

ORIGINAL ARTICLE

Open Access



# Self-employment in an equilibrium model of the labor market

Jake Bradley

Correspondence: [jb683@cam.ac.uk](mailto:jb683@cam.ac.uk)  
Faculty of Economics, University of  
Cambridge, Sidgwick Ave, CB3 9DD  
Cambