non-monotonicity in the marginal tax rate. Under Nash bargaining, multiple equilibria may exist if the marginal tax rate is not continuously increasing. Nash and proportional bargaining yield the same results if the payroll tax does not depend on wage; with a tax that depend on labor income, match surplus is not perfectly transferrable between workers and firms and thus the two schemes diverge (l'Haridon et al., 2013; Jacquet et al., 2014). See Appendix A for more details.

Wage bargaining is only initiated if the losing firm y' can make the worker better-off compared to her current state, i.e. $W_e(\bar{\phi}(x,y'),x,y') \geq W_e(w_0,x,y_0)$, where a subscript 0 indicates the incumbent firm. If wage bargaining is not initiated, the worker remains in firm y_0 at wage w_0. Otherwise, the worker chooses to work at the winning firm y and use $W_e(\bar{\phi}(x,y'),x,y')$ as her new outside option in bargaining with the firm. The modified net surplus is

$$S_e(w,x,y,y') = W_e(w,x,y) - W_e(\bar{\phi}(x,y'),x,y') + J_f(w,x,y) - J_u(y) \tag{3}$$

The bargained wage ϕ must solve the following system:

$$\begin{cases} W_e(\phi,x,y) - W_e(\bar{\phi}(x,y'),x,y') = \alpha S_e(\phi,x,y,y') \\ S_e(\phi,x,y,y') \geq 0 \end{cases} \tag{4}$$

If ϕ exists, the match wage is $\phi_e(x,y,y') = \max\{\phi, w_{min}\}$.[13] If the poaching firm wins the competition, the worker makes a job-to-job transition.

We define two useful sets. First, let $\mathcal{A}_{e1}(x,y_0) \subseteq [0,y_h]$ be the subset of firms that can poach the worker from a match (x,y_0). Then, let $\mathcal{A}_{e2}(w_0,x,y_0) \subseteq [0,y_h]$ be the subset of firms not in the subset $\mathcal{A}_{e1}(x,y_0)$ that can make worker x strictly better-off than her current state. This is the subset of firms that can trigger a wage negotiation for the match (x,y_0).

2.5 Value functions

Given the match wages ϕ_u and ϕ_e and the sets $\mathcal{A}_u$, $\mathcal{A}_{e1}$, and $\mathcal{A}_{e2}$, we can characterize the value functions.

2.5.1 Value of Non-employment

A non-employed worker receives a flow income $B(x)$ regardless of her job search decision. The decision to search is time-invariant because it only depends on the permanent type x of the worker. The present value of non-participation is

$$rW_{np}(x) = B(x)$$

The present value of unemployment is defined by

$$rW_u(x) = B(x) - q + \kappa \int_{y' \in \mathcal{A}_u(x)} [W_e(\phi_u(x,y'),x,y') - W_n(x)]v(y')dy'$$

[13] We show later that the winning firm can always afford w_{min}.

A worker chooses whether or not to search in order to maximize her present value of non-employment $W_n(x)$:

$$W_n(x) = \max\{W_{np}(x), W_u(x)\} \tag{5}$$

Let $s(x)$ denote the optimal search decision of a non-employed worker, which is a binary choice such that $s(x) = 1$ indicates unemployment and $s(x) = 0$ indicates non-participation. We have $s(x) = 1$ if and only if

$$q \leq \kappa \int_{y' \in \mathcal{A}_u(x)} [W_e(\phi_u(x, y'), x, y') - W_n(x)] v(y') dy'$$

Note that we have ruled out mixed strategies because each worker type x is atomless and thus the search decision $s(x)$ does not influence κ.

2.5.2 Values of employment and a filled job

In a match, the employed worker receives a flow income w, and the firm receives a flow income of $F(x, y) - w - T(w)$. The worker may either become separated or meets a poaching firm. Given the latter, one of three events must happen. First, the worker moves to the poaching firm and negotiates the wage with it. Second, the worker stays and renegotiates wage with the incumbent firm. Third, the worker stays at the same wage. The present value function of employment is defined by

$$\begin{aligned} [r + \delta + s_1 \kappa V] W_e(w, x, y) &= w + \delta W_n(x) \\ &+ s_1 \kappa \int_{y' \in \mathcal{A}_{e1}(x,y)} W_e(\phi_e(x, y', y), x, y') v(y') dy' \\ &+ s_1 \kappa \int_{y' \in \mathcal{A}_{e2}(w,x,y)} W_e(\phi_e(x, y, y'), x, y) v(y') dy' \\ &+ s_1 \kappa \int_{y \in [0, y_h] \setminus \{\mathcal{A}_{e1}(x,y) \cup \mathcal{A}_{e2}(w,x,y)\}} W_e(w, x, y) v(y') dy' \end{aligned} \tag{6}$$

The present value of a filled job is defined by

$$\begin{aligned} &(r + \delta + s_1 \kappa V) J_f(w, x, y) = F(x, y) - w - T(w) + \delta J_u(y) \\ &+ s_1 \kappa \int_{y' \in \mathcal{A}_{e1}(x,y)} J_u(y) v(y') dy' \\ &+ s_1 \kappa \int_{y' \in \mathcal{A}_{e2}(w,x,y)} J_f(\phi_e(x, y', y), x, y) v(y') dy' \\ &+ s_1 \kappa \int_{y' \in [y_l, y_h] \setminus \{\mathcal{A}_{e1}(x,y) \cup \mathcal{A}_{e2}(w,x,y)\}} J_f(w, x, y) v(y') dy' \end{aligned} \tag{7}$$

We can show that the values of employment and a filled position both monotonically increase with wage. Given this result, the bargained wage must be unique. These are summarized in the following proposition and corollary.

Proposition 2. *For all $w \geq w_{min}$, x and y, $\partial W_e(w,x,y)/\partial w > 0$.* Moreover, $\lim_{w\to-\infty} W_e(w,x,y) = -\infty$ and $\lim_{w\to\infty} W_e(w,x,y) = \infty$.

For all x and y, $\frac{\partial J_f(w,x,y)}{\partial w} < 0$. Moreover, $\lim_{w\to-\infty} J_f(w,x,y) = +\infty$ and $\lim_{w\to\infty} J_f(w,x,y) = -\infty$.

Corollary 3. *The bargained wages satisfying the systems 2 and 4 are unique if they exist.*

Heuristic proofs can be found in Appendix B.

2.5.3 Vacancy

A vacant job pays a flow cost $c(v(y))$, which depends on the total measure of vacancies of type y. The job can either match with an unemployed worker or poach a worker from an existing match. Let $\mathcal{B}_u(y) = \{x : s(x) = 1 \text{ and } y \in \mathcal{A}_u(x)\}$ be the set of unemployed worker x with whom firm y can form a match, and let $\mathcal{B}_e(y) = \{(x,y') : s(x) = 1 \text{ and } y \in \mathcal{A}_{e1}(x,y')\}$ be the set of matches from which firm y can successfully poach. The present value of a vacancy is defined by

$$\begin{aligned} rJ_u(y) &= -c(v(y)) \\ &+\kappa \int_{x\in\mathcal{B}_u(y)} [J_f(\phi_u(x,y),x,y) - J_u(y)]u(x)dx \\ &+\kappa \iint_{(x,y')\in\mathcal{B}_e(y)} s_1[J_f(\phi_e(x,y,y'),x,y) - J_u(y)]h(x,y')dy'dx \end{aligned} \tag{8}$$

2.6 Equilibrium

2.6.1 Free-entry condition

We assume a competitive market for opportunities to post each type of vacancy y, in which $c(v(y))$ is the price of a vacancy. The market entry decision is made independent in each job. In equilibrium, the free-entry condition holds, so that the present value of a vacancy is equal to zero. For any give

y, if $\mathcal{B}_u(y) = \emptyset$ and $\mathcal{B}_e(y) = \emptyset$, we have $v(y) = 0$, in which case the firm y is inactive. Otherwise, we have

$$\begin{aligned} c(v(y)) = & \\ +\kappa \int_{x \in \mathcal{B}_u(y)} & [J_f(\phi_u(x,y), x, y) - J_u(y)]u(x)dx \\ +\kappa \iint_{(x,y') \in \mathcal{B}_e(y)} & s_1 [J_f(\phi_e(x,y,y'), x, y) - J_u(y)]h(x,y')dy'dx \end{aligned} \tag{9}$$

which gives the vacancy distribution.

2.6.2 Characterizations of $\mathcal{A}_u$, $\mathcal{A}_{e1}$, and $\mathcal{A}_{e2}$.

Given our assumptions that $\frac{\partial F(x,y)}{\partial y} > 0$ and $T'(\bar{\phi}(x,y)) \geq 0$, we can show that the best match wage $\bar{\phi}$ increases in the firm type.

Proposition 4. *For all x, y, $\partial \bar{\phi}(x,y)/\partial y > 0$.*

We prove the above proposition in Appendix B.

In a competition between two firms, the firm that can offer a higher $\bar{\phi}(x,y)$ is the winner. This is because, at $\bar{\phi}(x,y)$, the only possible events that may affect the value of employment are an exogenous separation or a job-to-job transition. The former is independent of y. As for the latter, the worker always gets a higher value when moving from a firm that can offer a higher $\bar{\phi}(x,y)$ because of the free-disposal condition. The free-disposal condition implies that the winning firm can always afford to pay any wage firm y' pays to the worker, so that it can mimic the losing firm in future firm competitions and negotiations. Because of Proposition 4, we have that firm y beats y' if and only if $y > y'$.

The free-disposal condition also implies that whenever the match (x, y') is viable, the match (x, y) such that $y > y'$ must also be viable. This allows us to define $\mathcal{A}_u(x)$ as an interval such that $\mathcal{A}_u(x) = [\underline{y}(x), y_h]$. The lower bound arises from two constraints. The first constraint is $y_u(x)$, which is the lowest y such that the system in Eq. 2 is satisfied.[14] The second constraint is $y_{min}(x)$, which

[14]It may be useful to consider the static counterparts of the constraint. In a static environment, the system of equations in 2 becomes

$$\begin{cases} \phi + B(x) = \alpha \left[F(x,y) - T(\phi) - B(x) \right] \\ F(x,y) - T(\phi) - B(x) \geq 0 \end{cases}$$

The static counterpart for the first constraint is $\arg\min_{y \in [y_l, y_h]} \{y : F(x,y) \geq B(x) + T(B(x))\}$, which states that the match productivity must be high enough to pay the worker's non-working income plus taxes. It is clear that higher taxes endanger match viability.

is the lowest y that can afford to pay at least the minimum wage to a worker x such that $y_{min}(x) = \arg\min_{y \in [y_l, y_h]} \{y : F(x,y) \geq w_{min} + T(w_{min})\}$. The lower bound of $\mathcal{A}_u(x)$ is the maximum of the two constraints: $\underline{y}(x) = \max\{y_u(x), y_{min}(x)\}$. From the two conditions, we can see that higher taxes and minimum wages endanger the viability of a match.

We can also define least productive active firm as

$$y_l = \arg\min_y \left\{ y : \exists x \text{ such that } s(x) = 1 \text{ and } y \geq \underline{y}(x) \right\}$$

Proposition 4 also implies that all workers make job-to-job transitions toward the firm with a higher y. This allows us to express the sets $\mathcal{A}_{e1}$ and $\mathcal{A}_{e2}$ as convex intervals on $[y_l, y_h]$ such that:

$$\mathcal{A}_{e1}(x,y) = \left\{ y' \in [y_l, y_h] : y' > y \right\} \tag{10}$$

$$\mathcal{A}_{e2}(w,x,y) = \left\{ y' \in [y_l, y_h] : y > y' > y_0(w,x,y) \right\} \tag{11}$$

where $y_0(w,x,y)$ is the lowest y' that can trigger a wage renegotiation, i.e. $W_e(y_0(w,x,y),x,y) = W_e(\bar{\phi}(x,y'),x,y')$.

2.6.3 Steady state conditions on the distribution of workers.

The steady state condition assumes that for each type of worker x, the distributions across labor force state (employment or non-employment), type of employer y, and wage are constant. If $s(x) = 0$, worker x is always out of the labor force, so that $h(x,y) = 0$ and $u(x) = 0$.

Otherwise, if $s(x) = 1$, the steady state levels of $h(x,y)$ and $u(x)$ are determined by equating inflows with outflows. The outflow from type (x,y) matches is due to exogenous separations at rate δ and voluntary job-to-job transitions. The inflow into type (x,y) matches is due to unemployed workers x matching with firm y and employed worker moving to firm y. The steady state condition for $h(x,y)$ is the following:

$$h(x,y) = \frac{v(y)[\kappa u(x) + s_1 \kappa \int_{\underline{y}(x)}^{y} h(x,y') dy']}{(\delta + s_1 \kappa \int_{y}^{y_h} v(y') dy')} \tag{12}$$

The outflow from unemployment is due to job finding. The inflow is due to exogenous separation. The stationary measure of employed workers $u(x)$ is given as follows

$$u(x) = \frac{\delta}{\delta + \kappa \int_{\underline{y}(x)}^{y_h} v(y')dy'} \tag{13}$$

The wage distribution is stationary if the distribution of the best past reference firm y' conditional on the type of match (x, y) is stationary. Let $\bar{G}(y'|x, y)$ represent the fraction of type (x, y) matches with a reference firm of type y' or higher, and let $G(0|x, y)$ denote the probability that the match is formed with a worker from unemployment. The outflow from a match (x, y) with a reference firm of at least type y' is due to either job separation or a meeting with a firm $y'' > y$. There are two types of inflows. The first is that a worker x who is matched with the firm of type between y' and y meets the firm y and makes a job-to-job transition. The second is that the worker x who is matched with the firm y and whose reference firm is of type less than y' meets the firm of type between y' and y. The steady state condition for $G(y'|x, y)$ is given by

$$G(y'|x, y)h(x, y) = \frac{h(x, y)\left[\delta + \int_{y}^{y_h} v(y'')dy''\right] - s_1 \kappa v(y) \int_{y'}^{y} h(x, y'')dy''}{\delta + \int_{y'}^{y_h} v(y'')dy''} \tag{14}$$

2.6.4 Steady State Equilibrium

The steady state equilibrium is characterized by the distributions $\{u(\cdot), h(\cdot, \cdot)\}$, and the decisions $\{s(\cdot), \underline{y}(\cdot), \phi_u(\cdot, \cdot), y_0(\cdot, \cdot, \cdot), \phi_e(\cdot, \cdot, \cdot), v(\cdot)\}$ such that

1. Non-employed workers choose the optimal search strategy $s(\cdot)$.
2. Unemployed workers choose the optimal threshold firm type $\underline{y}(\cdot)$.
3. Employed workers choose the optimal threshold firm type $y_0(\cdot, \cdot, \cdot)$ for wage renegotiation. In addition, they move to firms with a higher y.
4. $\phi_u(\cdot, \cdot)$ and $\phi_e(\cdot, \cdot, \cdot)$ specify the wages of matches.
5. There is free entry of firms such that Eq. 9 holds, and $v(\cdot)$ solves the equation.
6. Steady state conditions (Eq. 12, 13 and 14) hold.

Solving the model relies on iterating the value functions because the surplus function depends on wage due to the presence of labor market policies. The continuation value of the match surplus depends on wages of potential future matches, which can only be solved for if the value functions are known.[15] We describe the procedure to solve for the steady state equilibrium numerically in Appendix C.

3 Data

Our main source of data is the Annual Declaration of Social Data (DADS), a French administrative data maintained by the French National Statistical Institute (INSEE). The DADS is based on mandatory employer declarations of the earnings of employees who contribute to the social security system. Data in the DADS database are organized into several datasets with different sampling schemes and structures for data security reasons. We access three of them that are relevant for our empirical applications: *panel DADS, panel tous salariés,* and *fichier Postes*.

The first two, *panel DADS* and *panel tous salariés* are linked employer-employee datasets, containing employees who were born in Octobers of even-numbered years.[16] The *panel DADS* mainly covers non-public sector workers, while *panel tous salariés* mainly covers public sector workers. We construct our panel data by combining the two DADS panels and merging social security contributions for private sector employees computed by the tax simulator developed by the Institute for Public Policies (TAXIPP). While the panel data contains full employment information on individuals, the individual-level sampling scheme renders it insufficient for observing within-firm wage distributions, which are important for our estimation strategy. We thus compute firm-level wage variables using our third dataset *fichier Postes* (hereafter, POST). POST contains all workers in the employer declarations even though workers are not linked across different job spells. Moreover, firms can be linked between the panel and POST.

Both the panel and POST contain information on the gender and age of workers and characteristics of jobs. We keep men age 30-55, and only focus on those who work mainly in full-time non-executive jobs in the private sector. This relatively homogenous group faces similar social security tax schedules.[17]

Individuals may be missing from our data for one of several reasons: unemployment, self-employment, and non-participation. The period that a worker is not observed in the panel will be referred to as a gap spell. We impute the status of individuals in gap spells using the duration of the gap,

[15]Because of taxes and the minimum wage, we are unable to derive an analytical wage equation as in Cahuc et al. (2006).

[16]From 2002 onwards, the dataset contains those born in Octobers of each year.

[17]Individuals working primarily in full-time jobs have a relatively weaker taste for non-working time. This frees us from considering the within period decision of consumption and leisure. Although the part-time labor market are more likely low-wage, among men of age 30-55, as much as 84% of the jobs are full-time, and a substantial fraction of these jobs are low-wage.

age of the individual, and the type of job that follows the gap (public or private sector, industry, and occupation). The parameters of the imputation model are estimated based on a similar sample drawn from the French labor force survey *Enquête Emploi* (EE).[18] We furthermore impute non-employment incomes for gap spells by employing a formula that depends on the gross earnings in all jobs during the year that precedes the unemployment spell.[19]

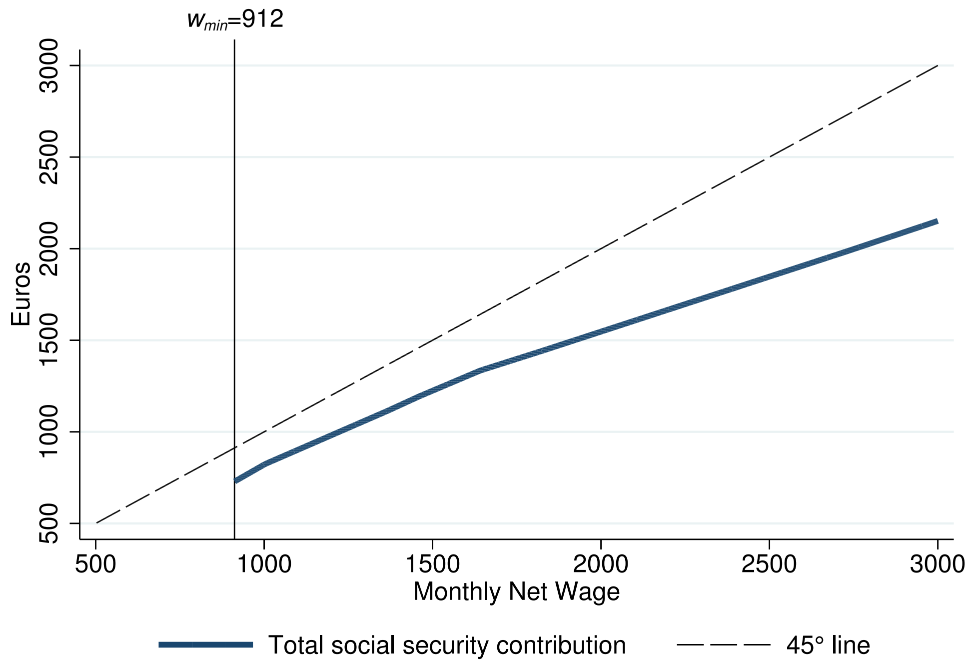
We took several steps to organize the panel data, which is based on employment and gap spells, into a monthly dataset that contains one observation per individual-month. The details of these steps are described in Appendix F. Wages are inflated to the 2010 price level. To convert between Euros and French francs, we apply the conversion rate of 1 € = 6.55957 FF.

We extract the aggregate labor market statistics used for model estimation from various sources other than the DADS. We do not compute the labor force participation rate and the unemployment rate using the DADS because the dataset does not include non-participants. Instead, we use the rates computed by INSEE using the Labour Force survey. Both rates are computed according to the International Labour Organization (ILO) definitions. According to the ILO, a labor force participant is defined as one who is either employed or unemployed, and an unemployed person is defined as a working age person who has not worked, is available for work, and has actively looked for a job in the previous month or will start a job in the next 3 months. We use the series for men age 25-49 in Metropolitan France. Our vacancy rate is taken from the Employment Orientation Board (*Conseil d'orientation pour l'emploi*).[20] The vacancy rate is defined as the number of vacancies divided by the sum of vacancies and jobs. We use the European vacancy definition, according to which a vacancy is a job to be filled immediately or at short notice, and there must be active search for candidates outside of the concerned firm. We consider non-public sector, non-agricultural vacancies in France.

Finally, we obtain a function the total social security contribution (SSC) in net wages by regressing the SSC on the net wage using a linear spline model. Since SSC is primarily based on individual wages rather than household income, this methods provides us with a good approximation for the SSC. We use data from the period between January 1993 to August 1995 for our steady state estimation because this is a period of relative stability in terms of minimum wage and tax policies. Figure 3.1 shows the approximated SSC for this period. The net statutory minimum wage for this period is 912 in 2010 Euros.

[18] See Appendix D for details on the imputation procedure.
[19] See Appendix E for details.
[20] See Conseil d'Orientation pour l'Emploi (2013).

Figure 3.1: Payroll tax by net wage, January 1993 to August 1995. The relationship is estimated based on a sample of non-executive, full-time, and private sector jobs from the DADS by regressing total payroll tax (social security contributions) on net wage using a linear spline model.



4 Estimation

We estimate the structural parameters of the model presented in Section 2 based on the data described in Section 3 using the Simulated Method of Moments.

4.1 Simulated method of moments

We estimate the structural parameters θ of our model using the Simulated Method of Moments (SMM), which involves finding the best parameters to match a set of statistics that capture chosen features of the data. More precisely, for each set of parameters, we first solve the model numerically for the distributions $\{u(\cdot), h(\cdot,\cdot)\}$, and the decisions $\{s(\cdot), \underline{y}(\cdot), \phi_u(\cdot,\cdot), y_0(\cdot,\cdot,\cdot), \phi_e(\cdot,\cdot,\cdot), v(\cdot)\}$. We then simulated data from the model and compute moments.

The SMM estimator is the set of parameters that minimizes the distance between the moments based on actual and simulated data, $\hat{m}_{data}$ and $\hat{m}_{sim}(\theta)$. It is defined as:

$$\hat{\theta} = \arg\min_{\theta} \left\{ [\hat{m}_{data} - \hat{m}_{sim}(\theta)]' \, \Omega \, [\hat{m}_{data} - \hat{m}_{sim}(\theta)] \right\}$$

where $\hat{m}_{data}$ and $\hat{m}_{sim}(\theta)$ are $M \times 1$ vectors, and the weighting matrix Ω is a symmetric positive definite $M \times M$ matrix.

4.2 Identification and Moments

Our model primitives include worker and firm productivity distributions, $\mathfrak{h}(x)$ and $\mathfrak{p}(y)$, the production function, $F(x, y)$, the bargaining power, α, the non-employment benefit function, $B(x)$, the search cost, q, the relative search efficiency of employed workers, s_1, the meeting technology, $M(\xi, V)$, the vacancy cost function, $c(v(y))$, and the separation rate, δ. Although all parameters are jointly estimated in the SMM procedure, some moments are particularly informative about certain model primitives.

The wage dispersions across different types of workers and firms are informative about the dispersions of the productivities $\mathfrak{h}(x)$ and $\mathfrak{p}(y)$. Since wages are always determined in matches between workers and firms, without additional information, we cannot identify the average contributions of workers and firms in the production output. We view $B(x)$ as the non-employment benefit that worker x receives. Given $\mathfrak{h}(x)$, the function $B(\cdot)$ can be identified by matching the variation in non-employment benefit in worker type.[21]

We next turn to the production function $F(x, y)$. The average level of production output can be identified from the wage level. We also allow workers and firms to be substitutable or complementary in the production. The degree to which they are complementary can be identified from the matching pattern, which is summarized by the difference in job finding rate across worker types. The job finding rate for an unemployed worker of rank x is $\kappa \int_{\underline{y}(x)}^{1} v(y')dy'$. The parameter γ influences the shape of the function $\underline{y}(\cdot)$, which in turn influences how the job finding rate varies workers of different ranks.

If we observe the type of workers and firms (x and y), we are able to construct moments to identify the productivity distributions, the production function, and the non-employment benefit function. However, in the data, we only observe the identity of workers and firms, i and j, but not their ranks. Nevertheless, our model informs us about which statistics are consistent with the ranks. Using these statistics, we can first estimate the ranks and then construct statistics based on the estimated ranks which we refer to as bins. In Subsection 4.3, we the procedure to construct the empirical ranks of workers and firms from our DADS datasets. The procedure involves assigning workers and firms into discrete bins according to the ranking statistics. Let $b_x(i)$ and $b_y(j)$ the bins of workers and firms, and let N_{xbin} and N_{ybin} denote the total number of worker and firm bins.

In wage bargaining, an unemployed worker's wage is determined by her receiving α share of the match surplus, unless the minimum wage is binding. As a worker receives one-the-job offers, the worker receives an increasing share of the match surplus through a wage closer to the "best match wage". The higher the worker's bargaining power α is, the smaller the difference is between the out-of-unemployment wage and her later wages, and also the smaller the overall wage dispersion is. Therefore, out-of-unemployment wages and the overall wage dispersion are informative about α.

The unemployment search q affects the labor force participation decision, and thus can be identified from the labor force participation rate. The vacancy cost function $c(\cdot)$ influences the vacancy distribution. However, as we only observe the aggregate level of vacancies, we can only identify the average level of the vacancy cost using the vacancy rate.

The parameter s_1 determines the efficiency of on-the-job search relative to unemployment search. The rate of job-to-job transitions to the unemployment-to-job transitions is informative about s_1. Moreover, a higher s_1 implies workers move faster to more productive firms, resulting in a more skewed firm size distribution. The variation in firm size across firm types is thus also informative

[21]Note that our data does not include individuals who have never been in the labor force, thus our computation of the simulated moments of the benefit levels are also based solely on those in the labor force. Predicting the benefit levels of labor force non-participants relies on the functional form assumption.

about s_1. The average efficiency of the matching technology $M(\xi, V)$ determines the how frequent workers and firms meet. An efficient matching technology leads to more matches and thus less unemployment. We use the unemployment rate (UR) as a moment. However, since we do not observe variations in ξ nor V in the steady state, we cannot identify how each of them separately influence the number of meetings. Finally, the exogenous separation rate δ can be directly estimated from the employment separation rate in the data.

In summary, we use the following moments in the SMM estimation: the median wage by worker and firm bin $\{\text{median}(w_{i,t}|b_x(i) = b)\}_{b=1,...,N_{xbin}}$ and $\{\text{median}(w_{i,t}|b_y(\chi_{i,t}) = b)\}_{b=1,...,N_{ybin}}$, where $\chi_{i,t}$ is the index of the firm in which worker i is employed in period t; the median non-employment benefit by worker bin $\{\text{median}(B_{i,t}|b_x(i) = b)\}_{b=1,...,N_{xbin}}$; the median out-of-unemployment wage $\text{median}(w_{i,t}|\chi_{i,t-1} = 0)$; and the wage percentiles $w(p10)$, $w(p20)$,...,$w(p90)$; the wage distribution over the policy-relevant intervals: $Pr(w \leq 1.05w_{min})$, $Pr(1.05w_{min} < w \leq 1.3w_{min})$, $Pr(1.3w_{min} < w \leq 1.6w_{min})$ and $Pr(1.6w_{min} < w \leq 2.5w_{min})$; the job finding rates by worker bin relative to that of the top bin, $\left\{\frac{Pr(\chi_{i,t}>0|\chi_{i,t-1}=0 \text{ and } b_x(i)=b)}{Pr(\chi_{i,t}>0|\chi_{i,t-1}=0 \text{ and } b_x(i)=N_{xbin})}\right\}_{b=1,...,N_{xbin-1}}$; the vacancy rate ($VR$), defined as $VR \equiv \frac{V}{V+E}$; the labor force participation rate ($LFPR$) and unemployment rate (UR); the job-to-job transition rate relative to the job finding rate (JJ/UE); and the employment share by firm bin, denoted by $\{Emp(b_y = b)/Emp\}_{b=1,...,N_{ybin}}$, where $Emp(b_y)$ is the number of employed workers in firm bin b_y, and Emp is the total number of employed workers.

4.3 Empirical Ranks of Workers and Firms

We discuss the statistics we use to recover the empirical ranks of workers and firms from the data and method of aggregating ranks.

4.3.1 Workers

We use three statistics to rank workers: lifetime earnings, and lifetime minimum and maximum wages. In our model, individuals are risk-neutral and make decisions to maximize their discounted lifetime net income. By assumption, both $F(x, y)$ and $B(x)$ increase in x. High-x workers thus have a greater earnings capacity; lifetime earnings accounting for both labor income and non-employment benefits is a statistic that can be used to rank workers in the data. Because workers appear in our sample for various lengths of time, we take the average daily net earnings, including labor income and non-employment transfers.[22]

[22]Lifetime earnings that does not account for periods of non-employment may not be consistent with worker productivity because the set of firms with which a worker can form viable matches does not necessarily expand with x: a high-x worker may be more selective because of her higher outside options and spend more time in unemployment, and thus using lifetime labor earnings may under-predict the productivity of high-x workers.

Lifetime maximum and minimum wages are also monotonically increasing in x. First consider the lifetime maximum wage. Since all workers labor force participants (employed or unemployed workers) are able to form a viable match with the most productive firm and the maximum match wage increases in x ($\partial\bar{\phi}(x,y)/\partial x > 0$), the lifetime maximum wage is a consistent ranking statistic. Then, consider the lifetime minimum wage. The lowest wage that a worker x can attain in a given firm $y \in \mathcal{A}_u(x)$ is $\max\{w_{min}, \phi_u(x,y)\}$. If $\underline{y}(x)$, the lower bound of $\mathcal{A}_u(x)$, is constant or increasing in x, a less productive worker is able to match with all firms that a more productive worker can match with. Given that $\phi_u(x,y)$ increases in x, the lowest wage a worker can ever attain must be non-decreasing in x. $\underline{y}(x)$ may also decrease in x due to the minimum wage constraint. In this case, the lowest wage that the worker can ever attain must be w_{min}.

In the panel data of DADS, we observe almost complete employment history of each individual over the period our data is available. This allows to construct statistics that are functions of lifetime income or earnings. In order for these statistics to capture the permanent rather than transitory effect of individuals, we need to base the computation on an extended period of time. To rank individuals in DADS, we use the panel data from the period between 1991 to 2008, and we exclude individuals for whom we do not have sufficient observations.[23] To rank individuals in our simulated data, we base on 10 years of simulated data since the average sample duration of workers in the DADS sample is around 10 years.

All three statistics allow for a global ranking of workers and may contain conflicts. We aggregate the three rankings by iterative bisections. At the end of each iteration $n_{xiter} = 1, ..., N_{xiter}$, workers are assigned to $N_{xbin} = 2^{n_{xiter}}$ ranked bins. We use the lifetime net earnings as the primary statistic as it is measured with the least noise. In case of a disagreement among the three rankings, the mechanism follows the primary statistic unless the other two statistics simultaneously show a strong indication for the alternative. We choose $N_{xiter} = 3$, resulting in $N_{xbin} = 8$ worker bins. Table 4.1 shows that the average values of all three statistics increase in the discrete worker bin b_x.

4.3.2 Ranking Firms

Our model predicts that the highest worker type a firm can attract increases with firm productivity. In addition, the "best match wage" also monotonically increases in both worker and firm types ($\partial\bar{\phi}(x,y)/\partial x \geq 0$ and $\partial\bar{\phi}(x,y)/\partial y \geq 0$). Thus, in probability, the highest wage observed in a firm increases with the firm type.

[23] In Appendix G, we describe details on sample selection and the computation of the ranking statistics.

Table 4.1: Ranking statistics of workers by discrete worker bins (b_x). All statistics are monthly in 2010 euros. For each individual, the statistics are computed based on the panel data from the DADS from 1991 to 2008, restricted to men age 30-55 who are primarily employed in non-executive, full-time, private sector jobs. "Lifetime income" refers to the net-of-tax average income per month, accounting for net wage and imputed non-employment benefits. "Lifetime min. wage" and "Lifetime max. wage" refer to the lowest and the highest net wage from employment that an individual obtains while in sample. Individuals are assigned to bins based on the three statistics such that those with higher income and wages are put to a higher bin.

Discrete worker bin (b_x)	Lifetime income	Lifetime min. wage	Lifetime max. wage
1	915.6	328.8	1973.4
2	1177.5	632.1	2160.6
3	1332.6	847.8	2265.3
4	1475.4	1009.8	2324.1
5	1631.4	1179.9	2541.0
6	1828.2	1366.2	2727.6
7	2114.4	1613.4	3151.2
8	2896.2	2146.8	5184.0

In the panel data of DADS, we only observe 1/12 or 1/24 of all employee earnings within each firm on average. It is thus impossible to construct from the panel data reliable firm-level statistics of the within-firm wage distribution. In the POST data, however, we observe wages of all employees within each firm.[24]

In practice, we use the 99th within-firm wage percentile to rank firms that have an average firm size of at least 10 workers, where firm size is defined as the number of male workers age 30-55 in non-executive, full-time, private sector positions. In this way, about 30% of firms in POST can be ranked, accounting for over 80% of employment.[25] Firms satisfying our ranking criteria are assigned to a total of $N_{ybin} = 4$ ranked bins such that a firm with a lower 99th wage percentile measure is assigned to a lower bin.

Table 4.2 compares firm wages and firm size between ranked (larger) and unranked (smaller) firms; it is clear that wages are lower in the unranked firms. According to our model, firm size monotonically increases in firm productivity. The set of ranked firms is thus not randomly selected, but rather represents more productive firms. We replicate this selection in our simulation by using the same firm size criteria to determine the set of ranked firms and matching the share of employment accounted for by ranked firms. In addition, to rank firms in POST, we use data from the period between 1991

[24] We are not able to construct the poaching index, which is the fraction of hires that were poached from another firm, used by Bagger and Lentz (2014) in either dataset. This is because of the incomplete within-firm observations in the panel and the fact that we cannot track workers across years in POST. The poaching index is

[25] See Appendix G.1 for details on firm size distribution.

Table 4.2: Firm statistics for ranked and unranked firms. Statistics are computed based on POST, and restricted to jobs that are non-executive, full-time, in the private sector, and are filled by men of age 30-55. The "Highest firm wage" of a firm is the 99th monthly wage percentile ever reported by the firm; the "Mean firm wage" of a firm is the average monthly wage ever reported by the firm; and "Firm size" is the total number of employee-days divided by the total number of days a firm is observed in POST. All wages are monthly gross wage measured in 2010 Euros.

Firm statistic:	(1) Highest firm wage	(2) Mean firm wage	(3) Firm size
Ranked firms (firm size ≥ 10)	5277.8	2306.0	54.8
Unranked firms (firm size < 10)	3110.7	2106.7	1.50
All firms	3236.8	2116.4	4.10

to 2008. To rank firms in our simulated data, we base on 10 years of simulated data since it is the average sample duration of firms.

4.4 Parametrization

We assume that worker and firm productivities follow log-Normal distributions. More precisely, we have $\mathfrak{h}(x) = \exp\left[\Phi_x^{-1}(x)\right]$ and $\mathfrak{p}(y) = \exp\left[\Phi_y^{-1}(y)\right]$ with support $[0,1]$. Φ_x is the cumulative distribution function of the Normal distribution $N(\tilde{\mu_x}, \tilde{\sigma}_x)$. The mean and standard deviation of the worker productivity are thus $\mu_x = \exp(\tilde{\mu}_x + \tilde{\sigma}_x/2)$ and $\sigma_x = \sqrt{[\exp(\tilde{\sigma}_x) - 1]\exp(2\tilde{\mu}_x + \tilde{\sigma}_x^2)}$. Similarly, Φ_y is the cumulative distribution function of the Normal distribution $N(\tilde{\mu_y}, \tilde{\sigma}_y)$, and the mean and standard deviation are $\mu_y = \exp(\tilde{\mu}_y + \tilde{\sigma}_y/2)$ and $\sigma_y = \sqrt{[\exp(\tilde{\sigma}_y) - 1]\exp(2\tilde{\mu}_y + \tilde{\sigma}_y^2)}$. We normalize $\mu_x = 1$ and $\mu_y = 1$. Note that not all firms on the support $[0,1]$ are active. We impose an exogenous upper bound $p_h = 4$ on firm productivity. Given Φ_y, the upper bound on firm type is $y_h = \Phi_y(\ln(p_h))$. In addition, there is also the endogenous lower bound on the firm distribution. As a result of the bounds, σ_y may be greater than the standard deviation of the observed firm productivity distribution.[26]

The production function admits constant elasticity of substitution (CES) between $\mathfrak{h}$ and $\mathfrak{p}$:

$$F(x,y) = \begin{cases} f_0\left[\frac{1}{2}\mathfrak{h}(x)^\gamma + \frac{1}{2}\mathfrak{p}(y)^\gamma\right]^{1/\gamma} & \text{if } \gamma \neq 0 \\ f_0 \mathfrak{h}(x)^{\frac{1}{2}}\mathfrak{p}(y)^{\frac{1}{2}} & \text{if } \gamma = 0 \end{cases}$$

[26] While σ_y is related the wage dispersion across firms, the relationship is not necessarily monotone because not all firms are always active. On one hand, a very low σ_y implies little firm productivity dispersion. On the other hand, a high σ_y may cause individuals to become more selective in accepting matches, and this increases the productivity of the least productivity firm p_l. This compresses the interval over which firms are active.

with $f_0 > 0$ and $\gamma \leq 1$, where f_0 is total factor productivity and $1/(1-\gamma)$ is the elasticity of substitution between worker and firm productivities. If $\gamma = 1$ worker and firm productivities are perfect substitutes. If $\gamma = -\infty$, the productivities are perfect complements. If $\gamma = 0$, the production function is Cobb-Douglas.

We assume that the non-employment benefit is a linear function in worker's productivity such that

$$B(x) = b_0 + b_1 \mathfrak{h}(x)$$

Simulated benefit levels under the linear specification provides a close match to the data, so we do not include terms of higher orders in $B(\cdot)$.

We assume that the following functional form for the vacancy cost function:

$$c(v) = (c_0 v)^{1+1/c_1}$$

We cannot identify c_1 due to the lack of data on the distribution of vacancy across firm ranks, we fix it at 0.01 so that it allows for a plausible firm size distribution.[27]

Finally, we assume that the meeting technology $M(\xi, V)$ takes the following form:

$$M = m_0 \sqrt{\xi V}$$

4.5 Estimation results

In the SMM estimation, we simulate 100,000 individuals whose types are drawn from the discretized worker productivity distribution with 100 grid points. We simulate 2,000 firms drawn from the discretized firm productivity distribution with 50 grid points. The computation of simulated moments is based on 36 months of simulated data, consistent with our moment computation using the DADS data. We consider a discrete time version of our model by aggregating to the monthly level.

Before carrying out the SMM estimation, we estimate the separation rate δ to be 0.00855 per month, which implies an annual separation rate of 0.098.[28] Table 4.3 shows parameter estimates; we discuss their interpretation in turn.[29] The estimate of the parameter γ indicates that workers and firms

[27] Bagger and Lentz (2014), who estimate a search-and-matching model based on Danish data, estimate a c_1 on the same order of magnitude.

[28] The separation rate is computed based on monthly data of ranked workers. It is equal to the fraction of workers who are employed in any type of job in the current month that become unemployment in the following month.

[29] The standard errors can be obtained using the bootstrap method; they will be made available soon.

are complementary in production at a degree slightly greater than a Cobb-Douglas specification. Given the estimated values of γ, f_0, σ_x, and σ_y, we plot the production function in Figure 4.1. It shows that the output of a median worker in terms of her productivity at a median active firm in terms of its productivity is 2633 euros per month. While the upper bound of the firm productivity distribution is fixed at $p_h = 4$, the endogenous lower bound is $p_l = 2.85$. Given the estimated σ_y, the standard deviation of the productivity of active firms is 0.33, which is smaller than the estimated standard deviation of worker productivity, $\sigma_x = 0.42$. If matched with the median active firm, the worker with at the 90th percentile of the productivity distribution produces 1.49 times the output of the median worker, while the median worker produces 1.51 times the output of the worker at the 10th percentile of the productivity distribution. Given the median worker, the firm at the 90th percentile of productivity distribution of active firms produces 1.07 times the output of the median firm, while the median firm produces 1.06 times the output as the firm at the 10th percentile of the same distribution.

The parameter estimates of b_0 and b_1 indicate that while all non-employed workers receive a basic level of transfer of 590.5 euros per month, the transfer increases with worker productivity. In particular, the marginal increase in the transfer with respect to $\mathfrak{h}$ is 654.9 euros for all workers. For the median worker, the marginal increase in production output with respect to $\mathfrak{h}$ ranges from 1574 to 1873 euros per month.

The job search parameters m_0 and s_1 imply that, on average, an unemployed worker meets a vacancy every 9.0 months, an employed worker meets a vacancy every 14.7 months, and a vacancy meets a worker every month.[30] The search cost q is small. Estimated as 13.2 euros per month, it is only 1.4% of the minimum wage. The vacancy cost parameter c_1 implies that the cost of posting one additional vacancy for the least productive firm is 3.15 euros per month, while the cost for the most productive firm is 1472 euros per month.

Finally, our estimated workers' bargaining power α is 0.729, which is higher than the common findings in the related literature due to differences in the model environment and identification strategy.[31] Our estimate of α is driven by a relatively low dispersion in the net wage and a high out-of-unemployment wage. If we assume that individuals care about the labor cost instead of the net wage in wage bargaining, we would end up with a lower estimate of α because there would be a greater level of dispersion in the labor cost.

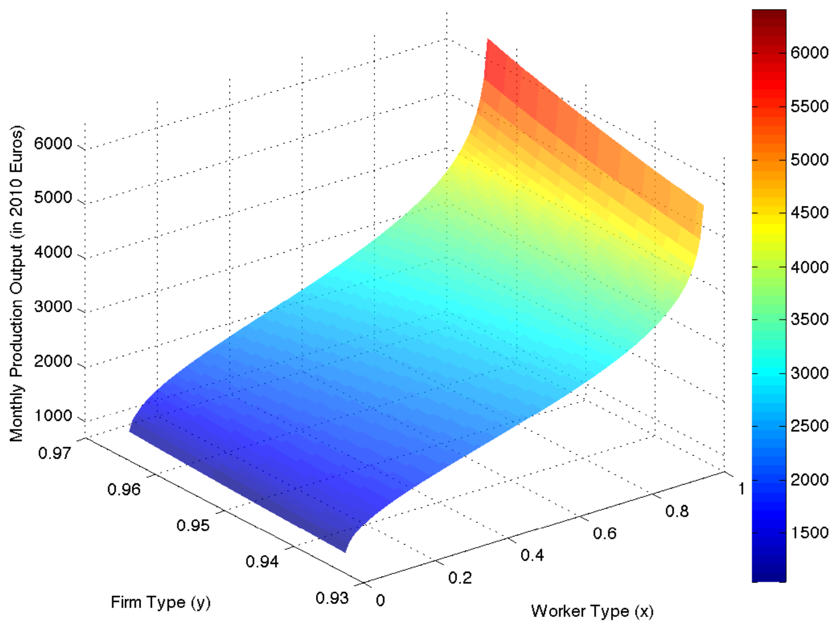
[30] Note that these are meeting rates. The job finding rate also depends on the probability of forming a viable match.

[31] For example, using the relationship between job productivity and labor cost, Cahuc et al. (2006) estimate the bargaining power to be close to zero for workers in the two least skilled categories.

Table 4.3: Parameter Estimates.

Parameter	Value
Production function: $F(x,y) = f_0(\frac{1}{2}\mathfrak{h}(x)^\gamma + \frac{1}{2}\mathfrak{p}(y)^\gamma)^{1/\gamma}$	
γ	-0.12
f_0	1626.27
Dispersion parameter of the worker productivity distribution	
σ_x	0.424
Dispersion parameter of the firm productivity distribution	
σ_y	2.49
Non-employment benefit: $B(x) = b_0 + b_1\mathfrak{h}(x)$	
b_0	590.5
b_1	654.9
Efficiency of on-the-job search relative to unemployment search:	
s_1	0.593
Meeting technology: $M(\xi, V) = m_0\sqrt{\xi V}$	
m_0	0.914
Cost of unemployment search:	
q	13.2
Vacancy price: $c(v) = (c_0 v)^{100}$	
c_0	1980.7
Worker's share of surplus:	
α	0.729

Figure 4.1: Estimated production function. Parameter values are shown in Table 4.3.



4.6 Model Fit

Before showing the fit of the moments, we note that our firm ranking scheme leaves us with a similar fraction of ranked firms in the simulated data as in the actual data. Of the 2000 simulated firms, 31.25% are ranked. These firms account for 81.9% of total employment. In comparison, 29% of firms in POST are ranked, accounting for 88% of total employment.

Table 4.4 shows the fit of the targeted moments. Overall our model is able to fit the moments well. We replicate the hump-shape pattern in the job finding rate: it is lower among the lowest and the highest ranked workers and higher for those in the middle.

Our model also closely matches the unconditional wage distributions and wage dispersion across worker bins. We under-predict the wage dispersion across firms ranks (moments 29-32). It should be noted, however, that our simulation reveals that the firm ranking is substantially less accurate than worker ranking. The correlation between our simulated workers' rank b_x and true type x is 0.983. 77% of the simulated workers are correctly ranked, 23% are ranked higher or lower by one bin. Of the ranked firms, the correlation between firms' rank b_y and type y is 0.390. 33.6% are correctly ranked, 44.6% are ranked higher or lower by one bin.

The rest of the moments, including the median out-of-unemployment wage, non-employment transfers, and measures of labor force stocks and flows, all closely match their data counterparts.

5 Equilibrium effects of payroll tax reductions

In this section, we examine the equilibrium effects of a low-wage payroll tax reduction. A low-wage tax reduction relaxes the minimum labor cost constraint on the set of viable matches. As $y_{min}(x)$ shifts downward for all x, $\underline{y}(x)$ also shifts downward for low-productivity workers such that $\underline{y}(x) = y_{min}(x)$., and certain low-productivity workers who were previously non-participants may decide to choose $s(x) = 1$. As a result of these, the set $\mathcal{B}_u(y)$ expands for certain low-productivity firms so that they can match with more workers, which incentives them to post more vacancies. In addition, previously inactive low-productivity firms may become active by posting a positive level of vacancies.

The labor force entry of low-productivity workers and the increased vacancies from low-productivity firm can intensify the congestion externality. Hold-up and congestion are two classical externalities in a frictional labor market. Hold-up happens when agents fail to internalize the full social value of a match, and congestion happens when agents fail to internalize that their search activity lowers the meeting opportunities of others. Shimer and Smith (2001) show that, in an labor market with heterogeneity agents, the congestion externality dominates the hold-up externality for low-productivity agents because there is a low social value of matching with them. A policy that biases toward low-productivity agents would thus lead to more congestion.

In the following, we examine the extent to which congestion affects employment and aggregate production by simulating a tax reduction. Starting from late-1995, the then French prime minister Alain Juppé enforced tax reductions that targeted jobs with a wage less than 1.3 times the minimum wage.[32] There are also increases in the taxes in other parts of the wage distribution. Figure 5.1 shows the overall changes in the tax schedule. We simulate these changes by imposing the tax schedule in 1997 to the baseline environment.

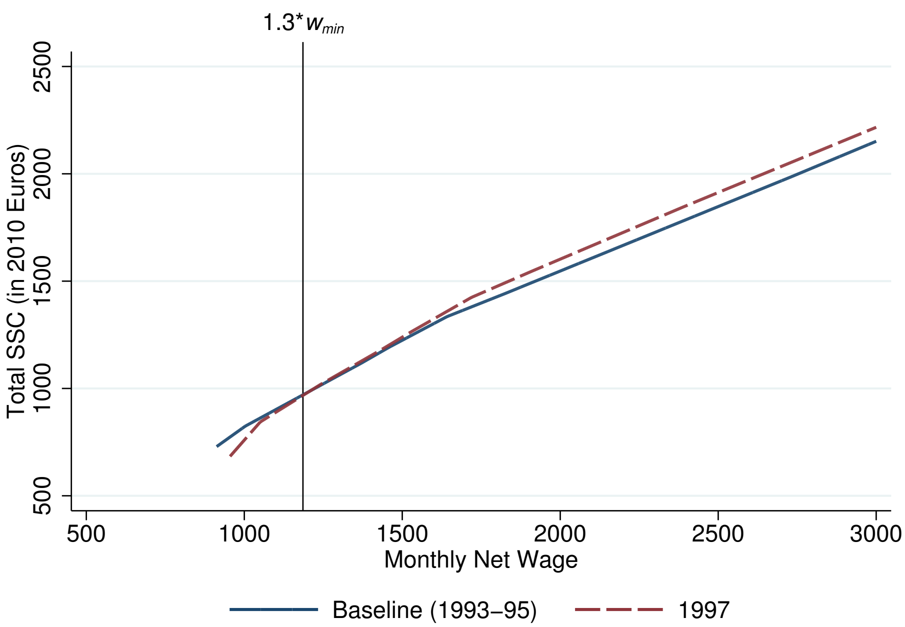
We impose the assumption that the government keeps a neutral budget: tax revenues are first used to finance non-employment benefit payments, and the remaining revenue is redistributed to the

[32] In practice, the tax changes are applied to the employer share of the payroll tax. Therefore, the tax reductions do not increases the statutory minimum income for workers.

Table 4.4: Moments.

	Moment	Data	Simulation		Moment	Data	Simulation
					Median wage by firm bin:		
	Job finding rate of workers in bin b_x				median$(w_{i,t} \mid b_y(\chi_{i,t}) = b_y)$		
	relative to that of workers in the top bin:						
	$\frac{Pr(\chi_{i,t}>0 \mid \chi_{i,t-1}=0 \text{ and } b_x(i)=b_x)}{Pr(\chi_{i,t}>0 \mid \chi_{i,t-1}=0 \text{ and } b_x(i)=8)}$						
1	$b_x = 1$	0.472	0.518	29	$b_y = 1$	1358	1497
2	$b_x = 2$	0.870	1.004	30	$b_y = 2$	1527	1557
3	$b_x = 3$	1.048	1.089	31	$b_y = 3$	1717	1589
4	$b_x = 4$	1.200	1.127	32	$b_y = 4$	1861	1621
5	$b_x = 5$	1.201	1.143		*Median out-of-unemployment wage*		
6	$b_x = 6$	1.209	1.142	33	median$(w_{i,t} \mid \chi_{i,t-1} = 0)$	1492	1491
					Median non-employment benefit by worker		
7	$b_x = 7$	1.196	1.129		*bin:* median$(B_{i,t} \mid b_x(i) = b_x)$		
	Wage percentiles:			34	$b_x = 1$	811	833
8	$w(p10)$	1131	1090	35	$b_x = 2$	941	961
9	$w(p20)$	1280	1226	36	$b_x = 3$	1017	1044
10	$w(p30)$	1401	1340	37	$b_x = 4$	1090	1143
11	$w(p40)$	1518	1460	38	$b_x = 5$	1196	1243
12	$w(p50)$	1638	1575	39	$b_x = 6$	1338	1388
13	$w(p60)$	1769	1716	40	$b_x = 7$	1540	1594
14	$w(p70)$	1929	1885	41	$b_x = 8$	2181	2095
15	$w(p80)$	2155	2107		*Labor force participation rate*		
16	$w(p90)$	2502	2438	42	$LFPR$	0.947	0.95
	Wage distribution relative to w_{min}				*Unemployment rate*		
17	$Pr(w \leq 1.05 w_{min})$	0.040	0.033	43	UR	0.077	0.085
18	$Pr(1.05 w_{min} < w \leq 1.3 w_{min})$	0.093	0.148		*Job-to-job transition rate relative to*		
					unemployment-to-job transition rate		
19	$Pr(1.3 w_{min} < w \leq 1.6 w_{min})$	0.217	0.166	44	JJ/UE	0.091	0.105
20	$Pr(1.6 w_{min} < w \leq 2.5 w_{min})$	0.494	0.499		*Vacancy rate (vacancies over vacant and filled jobs)*		
	Median wage by worker bin:			45	VR	0.011	0.011
	median$(w_{i,t} \mid b_x(i) = b_x)$						
21	$b_x = 1$	998	1014		*Employment share by firm bin: $Emp(b_y)/Emp$*		
22	$b_x = 2$	1176	1191	46	$b_y = 1$	0.140	0.127
23	$b_x = 3$	1293	1335	47	$b_y = 2$	0.256	0.240
24	$b_x = 4$	1418	1477	48	$b_y = 3$	0.331	0.325
25	$b_x = 5$	1562	1638	49	$b_y = 4$	0.272	0.308
26	$b_x = 6$	1731	1837				
27	$b_x = 7$	1978	2107				
28	$b_x = 8$	2496	2620				

Figure 5.1: Payroll tax schedule in the baseline period (January 1995 to August 1997) and in 1997. The relationship between the payroll tax (social security contributions, or SSC, in France) and net wage is estimated based on data from the DADS by regressing SSC on net wages using a linear spline model. SSC for each job spell is computed by TAXIPP. We restrict our sample to non-executive, full-time, and private sector jobs.



entire population as a lump-sum transfer, which we denote by D_t. More precisely,

$$D_t = \frac{1}{N_i}\left[\sum_{i:\chi_{i,t}>0} T(w_{i,t}) - \sum_{i:\chi_{i,t}\leq 0} B(x_{i,t})\right] \tag{15}$$

where $\chi_{i,t} > 0$ if the worker i is employed in period t, and $\chi_{i,t} \leq 0$ if the worker is non-employed. Since individuals in our model are risk-neutral and do not save, their consumption is equal to the sum of labor income and transfers:

$$c_{i,t} = \mathbf{1}_{\{\chi_{i,t}\leq 0\}} B(x_{i,t}) + \mathbf{1}_{\{\chi_{i,t}>0\}} w_{i,t} + D_t$$

Table 5.1a shows the simulated equilibrium effects on the aggregate level. Overall, our simulation suggests that the payroll tax changes lead to a 2.5% increase in employment. Despite the drop in average job productivity, there is a 1.2% increase in aggregate production. The effects on employment and production are in line with Crépon and Desplatz (2003) and Chéron et al. (2008). The increase in employment and the increase in the tax rate for highly paid jobs lead to higher tax revenue and thus higher tax redistribution. On average, individuals are better-off with an 1.2% increase in consumption.

There is a 2.1% increase in labor force participation; the new entrants are less productive workers whose employment opportunities were previously constrained by the minimum labor cost. Moreover, there is a 4.8% increase in vacancies. Figure 5.2 shows that the increase in vacancies is concentrated in low-productivity firms. To see how the increase in labor force participation and vacancy creation generate congestion effects, we examine the effects on workers in different productivity quartiles. Table 5.1b shows the results.

Due to the increased employment, there is a greater tax revenue and less non-employment benefit payments. These in turn results in an increase in the tax redistribution D_t. As a result, all workers gain from the tax changes. However, the least productive workers benefit the most in terms of consumption, which can be explained by two main factors. First, as tax increases for low-wage jobs and decreases for high-wage jobs, low-productivity individuals benefit from higher net wages and an expansion of the set of viable matches, whereas high-productivity workers suffer from lower wages. Nevertheless, there is little change to the set of viable matches for high productivity workers. This can be seen from the fact that there is little change in average job productivity for the most productive quartile (Q4).

The second factor is related to the congestion externality exerted by low-productivity workers and firms. This can be seen from that the most productive quartile experiences a 0.6% drop in their employment. To understand the congestion effect in our framework, note that all workers and firms engage in random search in the same market and the meeting technology is constant returns to scale. The increased search activity from low-productivity workers negatively affects the changes that a high-productivity worker can meet a vacancy. Moreover, as low-productivity firms cannot form viable matches with high-productivity workers, their vacancies negatively affect the chance that a high-productivity worker can find a viable match. Besides these two factors, a change in the tax schedule also affect the present value of non-employment $W_n(x)$, which in turn affects wage bargaining between with non-employed workers. The change in $W_n(x)$ is qualitatively consistent with the above two factors: it increases for low-productivity workers and decreases for high-productivity workers.

As workers and firms are estimated to be complementary in production, the congestion effect has strong implications for aggregate production. Despite a relatively moderate decrease in employment, the most productive quartile of workers accounts for a 0.2% drop in aggregate production, which brings the overall effect on aggregate production to 1.2% rather than 1.4%.

The results shown in this section indicate that even a relatively moderate low-wage tax reduction policy can lead to quantitatively important congestion effects. Therefore, policy evaluations and designs should account for not only the direct effects on the targeted population, but also the effects on those who are not directly targeted.

6 Optimal coverage of payroll tax reductions

Since the Juppé reform in the 1990s, France has continued its efforts to reduce payroll tax for low-paid work. These tax reductions cover an increasingly broader range of jobs.[33] The coverage of a tax reduction will likely affect the balance between the positive direct effects on match value and viability and the negative indirect effects that result from congestion by low-productivity workers and firms. A trade-off between aggregate production and equality when considering the coverage of a tax reduction. On one hand, a policy that only reduces tax for a narrow range of minimally-paid jobs may be more likely to generate stronger congestion externalities. On the other hand, a narrower tax reduction may be more effective in redistribution toward low-productivity workers.

[33]Most notable tax reductions are the Fillon reform in 2003-2005 that provided reductions in social security contributions (French payroll tax) for jobs with a wage of up to 1.6 times the minimum wage, and the “tax credit for competitiveness and employment (CICE)” announced in 2012 that reduces payroll taxes for jobs earning up to 2.5 times the minimum wage. For institutional details, see Bunel and L’Horty (2012) and André et al. (2015).

Table 5.1: Simulated effects of tax changes between the baseline period and 1997. The baseline period is between January 1993 to August 1995.

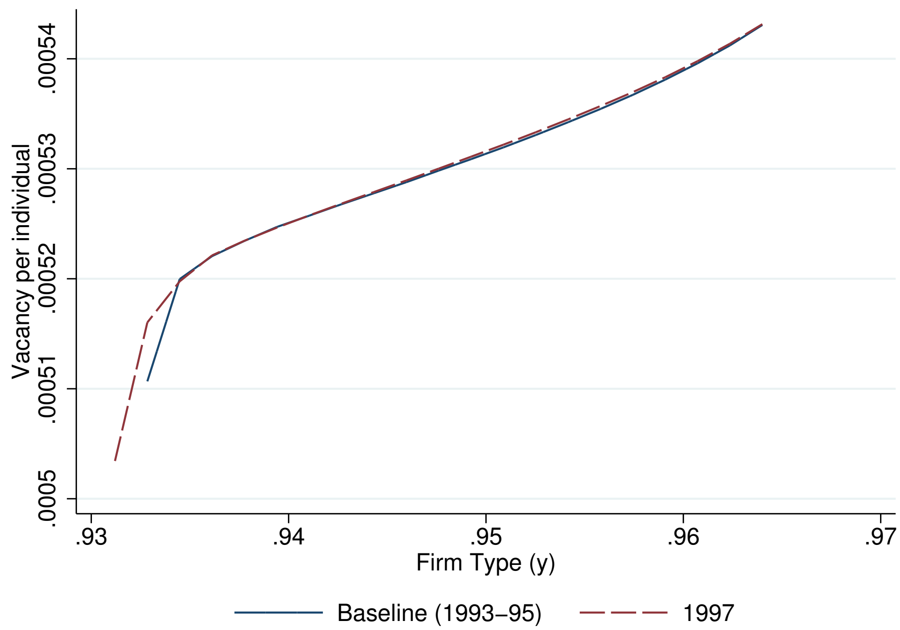
(a) Aggregate effects. The values shown are percentage changes from the baseline.

Vacancies	4.8%
Employment	2.55%
LF participation	2.1%
Job productivity	-1.35%
Aggregate production	1.16%
Tax Redistribution	3.85%
Consumption	1.15%

(b) Distributional effects. Worker quartiles represent the quartiles of worker productivity, with Q1 being the least productive. "Emp." refers to employment, "Job Prod." refers to average job productivity, and "Agg. Prod." refers to aggregate production output. The values in Columns (1)-(3) are percentage changes from the baseline for the worker quartile. In Column (4), the value refers to the percentage change in the overall production output that is accounted for by the worker quartile.

Worker Quartiles	(1) Consumption	(2) Emp.	(3) Job Prod.	(4) Agg. Prod.
Q1	2.92%	13.58%	-2.47%	1.43%
Q2	1.47%	-0.09%	-0.10%	-0.04%
Q3	0.75%	-0.07%	-0.10%	-0.05%
Q4	0.16%	-0.59%	0.06%	-0.20%

Figure 5.2: Simulated effects of changes in payroll taxes between the baseline period and 1997 on the vacancy distribution. The baseline period is between January 1993 to August 1995.



6.1 Framework

To simplify our analysis, we restrict our attention to programs that offer the maximum tax reduction for minimum wage jobs and phase out at the upper bound of the coverage threshold. This is a framework that encompasses the two large payroll reduction programs implemented in France: the Fillon reform and the CICE reform. We refer to the tax reduction as a payroll subsidy, and parametrize the magnitude of the subsidy with the following function:

$$\text{Subsidy}(w) = \begin{cases} (subb \times w_{min} - w)suba & \text{if } w \leq subb \times w_{min} \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

where $suba$ is the generosity of the subsidy and $subb$ is the tax coverage threshold with respect to the statutory minimum wage. Taxes are reduced by the amount of the subsidy, and the effective tax schedule becomes

$$T(w; suba, subb) = \max\{0, T(w) - \text{Subsidy}(w)\} \tag{17}$$

where $T(w)$ is the baseline tax schedule (Fig. 3.1).

We also consider a criteria for evaluating alternative tax reduction programs, which is a social welfare criterion that allows policy makers to care more about the consumption of the less well-off than others. Based on a sample of N_i individuals over T periods, the social welfare level $\mathcal{W}$ is computed as follows:

$$\mathcal{W} = \frac{1}{TN_i} \sum_{t=1}^{T} \sum_{i=1}^{N_i} welfare(c_{i,t}) \tag{18}$$

where $welfare(c_{i,t})$ is the individual welfare weight placed on worker i in period t if her consumption is $c_{i,t}$. We assume a constant-relative-risk-aversion formulation for $welfare(\cdot)$:

$$welfare(c_{i,t}) = \begin{cases} \frac{c_{i,t}^{1-\rho}}{1-\rho} & \text{if } \rho \neq 1 \text{ and } \rho > 0 \\ \log c_{i,t} & \text{if } \rho = 1 \end{cases} \tag{19}$$

We consider the case when $\rho = 4$, in which the policy maker is concerned with the distribution of consumption as well as the levels of consumption.

6.2 Results

We begin by presenting equilibrium effects of payroll tax reductions that raise the baseline employment by 5%. Figure 6.1 plots the percentage change in average consumption against the subsidy coverage. The subsidy generosity is adjusted so that the equilibrium employment is 5% higher than the baseline level. We show the average effects as well as effects by productivity quartiles. The results indicates that individuals in the least productive quartile are better-off under subsidies of any coverage, with an consumption increase of around 5%. They are slightly more better-off under a narrower subsidy than a broader one. In contrast, individuals in the top productivity quartile are strictly worse-off under a narrower subsidy. As the coverage broadens, the negative effects dissipates.

The negative effects of a narrow subsidy on high-productivity individuals is due to congestions from low-productivity workers and jobs. The differing employment effects on different workers are evidence for the congestion effect. As shown in Figure 6.2, under a subsidy with a narrow coverage, the employment effect on the least productive workers is more positive, while the employment effect on the most productive workers is more negative. As the coverage broadens, the congestion becomes less severe and the negative effects on high-productivity workers disappears.

Figure 6.1: Effects of expanding subsidy coverage on consumption by quartiles of worker productivity. All subsidy programs raise baseline employment by 5%. The lines “Q1-Q4” represent quartiles of worker productivity. “Change in consumption” is the difference in the average consumption $c_{i,t}$ between the counterfactual and the baseline environments. The subsidy coverage threshold parameter *subb* is a multiple of the minimum wage (see Eq. 16). The graph has been smoothed using a third degree polynomial.

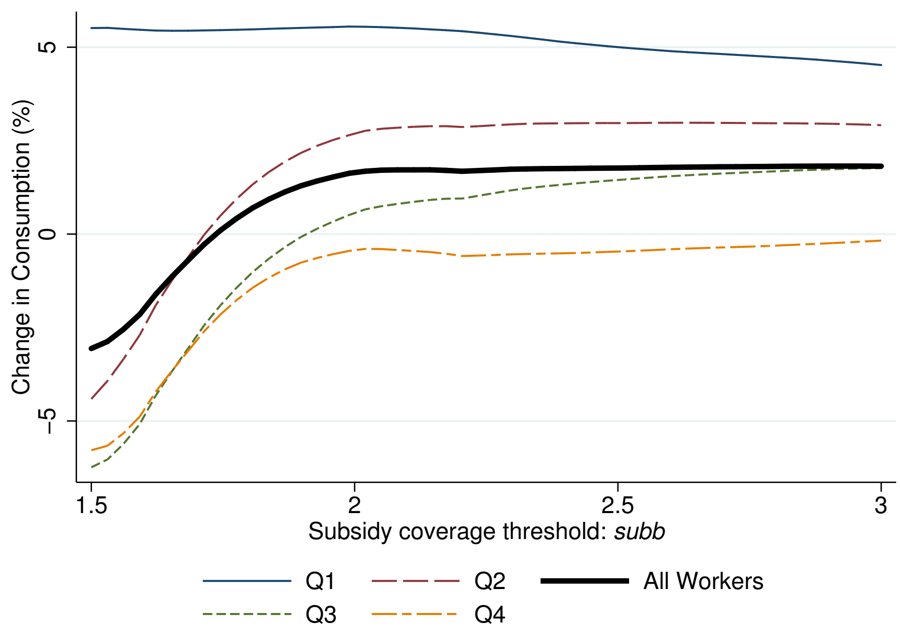


Figure 6.2: Effects of expanding subsidy coverage on employment by quartiles of worker productivity. All subsidy programs raise baseline employment by 5%. The lines "Q1-Q4" represent quartiles of worker productivity. "Change in employment rate" is the difference in the employment to population rate between the counterfactual and the baseline environments. The subsidy coverage threshold parameter *subb* is a multiple of the minimum wage (see Eq. 16). The graph has been smoothed using a third degree polynomial.

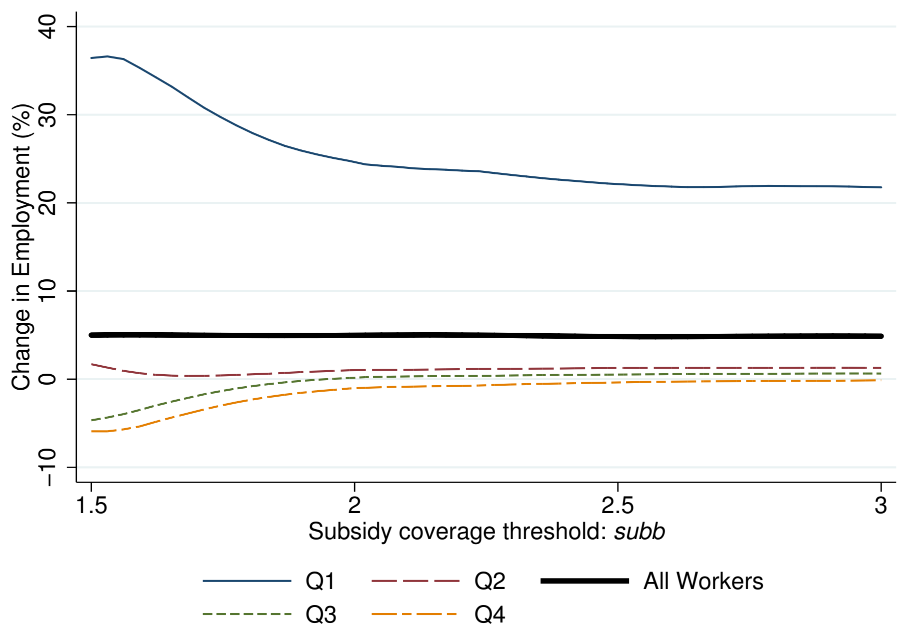
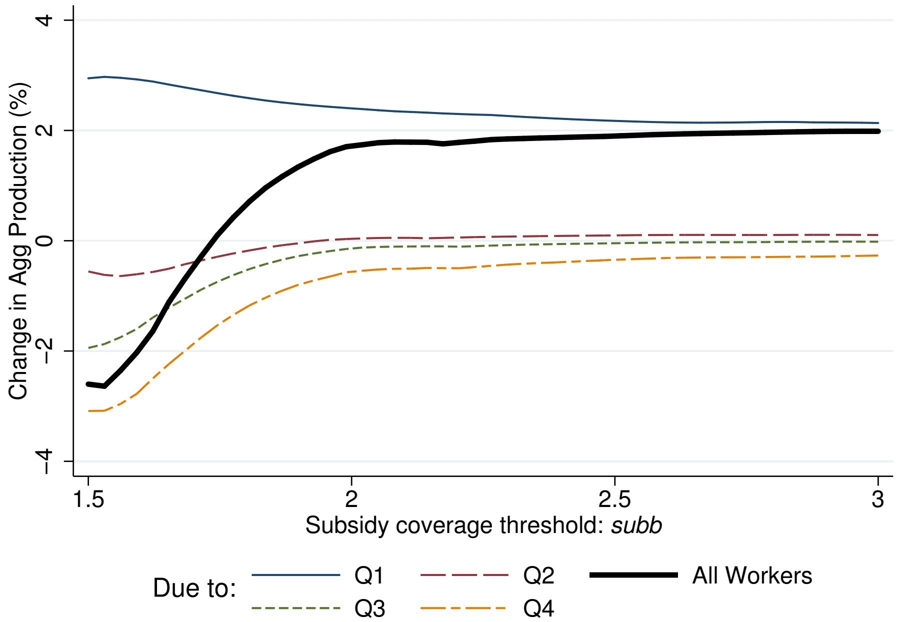


Figure 6.3: Effects of expanding tax reduction coverage on aggregate production by quartiles of worker productivity. All tax reduction programs raise baseline employment by 5%. The lines "Q1-Q4" represent quartiles of worker productivity. "Change in production" is the difference in the monthly total production output divided by the number of individuals between the counterfactual and the baseline environments. The subsidy coverage threshold parameter *subb* is a multiple of the minimum wage (see Eq. 16). The graph has been smoothed using a third degree polynomial.



Although the negative employment effects on the most productive workers seem moderate compared to the positive employment effects on the least productive workers, they are disproportionally costly for aggregate production. Figure 6.3 shows how the overall change in aggregate production is accounted for by each quartile of workers. With a subsidy that covers jobs of less than 1.5 times the minimum wage, the bottom quartile of workers accounts for a 3% increase in aggregate production, while the top quartile accounts for a 2% decrease in aggregate production. As the subsidy coverage broadens, the gap between the top and bottom workers shrinks.

The results above clearly indicate a trade-off between aggregate production and equality: a narrower subsidy coverage narrows the consumption gap between the most and the least productive workers, but it lowers the employment from the most productive workers, resulting in significant drops in aggregate production. Due to this trade-off, a certain level of subsidy coverage may be optimal in balancing efficiency and equity. We use the welfare criterion in Eq. 18 with $\rho = 4$ to evaluate alternative subsidies.[34]

Figure 6.4 shows the social welfare gains under subsidy programs that raises the baseline employment by 2, 3, and 5%. To increase employment by 5%, the optimal tax reduction program is one with a relatively broad coverage: jobs that make up to 2.1 times the minimum wage should benefit from the tax reduction. To reach the same level of employment increase, a narrower program leads to worse welfare outcomes compared to a tax reduction program with a more moderate employment goal. Overall, a higher employment goal requires a broader coverage.

Chéron et al. (2008) conduct a similar exercise regarding the subsidy coverage. However, instead of considering fixed employment goals, they are concerned with the restricting alternative subsidies to the same ex-ante fiscal cost (i.e. the same upfront sum of subsidies before equilibrium labour market reactions are taken into account). They conclude that the optimal coverage threshold is 1.36 times the minimum wage, which results in an employment rise of around 2%. Their result is not inconsistent with our findings. Note that having a fixed ex-ante fiscal cost does not imply that the equilibrium budget cost is unchanged. Based on our simulations, ex-post (equilibrium) fiscal budget can be drastically different from the ex-ante budget and their relationship can be non-monotone. Figure 6.5 shows the equilibrium fiscal costs of different subsidy programs, measured as decreases in D from the baseline. We can see that, although raising employment by 5% is highly costly ex-ante, at the optimal coverage threshold of around 2.1 times the minimum wage, the equilibrium fiscal cost is close to zero.

To close this section, we would like to point out that, as in most equilibrium studies, our analyses have focused exclusively on the steady state equilibrium and placed no weight on periods when the economy is in transition. The speed at which employment responds to a subsidy and the potential short-run welfare costs to individuals may be highly relevant for policy makers, particularly if the policy makers have a short horizon due to term limits of their office or if there are frequent policy changes. A policy evaluation conducted using the short-run data may not fully capture the employment effects. In addition, a subsidy program that maximized the equilibrium welfare may be less desirable in the short run. However, simulating the short-run dynamics is difficult because our model is intractable off the steady state equilibrium.[35] In Appendix H, we develop approximation

[34]Alternative specifications of $\rho = 1$ and 2 give similar results regarding the optimal subsidy coverage.

[35]Robin (2011) and Lise and Robin (2017) develop dynamic equilibrium search-and-matching models to study aggregate shocks. A key assumption that allow them to maintain model tractability is that firms have full bargaining power, so that decisions of search and match can be determined independent of wages. Since wage is important to studying payroll taxes, we cannot make the same assumption.

Figure 6.4: Simulated welfare gain under alternative subsidies. The welfare criterion is shown in Eq. 18 with $\rho = 4$. Subsidy generosity *suba* varies such that the subsidy program raises baseline employment by the corresponding percentage.

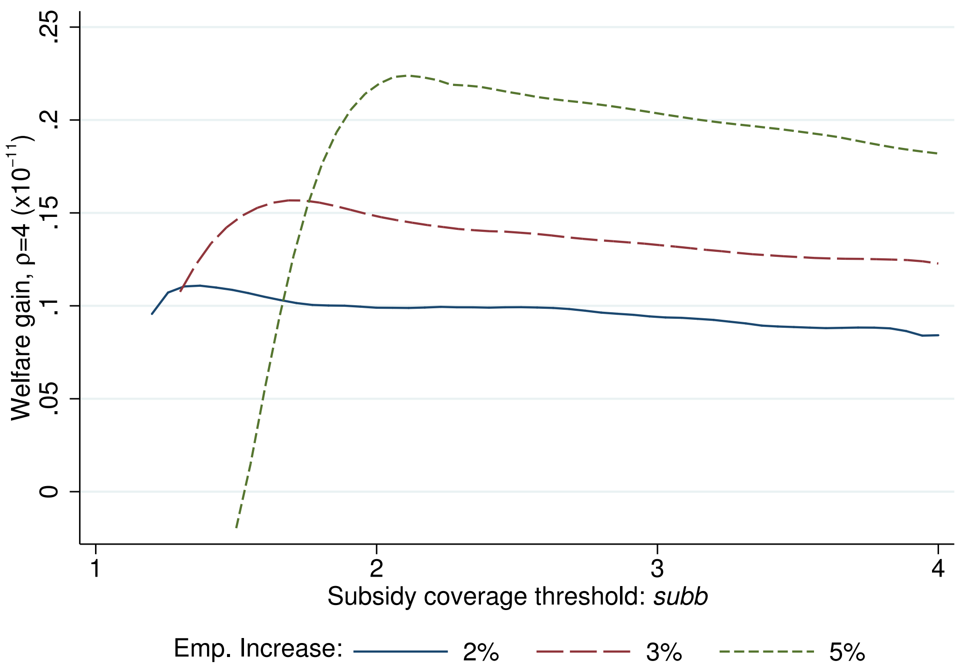
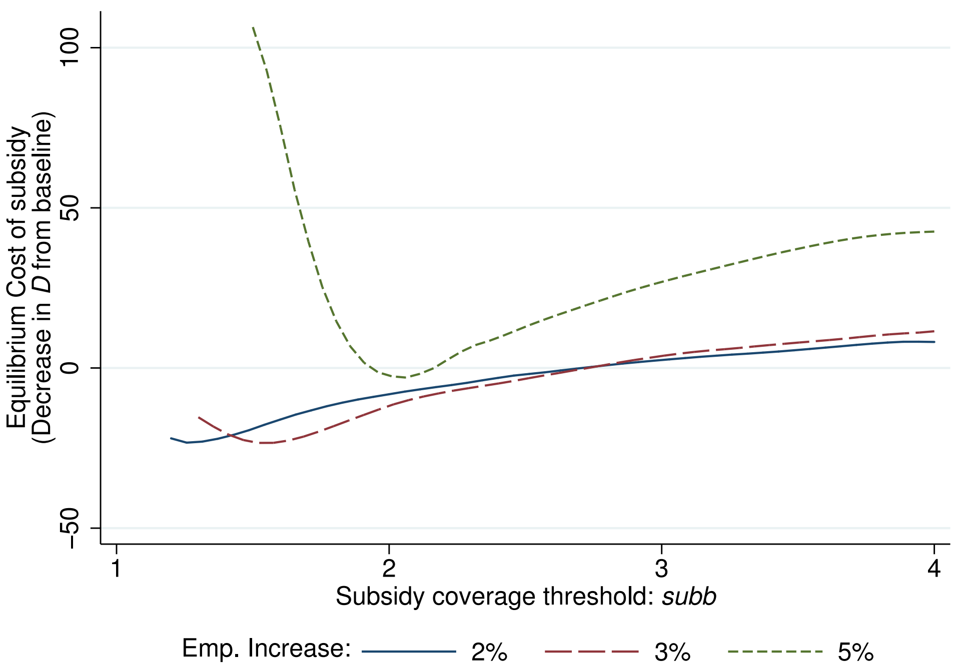


Figure 6.5: Equilibrium costs of subsidy programs measured as the decrease in tax revenue redistribution ($D_{baseline} - D$). We consider an inequality-averse welfare specification with $\rho = 4$. Subsidy generosity $suba$ varies such that the subsidy program raises baseline employment by the corresponding percentage.



methods to simulate the short-run behaviors of workers and firms. Based on our approximated transitional paths, we find that subsidy programs only achieve 50-65% of the equilibrium employment growth in the first year after implementation. Moreover, with subsidies applied to employers, workers may be strictly worse-off in the short run.

7 Conclusion

In this paper, we examine payroll tax reductions that target low-paid workers using an equilibrium search-and-matching model. We assume that the labor market is frictional and individuals and firms both have differing levels of productivity. We consider not only the effects of tax reductions on those directly targeted, but also the indirect effects on more productive individuals. We focus on labor markets that are constraint by a substantial minimum labor cost, which arises as a result of a high minimum wage and high taxes on low-paid work.

In our model, a low-productivity firm may not be able to form a viable match with a low-productivity worker because of the minimum labor cost constraint, while it may not be able to match with a high-productivity worker because of her good outside option. Reducing taxes for low-paid work expands job opportunities for low-productivity workers and allow them to earn higher wages. This induces higher labor force participation from these workers and more job creation from low-productivity firms. As different workers and firms are in competition for the opportunities to meet agents on the other side of the market, the search behavior creates congestion externalities that reduces the employment opportunities for the high-productivity workers.

Our model is particularly suited for studying payroll tax reductions in France, a country that features a high minimum wage and high taxes on low-paid work. Since the mid-1990s, France has implemented a series of tax reductions for low-paid work in an effort to open up employment opportunities for the low-skilled. These tax reductions also raise concerns about declining job productivities and the potential negative impacts on the employment opportunities for the high skilled workers.

To quantify the congestion effects, we estimate our model using French administrative data. We first simulate equilibrium effects of the Juppé tax reductions between 1995 and 1997. The overall effects on employment and aggregate production are positive and the magnitudes are in line with findings in the literature. However, there are large variations in these effect across different workers that are consistent with the congestion effects. The employment rate of highly productive workers decreases although their set of viable jobs has not significantly changed. The results indicate that highly productive workers find viable matches less frequently as a result of congestions from low-productivity job-seekers and vacancies. Moreover, since workers and firms are estimated to be complementary in production, the decrease in the employment of high-productivity workers is costly for aggregate production.

Which jobs be entitled to a payroll tax reduction? We consider tax reduction policies that offer the maximum reduction on minimum wage jobs and phase out at some coverage threshold; this type of policies encompasses the major tax reductions implemented in France. We also assume that tax revenue is redistributed in a lump-sum transfer. Our simulations indicate a clear trade-off between aggregate production and equality: we find that a tax reduction policy that is narrowly focused on minimum wage jobs is more likely to cause more severe congestion effects. By contrast, a narrow tax reduction is more effective in redistributing toward the less productive workers. To find the welfare-maximizing range of wages that ought to enjoy tax reductions, we adopt a social welfare criterion that includes a parameter for a more equal distribution of consumption. The key policy objective - increasing employment levels - calls for extending tax reductions to individuals with fairly high wages. To increase employment by 5% the optimal payroll tax reduction applies to jobs with wages up to twice the level of the minimum wage.

References

Aguiar, M., Bils, M., Charles, K. K., and Hurst, E. (2017). Leisure Luxuries and the Labor Supply of Young Men. NBER Working Papers 23552, National Bureau of Economic Research, Inc.

André, M., Bozio, A., Guillot, M., and Paul-Delvaux, L. (2015). French Public Finances through the Financial Crisis: It's a Long Way to Recovery. *Fiscal Studies*, 36:431–452.

Azmat, G. (2006). The Incidence of an Earned Income Tax Credit: Evaluating the Impact on Wages in the UK. CEP Discussion Papers dp0724, Centre for Economic Performance, LSE.

Bagger, J. and Lentz, R. (2014). An Empirical Model of Wage Dispersion with Sorting. NBER Working Papers 20031, National Bureau of Economic Research, Inc.

Bunel, M. and L'Horty, Y. (2012). The Effects of Reduced Social Security Contributions on Employment: An Evaluation of the 2003 French Reform. *Fiscal Studies*, 33(3):371–398.

Cahuc, P., Postel-Vinay, F., and Robin, J.-M. (2006). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, 74(2):pp. 323–364.

Chéron, A., Hairault, J.-O., and Langot, F. (2008). A quantitative evaluation of payroll tax subsidies for low-wage workers: An equilibrium search approach. *Journal of Public Economics*, 92(3-4):817–843.

Conseil d'Orientation pour l'Emploi (2013). Emplois Durablement Vacants et Difficultés de Recrutement. Technical report.

Crépon, B. and Desplatz, R. (2003). The Effets of Payroll Tax Subsidies for Low Wage Workers on Firms Level Decisions. Working Papers 2003-06, Centre de Recherche en Economie et Statistique.

Dey, M. S. and Flinn, C. J. (2005). An Equilibrium Model of Health Insurance Provision and Wage Determination. *Econometrica*, 73(2):571–627.

Flinn, C. J. (2006). Minimum Wage Effects on Labor Market Outcomes under Search, Matching, and Endogenous Contact Rates. *Econometrica*, 74(4):1013–1062.

Hotz, V. J. and Scholz, J. K. (2003). The Earned Income Tax Credit. In *Means-Tested Transfer Programs in the United States*, NBER Chapters, pages 141–198. National Bureau of Economic Research, Inc.

Jacquet, L., Lehmann, E., and Linden, B. (2014). Optimal income taxation with Kalai wage bargaining and endogenous participation. *Social Choice and Welfare*, 42(2):381–402.

Jardim, E., Long, M. C., Plotnick, R., van Inwegen, E., Vigdor, J., and Wething, H. (2017). Minimum Wage Increases, Wages, and Low-Wage Employment: Evidence from Seattle. NBER Working Papers 23532, National Bureau of Economic Research, Inc.

Leigh, A. (2010). Who Benefits from the Earned Income Tax Credit? Incidence among Recipients, Coworkers and Firms. *The B.E. Journal of Economic Analysis & Policy*, 10(1):1–43.

l'Haridon, O., Malherbet, F., and Pérez-Duarte, S. (2013). Does bargaining matter in the small firms matching model? *Labour Economics*, 21(C):42–58.

Lise, J., Meghir, C., and Robin, J.-M. (2013). Mismatch, Sorting and Wage Dynamics. NBER Working Papers 18719, National Bureau of Economic Research, Inc.

Lise, J. and Robin, J.-M. (2017). The Macrodynamics of Sorting between Workers and Firms. *American Economic Review*, 107(4):1104–1135.

Mabli, J. and Flinn, C. (2007). On-the-Job Search, Minimum Wages, and Labor Market Outcomes in an Equilibrium Bargaining Framework. 2007 Meeting Papers 791, Society for Economic Dynamics.

Mayneris, F., Poncet, S., and Zhang, T. (2016). Improving or Disappearing: Firm-Level Adjustments to Minimum Wages in China. Discussion Papers (IRES - Institut de Recherches Economiques et Sociales) 2016027, Université catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).

Moser, C. and Engbom, N. (2016). Earnings Inequality and the Minimum Wage: Evidence from Brazil. 2016 Meeting Papers 72, Society for Economic Dynamics.

Nichols, A. and Rothstein, J. (2015). The Earned Income Tax Credit. In *Economics of Means-Tested Transfer Programs in the United States, volume 1*, NBER Chapters, pages 137–218. National Bureau of Economic Research, Inc.

OECD (2016). 2015 Tax Burdens. In *Taxing Wages 2016*. OECD Publishing, Paris.

Riley, R. and Rosazza Bondibene, C. (2017). Raising the standard: Minimum wages and firm productivity. *Labour Economics*, 44(C):27–50.

Robin, J.-M. (2011). On the Dynamics of Unemployment and Wage Distributions. *Econometrica*, 79(5):1327–1355.

Rothstein, J. (2008). The Unintended Consequences of Encouraging Work: Tax Incidence and the EITC. Working Papers 1049, Princeton University, Department of Economics, Center for Economic Policy Studies.

Shephard, A. (2016). Equilibrium Search and Tax Credit Reform. Technical report.

Shimer, R. and Smith, L. (2001). Matching, Search, and Heterogeneity. *The B.E. Journal of Macroeconomics*, 1(1):1–18.

A Nash Bargaining

Consider the wage bargaining between an unemployed worker and a firm. The Nash bargaining wage maximizes the Nash product:

$$\phi_u^{Nash}(x,y) = \arg\max_w [W_e(w,x,y) - W_u(x)]^{\alpha_{Nash}} [J_f(w,x,y) - J_u(y)]^{(1-\alpha_{Nash})} \tag{20}$$

The Nash wage is characterized by the first order condition

$$W_e(w,x,y) - W_u(x) = \frac{\alpha_{Nash}}{1-\alpha_{Nash}} [J_f(w,x,y) - J_u(y)] \frac{\partial W_e/\partial w}{-\partial J_f/\partial w} \tag{21}$$

Without taxes, $\frac{\partial W_e/\partial w}{-\partial J_f/\partial w} = 1$, utility is perfectly transferrable between workers and firms, and thus the wage under Nash and that under our proportional bargaining scheme coincide ($\phi_u^{Nash}(x,y) = \phi_u(x,y)$ whenever $\alpha_{Nash} = \alpha$).[36] With taxes, the marginal tax rate matters.

Given values functions 6 and 7 and sets 10 and 11, derivatives can be written as

$$\begin{aligned} [r+\delta+s_1\kappa V]\frac{\partial W_e(w,x,y)}{\partial w} &= 1 \\ +s_1\kappa \frac{\partial \left[\int_{y_0(w,x,y)}^{y} W_e(\phi_e(x,y,y'),x,y) v(y') dy'\right]}{\partial w} & \\ +s_1\kappa \frac{\partial \left[\int_{y_l}^{y_0(w,x,y)} W_e(w,x,y) v(y') dy'\right]}{\partial w} & \end{aligned} \tag{22}$$

$$(r+\delta+s_1\kappa V)\frac{\partial J_f(w,x,y)}{\partial w} = -1 - \frac{dT(w)}{d}$$

$$+s_1\kappa \frac{\partial \left[\int_{y_0(w,x,y)}^{y} J_f(\phi_e(x,y,y'),x,y) v(y') dy'\right]}{\partial w} \tag{23}$$

$$+s_1\kappa \frac{\partial \left[\int_{y_l}^{y_0(w,x,y)} J_f(w,x,y) v(y') dy'\right]}{\partial w} \tag{24}$$

Applying the Leibniz integral rule, and noting that $\phi_e(x,y,y_0(w,x,y)) = w$, we get

[36] Consistent with l'Haridon et al. (2013) and Jacquet et al. (2014).

$$\frac{\partial W_e(w,x,y)}{\partial w} = \frac{1}{r+\delta+s_1\kappa\int_{y_0(w,x,y)}^{y_h} v(y')dy'} \tag{25}$$

and

$$-\frac{\partial J_f}{\partial w} = \frac{1+\frac{dT(w)}{dw}}{r+\delta+s_1\kappa\int_{y_0(w,x,y)}^{y_h} v(y')dy'} \tag{26}$$

The Nash equation can be rewritten as

$$\frac{W_e(w,x,y)-W_u(x)}{J_f(w,x,y)-J_u(y)} = \frac{\alpha}{[1-\alpha]}\frac{1}{\left[1+\frac{dT(w)}{dw}\right]} \tag{27}$$

which states that the ratio of the worker and firm surpluses is equal to the product of the ratio of their respective bargaining parameters and $\frac{1}{\left[1+\frac{dT(w)}{dw}\right]}$. Whether Equation 27 has a unique solution depends on how the marginal tax rate $\frac{dT(w)}{dw}$ varies with w. If the marginal tax rate is continuously increasing in w, a unique solution (interior or corner) exists. However, the French tax (SSC) schedule that we consider for our empirical section exhibits decreasing $\frac{dT(w)}{dw}$, and thus a unique Nash bargaining solution is not guaranteed. This poses theoretical and numerical challenges to solving the model and therefore we opt for the simpler proportional bargaining scheme that we describe in the main text.

By assumption, $dT(w)/dw \geq 0$, thus $\frac{1}{\left[1+\frac{dT(w)}{dw}\right]} \leq 1$. Given x and y, the Nash wage ϕ_u^{Nash} must be smaller than the proportionally bargained wage ϕ_u, implying that the α parameter we estimate must be smaller than the α_{Nash} if Nash bargaining were in place. The intuition is that, with Nash, the worker needs to compensate the firm knowing that increasing wage leads to increasing tax burden. In proportional bargaining, the two parties remain ignorant about how the tax burden comes about.

B Proofs

Proposition 2.

Proof. We first focus on W_e. Wage affects the value of employment both directly, through the current consumption, and in the continuation value, when the worker meets a poaching firm that cannot win the incumbent. The direct effect is necessarily a positive one as the marginal consumption equals the marginal increase in wage. We only need to examine the continuation effect. Consider an infinitesimal change in wage dw. At any $w > w_{min}$, we have:

$$\begin{aligned}
&(r+\delta+s_1\kappa V)\left[W_e(w,x,y)-W_e(w-dw,x,y)\right]=1\\
&+s_1\kappa\left[W_e\left(w,x,y\right)-W_e\left(\phi_e(x,y,y^*),x,y\right)\right]\int_{d\mathcal{A}_{e2}(w,x,y)} v(y')dy'
\end{aligned}$$

where y^* is some element of $d\mathcal{A}_{e2}(w,x,y) = \left\{y' : W_e(w-dw,x,y) < W_e(\bar{\phi}(x,y'),x,y') < W_e(w,x,y)\right\}$. Therefore, $\left[W_e\left(w,x,y\right)-W_e\left(\phi_e(x,y,y^*),x,y\right)\right] > 0$. This implies that $W_e(w,x,y) - W_e(w - dw,x,y) > 0$, and thus the partial derivative $\partial W_e(w,x,y)/\partial w > 0$.

A similar argument can be made for J_f. □

Corollary. *3*

Proof. Consider the case of wage bargaining between an unemployed worker and a firm. The bargained wage ϕ satisfies the following:

$$W_e(\phi,x,y)-W_n(x)=\frac{\alpha}{1-\alpha}[J_f(\phi,x,y)-J_u(y)]. \tag{28}$$

The left hand side $W_e(\phi,x,y) - W_n(x)$ monotonically increases in ϕ whereas $J_f(\phi,x,y) - J_u(y)$ is monotonically decreases. Moreover, as $\phi \to -\infty$, the left hand side goes to $-\infty$ whereas the right hand side goes to ∞. As $\phi \to \infty$, the left hand side goes to ∞ whereas the right hand side goes to $-\infty$. Therefore, there is always a unique solution Eq. 28. If the solution to Eq. 28 makes the net surplus S positive, a unique solution exists. Similar arguments can be made for the bargained wage between an employed worker and the winning firm of the firm competition. □

Proposition. *4*

Proof. In equilibrium, given that $J_u(y) = 0$, it must be that $J_f(\bar{\phi}(x,y),x,y) = 0$. Moreover, we know that $\mathcal{A}_{e2}(\bar{\phi}(x,y),x,y) = \emptyset$. Therefore, substituting w with $\bar{\phi}(x,y)$ in Eq. 7 gives

$$\bar{\phi}(x,y)+T(\bar{\phi}(x,y))=F(x,y) \tag{29}$$

Taking the derivative of Eq. 29 with respect to y, we have

$$\frac{\partial\bar{\phi}(x,y)}{\partial y}+T'(\bar{\phi}(x,y))\frac{\partial\bar{\phi}(x,y)}{\partial y}=\frac{\partial F(x,y)}{\partial y}$$

Rearranging, we have

$$\frac{\partial \bar{\phi}(x,y)}{\partial y} = \frac{\frac{\partial F(x,y)}{\partial y}}{1 + T'(\bar{\phi}(x,y))}$$

By assumption, $\frac{\partial F(x,y)}{\partial y} > 0$ and $T'(\bar{\phi}(x,y)) \geq 0$, thus $\frac{\partial \bar{\phi}(x,y)}{\partial y} > 0$. □

C Numerical Solution of Steady State Equilibrium

In this Section, we describe the procedure of the numerical solution. As we have explained, the surplus of every match in our model depends on current and future wages due to the fact that firms do not have full bargaining power. In addition, due to the presence of taxes and the minimum wage, we are unable to derive an analytical wage equation, and thus solving the model relies on iterating the value functions.

Before solving the model, we fix the exogenous components of the model and choose a tolerance level and a criterion function. We discretize the state space containing worker type, firm type, and wage into respective grids. We allow the grids for firm type and wage are to depend on model parameters for numerical efficiency.

We make initial guesses for the value functions W_e, W_n, and J_f such that W_e and J_f are increasing in wage. The initial guess for the match and unemployment distributions $h(\cdot,\cdot)$ and $u(\cdot)$ are such that the sum of them across all worker and firm times is equal to 1. We also make initial guess for the vacancy distribution $v(\cdot)$ such that the sum of vacancies is greater than 0. Given the initial guesses, we enter the loop for the fixed point solution. In each iteration, we take the $\tilde{W}_e, \tilde{W}_n, \tilde{J}_f$, $\tilde{h}$, $\tilde{u}$, and $\tilde{v}$ as given and solve for the optimal decisions. The tilde-notation refers to the initial guesses if we are in the first iteration, otherwise, it refers to the resulting objects from the previous iteration.

More specifically, each iteration can be broken down into several steps:

1. Given $\tilde{W}_e, \tilde{W}_n, \tilde{J}_f$, we solve for the set of viable matches, $\tilde{\Omega}$, such that

$$\tilde{\Omega} = \{(x,y) : \exists w \text{ s.t. } w \geq w_{min} \text{ and } \tilde{W}_e(w,x,y) - \tilde{W}_n(x) \geq 0 \text{ and } \tilde{J}_f(w,x,y) \geq 0\}$$

2. Solve for $\tilde{\phi}_u(x,y)$ for all $(x,y) \in \tilde{\Omega}$.

3. Define $\tilde{\bar{\phi}}(x, y)$ as the highest wage such that $\tilde{J}_f(w, x, y) = 0$. Solve for $\tilde{\phi}_e(x, y', y)$ for all (x, y', y) such that $(x, y) \in \tilde{\Omega}$ and $(x, y') \in \tilde{\Omega}$ from the following equation.

$$\tilde{W}_e(\tilde{\phi}_e(x, y', y), x, y') - \tilde{W}_e(\tilde{\bar{\phi}}(x, y), x, y) = \frac{\alpha}{1-\alpha}\tilde{J}_f(\tilde{\phi}_e(x, y', y), x, y')$$

4. Define $mobility(x, y', y) = 1$ if either of the following criteria is satisfied.

   (a) $(x, y) \in \tilde{\Omega}$, $(x, y') \in \tilde{\Omega}$, $\tilde{\phi}_e(x, y', y) \leq \tilde{\bar{\phi}}(x, y')$, and $\tilde{W}_e(\tilde{\bar{\phi}}(x, y'), x, y') - \tilde{W}_e(\tilde{\bar{\phi}}(x, y), x, y) \geq 0$.

   (b) $(x, y) \notin \tilde{\Omega}$ but $(x, y') \in \tilde{\Omega}$.

5. Solve for $\tilde{s}(x)$ for all x, such that

$$\tilde{s}(x) = \arg\max_{s=\{0,1\}} \left\{ B(x) - sq + s\tilde{\kappa} \int_{y' \in \tilde{\mathcal{A}}_u(x)} [\tilde{W}_e(\max\{w_{min}, \tilde{\phi}_u(x, y')\}, x, y') - \tilde{W}_n(x)]\tilde{v}(y')dy' \right\}$$

   where $\tilde{\kappa} = \frac{M(\tilde{\xi}, \tilde{V})}{\tilde{\xi}\tilde{V}}$, and $\tilde{\mathcal{A}}_u(x) = \left\{ y : (x, y) \in \tilde{\Omega} \right\}$.

6. Update the efficiency search units: $\tilde{\xi} = \int \left[\tilde{s}(x)\tilde{u}(x) + \int \tilde{h}(x, y)dy \right] dx$.

7. Solve for $\tilde{v}(\cdot)$ using Eq. 9, with $I_u(y)$ and $I_e(y)$ being replaced by their tilde-counterparts:

$$\begin{aligned} \tilde{I}_u(y) &= \int_{x \in \tilde{\mathcal{B}}_u(y)} \tilde{s}(x)\tilde{J}_f(\tilde{\phi}_u(x, y), x, y)\tilde{u}(x)dx \\ \tilde{I}_e(y) &= \iint_{(x, y') \in \tilde{\mathcal{B}}_e(y)} s_1 \tilde{J}_f(\tilde{\phi}_e(x, y, y'), x, y)\tilde{h}(x, y')dy'dx \end{aligned}$$

   where $\tilde{\mathcal{B}_u}(y) = \left\{ x : (x, y) \in \tilde{\Omega} \right\}$ and $\tilde{\mathcal{B}_e}(y) = \{(x, y') : mobility(x, y, y') = 1\}$.

8. Update $\tilde{\kappa} = \frac{M(\tilde{\xi}, \tilde{V})}{\tilde{\xi}\tilde{V}}$, and value functions the value functions $\tilde{W}_e, \tilde{W}_u, \tilde{J}_f$ using Eq. 6, 5, and 7 and the appropriate tilde-objects.

9. Update the unemployment distribution $\tilde{u}(\cdot)$ using Eq. 13, and the match distribution $\tilde{h}(\cdot, \cdot)$ by using Eq. 12.

10. Evaluate the criterion function and compare the value with the pre-set tolerance level. If the distance is within tolerance, terminate the loop.

D Imputing the status of individuals in gap spells

We use the *Enquête Emploi* (Hereafter, EE), French labor force survey, to impute the status of an individual in a gap spell in the DADS. Within EE , we label all spells that are not covered by the DADS panel as "not employed", with the indicator nw. This include in particular unemployment, but also self-employment and non-participants. The aim is to identify the probability of unemployment conditional on non-employment using individual and job characteristics that are available in both EE and DADS.

The first step is to select an EE sample to resemble the sample in DADS. This entails restricting to men of age 30-55 and dropping individuals who have never been employed prior to or following an nw spell. The latter restriction is related to the data structure in the DADS panel, in which a gap spell can only be observed if it is sandwiched between two employment spells. We also drop nw spells that last for more than 3 years.

We then estimate the likelihood of unemployment in EE. We use information on the individual's age, the duration of the nw spell, the social-professional status, industry, and sector (private or public) of the employment spell following the nw spell. We denote these information by Ω_s. Using a Probit model, we estimate $P(u_s|nw_s, \Omega_s)$, where $u_s = 1$ indicates unemployment.

The final step is to impute the unemployment status for gap, or nw, spells in DADS. We similarly construct Ω_s^{DADS} for each spell s, and compute the predicted likelihood that s is an unemployment spell using the estimated predictor from EE, $\hat{P}(u_s|nw_s, \Omega_s^{DADS})$. We draw the unemployment status of each nw spell from the distribution given by the predicted likelihood.

E Simulating non-employment benefits.

We simulate the benefit level, denote by $\tilde{B}$, as a function of the average daily gross wage $\tilde{w}$ in the year preceding the unemployment spell. Specifically, $\tilde{w}$ is equal to the total gross earnings during the preceding year divided by the number of days worked in that year. The procedure to compute $\tilde{B}$ is as follows:

1. compute $\tilde{B}_0(\tilde{w}) = \max\left\{\tilde{f} + \tilde{s}_0\tilde{w}, \tilde{s}_1\tilde{w}\right\}$;
2. compute $\tilde{B}_1(\tilde{w}) = \max\left\{\tilde{B}_0(\tilde{w}), \tilde{m}\right\}$;
3. if $\tilde{B}_1(\tilde{w}) = \tilde{m}$, the simulated benefit $\tilde{B} = \tilde{m}$. Otherwise, $\tilde{B} = \min\left\{\tilde{B}_0(\tilde{w}), \tilde{s}_2\tilde{w}\right\}$.

Table E.1: Values of the policy parameters $\tilde{f}$ and $\tilde{m}$ for simulating non-employment benefits. Values are nominal. Values prior to 2001 have been converted from French francs (FF) to Euros (€) using the conversion rule of 1€=6.55957FF.

Date effective	$\tilde{f}$	$\tilde{m}$	Date effective	$\tilde{f}$	$\tilde{m}$
7/1/10	11.17 €	27.25 €	7/1/00	9.56 €	23.32 €
7/1/09	11.04 €	26.93 €	7/1/99	9.38 €	22.86 €
7/1/08	10.93 €	26.66 €	7/1/98	9.26 €	22.58 €
7/1/07	10.66 €	26.01 €	7/1/97	9.09 €	22.16 €
7/1/06	10.46 €	25.51 €	7/1/96	8.90 €	21.68 €
7/1/05	10.25 €	25.01 €	7/1/95	8.68 €	21.17 €
7/1/04	10.25 €	25.01 €	7/1/94	8.43 €	20.39 €
7/1/03	10.15 €	24.76 €	7/1/92	8.26 €	19.97 €
7/1/02	9.94 €	24.24 €	7/1/91	8.04 €	19.45 €
7/1/01	9.79 €	23.88 €	10/1/90	7.87 €	19.02 €

The parameters $\tilde{f}$, $\tilde{m}$, $\tilde{s}_0$, $\tilde{s}_1$, and $\tilde{s}_2$ are policy parameters. $\tilde{f}$ and $\tilde{m}$ are time-varying, whose values are shown in Table E.1. The values of $\tilde{s}_0$, $\tilde{s}_1$, and $\tilde{s}_2$ are fixed in the entire sample period from 1991 to 2008, with $\tilde{s}_0 = 40.4\%$, $\tilde{s}_1 = 57.4\%$, and $\tilde{s}_2 = 75\%$.

F Panel data from the DADS

This section provides details on data cleaning procedures in dealing with the combined dataset from the two panels from the DADS, *panel DADS* and *panel tous salariés*. In Section F.1, we describe the procedures of converting the spell-based data in the original panel to monthly-based sample. In Section G, we explain sample restrictions and the calculation of individual ranking statistics.

F.1 Procedures to convert data into monthly data

The raw data is spell-based; there is one observation per individual-job-year. We took the following steps to convert the raw data into a monthly dataset.

F.1.1 Correcting missing spell dates.

Around 0.5% of employment spells contains missing start and end dates; the spell duration is available for over 99.998% of the spells. We infer the spell start and end dates using spell duration and the employment spells in the surrounding years. Let $spell(i, Y, j)$ denote an employment spell of worker i in year Y and firm j. Suppose we observe $spell(i, Y, j)$ with missing dates, and we also observe $spell(i, Y + 1, j)$ that starts on the first day of year $Y + 1$, and we do not observe $spell(i, Y - 1, j)$. In this case, the end date of $spell(i, Y, j)$ is the last day of year Y, and the start date is derived from the spell duration. In all other cases, we assume that the spell start date is day 1 of the spell year, and the end date is derived from spell duration. In the extremely rare cases that the spell duration is missing, we assume that the spell lasts for the entire year.

F.1.2 Correcting overlapping spells.

Multiple spells of the same worker at the same or different firms may have overlaps in time. About 40% of the individuals have held overlapping jobs. In these cases, we need to identify a main job and define the wage for the job. During the time window that two jobs overlap, the main job is the one that is full-time, private sector, and non-executive. If both or neither jobs satisfy these criteria, the main job is identified by a higher wage. Wages from overlapping jobs are only summed if they are in the same firm. Lastly, continuous employment spells within the same firm in a given year are concatenated and the wage is defined as the average wage over the concatenated spell.

F.1.3 Correcting whole-year gaps.

We notice that in years 1994, 2003, and 2005, there are high occurrences that individuals are missing for the entire year but are observable in the preceding or the following years; we refer to this as a whole-year gap. Over the period between 1991 and 2008, whole-year gaps occurs in 1.4% of the sample individuals. In 1994, 2003 and 2005, the occurrences are 10.3%, 3.0% and 1.4% respectively. A potential reason for the whole-year gaps may be missing data for these individuals in the three years. To correct for this problem, we replace the whole year gaps with employment spells if the worker is employed on the day before and after the gap year in the same firm. We take the average wages in the surrounding years as the wage for the new employment spells. Overall, 86.6% of the whole-year gaps in the three years are corrected.

F.1.4 Transforming spell data to monthly data

In the monthly data, there is one observation per individual-month. If more than one spells occupy the same month, we take the one that occupies the largest fraction of the month.

Table G.1: Descriptive Statistics. "DADS" refers to the merged DADS panels restricted to men age 30-55 between 1991-2008. "Sample" refers to our final sample after cleaning and restricting the panel. Sample duration is calculated as the difference between the start date of the first employment spell and the end date of the last employment spell. A sample job is one that is full-time, private-sector, and non-executive. Daily wage is the daily net wage.

	Data	Restricted sample
# Individuals	873,425	416,221
Mean sample duration	2122 days	3754 days
Median sample duration	1229 days	2879 days
% Full time jobs	84.01%	93.45%
% Private Sector jobs	69.32%	96.74%
% Non-executive jobs	74.28%	96.46%
% Sample jobs	44.51%	88.26%
25th daily wage percentile	46.43	45.68
50th daily wage percentile	57.47	55.18
75th daily wage percentile	73.11	68.26

G Ranking workers

In constructing individual statistics for ranking, we set two additional sample selection criteria. First, we exclude individuals whose sample duration is less than 5 years, where the sample duration is calculated as the difference between the start date of the first employment spell and the end date of the last employment spell. Second, we exclude individuals such that less than 50% of the sample duration is occupied by sample jobs or unemployment, where sample jobs are defined as full-time non-executive jobs in the private sector. Our final sample contains 416,221 men of age 30-55 between 1991-2008 who satisfy the sample selection criteria. Table G.1 compares the raw data from DADS and our final sample. As expected, individuals in the final sample have longer sample duration, and are more likely found in sample jobs. Since we exclude individuals who mainly work as executives, the average sample job wage is lower in our sample.

In computing the ranking statistics of an individual, we include all labor incomes including those from overlapping jobs and non-sample jobs. All income measures are net of taxes. The lifetime minimum and maximum wage are annualized wages in any job divided by 12. Since part-time jobs are included in these statistics, the lifetime minimum monthly wage may fall below the legal minimum wage.

Table G.2: Ranking statistics of workers. All statistics are monthly. For each individual, the statistics are computed based on the entire period between the first and last observations of the individual in the DADS sample from 1991 to 2008. "Lifetime income" refers to the net-of-tax average income per month, accounting for net wage and imputed non-employment benefits. "Lifetime min. wage" and "Lifetime max. wage" refer to the lowest and the highest net wage from employment that an individual obtains while in sample.

(a) Distributions of worker ranking statistics.

	Lifetime income	Lifetime min. wage	Lifetime max. wage
Mean	1671.3	1140.6	2790.6
P25	1260.9	556.2	1690.2
P50	1549.2	1182.9	2084.4
P75	1950.0	1590.6	2733.0

(b) Spearman rank correlations between worker ranking statistics.

	Lifetime income	Lifetime min. wage	Lifetime max. wage
Lifetime income	1	0.754	0.653
Lifetime min. wage	-	1	0.298

* all correlations are significant at the 1% level.

Table G.2a shows the distributions of the three ranking statistics, average individual earnings and lifetime maximum and minimum wages. G.2b shows the Spearman rank correlations between the statistics. Lifetime earnings is strongly correlated in rank with the lifetime minimum and maximum wages; the rank correlation between the latter is weaker but nevertheless significantly positive. The weaker correlation may be due to measurement and sampling errors, which pose a greater problem for the extrema measures.

G.1 Firms

We compute firm size from POST by counting the total number of employee-days divided by the total number of days a firm is in the sample. In the computation, we only consider jobs that are non-executive, full-time, in the private sector, and are filled by individuals who are male, age 30-55. Table G.3 shows the firm size distribution and the average sample duration for firms of different sizes. On average, larger firms have longer sample durations. In our estimation procedure, we only rank firms with a firm size of at least 10. For these firms, the average sample duration is over 10 years. Moreover, although only 29% of firms satisfy our criterion, they account for 88% of total employment.

Table G.3: Firm size distribution, firm duration in sample and employment share. The statistics are computed from the POST dataset of the DADS. Firm size is computed by counting the total number of employee-days divided by the total number of days a firm is in the sample. We restrict to jobs that are non-executive, full-time, in the private sector, and are filled by individuals who are male, age 30-55.

Firm size	Number of Firms	Fraction of firms	Firm duration in sample	Emp. Share
0 to 1	26,972	71.0%	4.68	12.2%
1 to 2	44,407		8.00	
3 to 5	73,717		9.99	
5 to 10	49,319		10.93	
10 to 50	63,617	29.0%	11.21	87.8%
50-100	8,473		10.93	
>100	7,311		11.16	
Total	273,816	100%	-	100%

H The short-run

In this section, we develop methods to approximate short-run labor market distributions by assuming that workers and firms do not perfectly correctly foresee the future. We compute social welfares at the time of policy implementation. When the time horizon for welfare calculations is sufficiently short, the conclusion we draw may be contrary to the one drawn from the steady state equilibrium.

H.1 Approximation methods for the short-run transitions.

The labor market is no longer in an equilibrium when a payroll tax reduction is implemented. Because of search frictions, the labor market distributions of unemployment, $u(\cdot)$, and matches, $h(\cdot,\cdot)$, take time to adjust. On the transitional path, the decisions $\{s(\cdot), \underline{y}(\cdot), \phi_u(\cdot,\cdot), y_0(\cdot,\cdot,\cdot), \phi_e(\cdot,\cdot,\cdot), v(\cdot)\}$ depend on future distributions.[37] For example, in order to decide whether or not to search, a non-employed worker needs to know not only the likelihood of finding a job in the current period, but also the likelihoods of receiving outside offers when she becomes employed. Similar examples can be made with firms' decisions. Making such decisions requires the knowledge of the current and future distributions of vacancies, unemployment and employment. However, as a result of the payroll

[37] If we can write down a surplus function that does not depend on future wages, solving for the decisions would not require knowing the future distributions, and consequently the system would be tractable in the off-equilibrium dynamics. This is the case in Lise and Robin (2017), which studies a search and matching model with aggregate shocks. In our model, we cannot achieve tractability if we let $\alpha > 0$, or if there were taxes and minimum wages.

tax reduction, these distributions change from period to period. The question thus arises, which distributions our worker uses to determine her search effort. Solving the model at every point in time on the transitional path is not feasible because the distributions are intractable. We believe that a good approximation for the transitional path is to assume that workers and firms misperceive the future with what they can observe around them and what they know about the steady state equilibrium with the new policies.

Let $D = \{u(\cdot), h(\cdot, \cdot)\}$ be the steady state distributions in the "old regime" prior to a policy reform. At $t = 0$, a reform is implemented. We assume the reform is unexpected, so that agents do not react the impending reform prior to $t = 0$. Let $D_t = \{u_t(\cdot), h_t(\cdot, \cdot)\}$ be the actual distributions of unemployment and matches in period $t > 0$, and $D' = \{u'(\cdot), h'(\cdot, \cdot)\}$ be the steady state distribution in the "new regime" under the policy reform. Instead of being correctly informed about D and $\{D_t\}_{t>0}$, we assume that workers and firms are misinformed but nevertheless are forward-looking and rational given their information. This allows us to simplify the model solution on the transitional path and allow us the simulate the short-run dynamics of then labor market distributions.

We consider two alternative assumptions: far-sightedness and short-sightedness. Let D_t^F and D_t^S denote agents' perception about the distributions in period $t > 0$ under the two far-sightedness and short-sightedness assumptions, respectively, and let Ψ_t^F and Ψ_t^S be the decisions in period $t > 0$.

Under the far-sightedness assumption, workers and firms believe that they are already in the new steady state equilibrium such that $D_t^F = D'$ for all $t > 0$. They follow the same decision rules under the new steady state such that $\Psi_t^F = \Psi'$ for all $t > 0$.

Under the short-sightedness assumption, at each period t, agents observe the actual distributions D_t. They perceive D_t to be the distributions that will last in all future periods, i.e. $D_{t+d}^S = D_t$ for all $d \geq 0$. They make optimal decisions Ψ_t^S consistent with a steady state equilibrium that features these distributions. Along the transitional path, agents' perceptions D_t^S are updated every period as D_t changes.

While we do not prove the convergence of the approximated transitions analytically, in the simulation exercises we conduct, the approximated transitional paths under the two assumptions always converge to the new regime steady state distributions. The discrepancy between the approximated paths and the actual path depends on the difference between the immediate impact and the new-regime equilibrium.

In addition to the search friction that is inherent to our model, there may be additional frictions in labor market transitions that only apply when policies are altered. In the steady state equilibrium, individual do not transition in and out of the labor force or transition from jobs to unemployment voluntarily. Therefore, we have not made any assumption about the speed or costs associated with

Table H.1: Parameters of three subsidy programs. "Emp. Goal" refers to the equilibrium increase in employment from the pre-reform level. *suba* and *subb* are parameters of the subsidy function in Eq. 16.

Subsidy	Emp. Goal	Generosity (*suba*)	Coverage (*subb*)
1	2%	0.34	1.35
2	3%	0.24	1.7
3	5%	0.30	2.1

such transitions. However, when the policy environment changes, such transitions are part of the transition process to the new steady state.[38]

We consider the possibility of a labor force entry (LFE) friction: Individuals have to wait before entering the labour market. We simulate the transitions both with and without the LFE friction. In the case with the LFE friction, we assume that the waiting time distribution is stochastic, and is on average 1 year. LFE friction may arise from the fact that there may be a waiting period for the non-participants to become eligible for jobs or for the worker to regain job search skills. It may also be the case that individuals are not aware of a tax reduction program immediately after its implementation.

H.2 Short-Run Simulation Results

In subsection 6.2, we maximize the equilibrium social welfare $\mathcal{W}$ by varying the broadness and generosity of payroll subsidy programs that raise employment to fixed goals. Table H.1 shows the subsidy parameters of the optimal programs that raise the baseline employment rate by 2, 3, and 5%. In this Subsection, we simulate the short-run employment and welfare transitions from the baseline to the steady state equilibria under the three subsidy programs.

We first simulate short-run transitions of employment based on our two approximation methods. A slow employment response may raise concerns for the efficacy of a subsidy. It may also explain the short-run unresponsiveness found in the empirical literature. Figure H.1a shows the transitional paths. A first observation is that the transitional paths under the far- and short-sightedness approximations closely trace one another. All simulated short-run transitions converge to the correct

[38]Note that we do not need to consider friction associated with job destructions and downward wage adjustments because we only consider tax reduction policies. If tax increases were being considered, it may be important to accout for frictions in those process, particularly in a country with strong labor protection.

steady state equilibrium. These provide assurance for our approximation methods and leave little room for the possibility for the multiplicity of transitional paths.

To quantify the speed of convergence, Table H.2 shows the completion of the employment goals in the short-run. The convergence to the new equilibrium employment is slow: without the LFE friction, only 50-65% of the employment goals is completed at the end of the first year after a subsidy is implemented. This one-year completion rate drops to 40% with the LFE friction. The number of new labor force entrants and whether these workers face the LFE friction are important determinants of the speed of employment growth. Of the three subsidy programs we focus on, the most ambitious program with a 5% employment target involves the most labor force entry. Consequently, without the LFE friction, it generates the fastest employment growth, both in terms of the completion rate and in absolute terms. The LFE friction also most strongly slows down the employment growth in this subsidy program.

The results imply that policy evaluations conducted shortly after the implementation of a tax reform would significantly under-estimate the long-term effects on employment. The short-run bias is stronger if non-participants face the LFE friction.

Next, we turn to the short-run welfare of individuals. Consistent with the equilibrium simulations, we assume that all tax revenues are redistributed, but firm profits are not redistributed to individuals and the policy maker only concerns with the welfare of individuals.[39] Furthermore, we assume that payroll subsidies are applied to employers, and they do not adjustment wages of existing jobs unless workers can make a credible threat to separate from the match. This implies that firms may benefit from a payroll subsidy in the short-run, allowing them to make a positive profit. The profits eventually disappear as existing jobs are destroyed either exogenously or as workers receive good outside offers that lead to wage renegotiations or job-to-job transitions. Given our redistribution assumptions, workers may suffer a short-run welfare loss even if they benefit from a welfare gain in the long run.

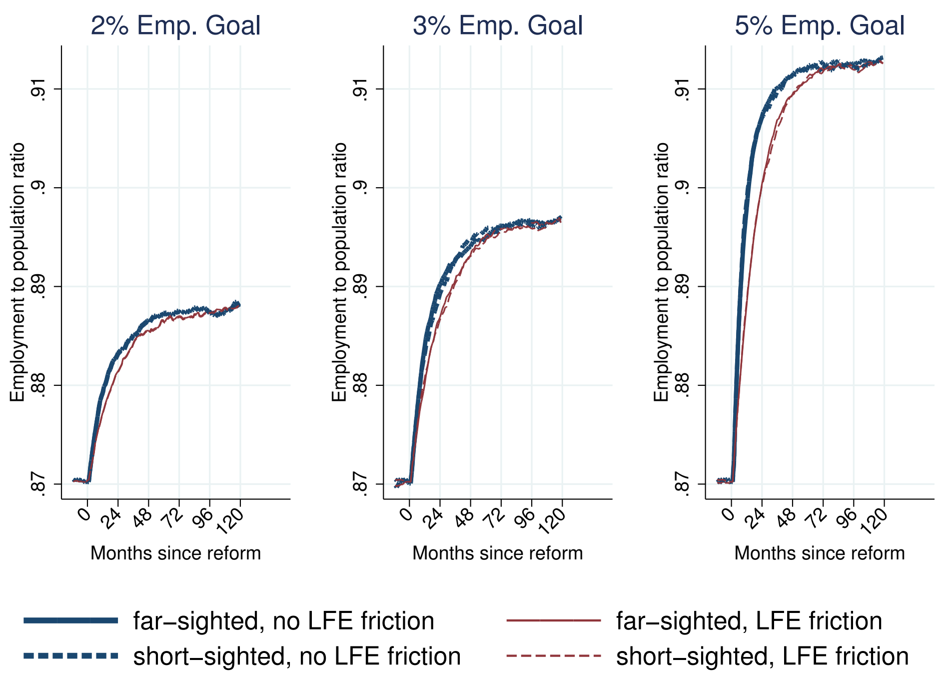
More specifically, we denote the per-period social welfare by $\mathcal{W}_t$:

$$\mathcal{W}_t = \frac{1}{N_i} \sum_{i=1}^{N_i} welfare(c_{i,t}) \tag{30}$$

[39]These are plausible assumptions because in our model firms do not invest in capital or other productive inputs.

Figure H.1: Simulated short-run transitions of welfare and employment. The simulations correspond to the three subsidy programs shown in Table H.1. "Emp. Goal" refers to the equilibrium increase in employment from the pre-reform level. "Far-sighted" and "short-sighted" refer to the two approximation methods. "LFE friction" is the labor force entry friction. In simulations with LFE friction, non-participants who wish to enter the labor force can do so with probability 1/12 per month.

(a) Employment to population ratio.



(b) Transition of social welfare of individuals ($\mathcal{W}_t$, $\rho = 4$).

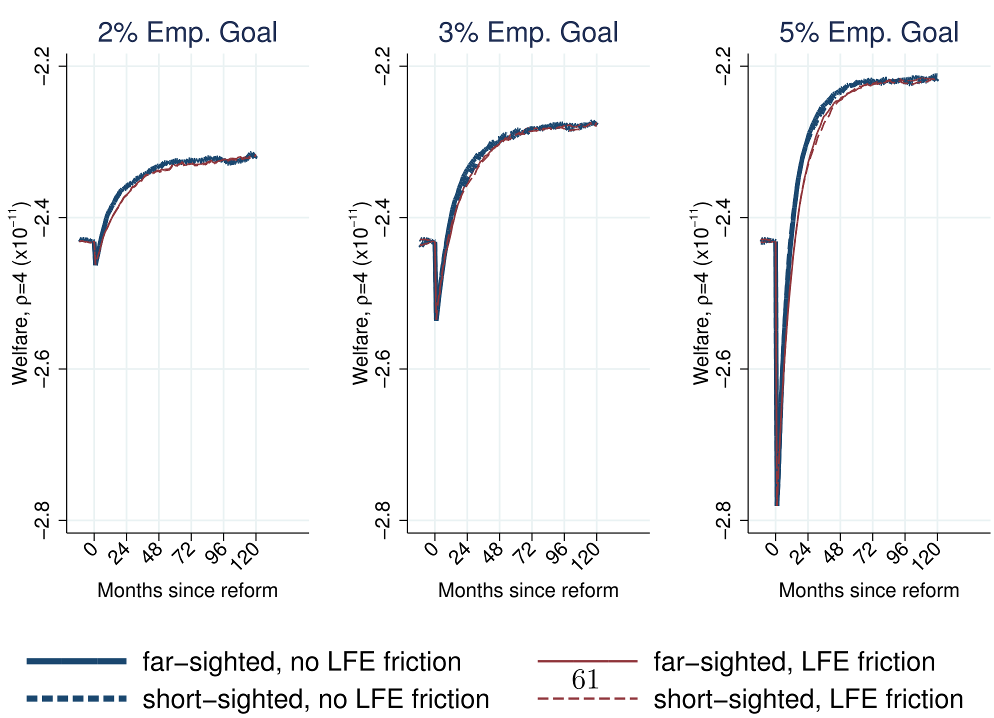


Table H.2: Simulated short-run employment rate based on far-sightedness approximation. Parameters of the subsidy programs are shown in Table H.1. "Emp. Goal" refers to the equilibrium increase in employment from the pre-reform level. "Completion of Emp. Goal" is the percentage of the employment goal that has been completed since the subsidy implementation at the end of each year. "Cum. Emp. Growth" is the cumulative employment growth from the baseline level. "LFE friction" is the labor force entry friction. In simulations with LFE friction, non-participants who wish to enter the labor force can do so with probability 1/12 per month.

(a) Subsidy 1: 2% Emp. Goal

	No LFE Friction		With LFE Friction	
Year	Completion of Emp. Goal	Cum. Emp. Growth	Completion of Emp. Goal	Cum. Emp. Growth
1	53.1%	1.1%	41.2%	0.8%
2	76.2%	1.5%	65.4%	1.3%
3	86.7%	1.7%	81.9%	1.6%
4	95.4%	1.9%	89.3%	1.8%
5	99.5%	2.0%	94.6%	1.9%

(b) Subsidy 2: 3% Emp. Goal

	No LFE Friction		With LFE Friction	
Year	Completion of Emp. Goal	Cum. Emp. Growth	Completion of Emp. Goal	Cum. Emp. Growth
1	53.2%	1.6%	38.5%	1.2%
2	77.4%	2.3%	64.6%	1.9%
3	86.7%	2.6%	79.1%	2.4%
4	92.2%	2.8%	88.9%	2.7%
5	95.5%	2.9%	96.6%	2.9%

(c) Subsidy 3: 5% Emp. Goal

	No LFE Friction		With LFE Friction	
Year	Completion of Emp. Goal	Cum. Emp. Growth	Completion of Emp. Goal	Cum. Emp. Growth
1	64.2%	3.2%	41.8%	2.1%
2	85.5%	4.3%	69.8%	3.5%
3	92.1%	4.6%	84.8%	4.2%
4	95.1%	4.7%	90.7%	4.5%
5	96.9%	4.8%	94.1%	4.7%

where D_t is the amount of tax redistribution (Equation 15) and $welfare(\cdot)$ is the individual welfare function (Equation 19). Figure H.1b shows $\mathcal{W}_t$ in the short run, with inequality-averse specification with $\rho = 4$. In our simulations, the initial drop in welfare is the largest in the most ambitious subsidy program that raises baseline employment by 5%.. This is because the more ambitious the subsidy program is in raising the employment goal, the more strongly and widely existing jobs are affected. As firms make profits from the payroll subsidies, workers suffer from a lower tax redistribution (D). This not only lowers the total consumption, but also has negative distributional impacts.

As a result of finite term limits of political offices and the constant evolution of public policies, policy makers may place a greater weight on the current period than the far future, and they may have a finite horizon when it comes to policy decisions. We have seen that the initial welfare response can be qualitatively different from the long-run effects, the time horizon of the policy maker may alter the her preference over alternative subsidy programs.

We define the continuation welfare $\mathcal{W}^c$ as the sum of discounted future welfare evaluated at $t = 1$, the period that a subsidy program is implemented. More specifically,

$$\mathcal{W}^c = \frac{(1 - \beta^T)}{1 - \beta} \sum_{t=1}^{T} \beta^{t-1} \mathcal{W}_t$$

where T is the time horizon of the policy maker. We assume that the discount rate of the policy maker coincides with that of agents in the labor market. Table H.3 shows the simulated continuation social welfare under different time horizons. Of the three subsidy programs we consider, the program that leads to a 5% increase in equilibrium employment results in the highest equilibrium welfare. However, the short-run welfare implications is different. Under the inequality-averse welfare specification with $\rho = 4$, we find that when the policy maker's horizon is 2 years, the most moderate subsidy program is preferred. With a horizon of 5 years, the most ambitious program is preferred if there were no LFE friction. With the friction, convergence of the welfare to the long-run level is slower.

Note that, if the subsidies were given to employees as in the EITC or the WFTC, the short-run welfare loss would not occur. In this case, the welfare transition would follow a similar path as the employment transition - welfare gradually increases as less-productive workers find jobs.

Table H.3: Simulated continuation welfare ($\mathcal{W}^c$ with $\rho = 4$) with different time horizons, inflated by a multiple of 10^{11}. Bolded value indicate the optimal amongst the three employment targets considered. The simulations are based on the far-sightedness approximation. Parameters of the subsidy programs are shown in Table H.1. "Emp. Goal" refers to the equilibrium increase in employment from the pre-reform level.

Subsidy	Emp. Goal	2 Year Horizon	5 Year Horizon	20 Year Horizon
		No LFE Friction		
1	2%	**-2.396**	-2.364	-2.336
2	3%	-2.413	-2.353	-2.304
3	5%	-2.460	**-2.338**	**-2.259**
		With LFE Friction		
4	2%	**-2.408**	-2.372	-2.340
5	3%	-2.426	**-2.362**	-2.307
6	5%	-2.496	-2.363	**-2.267**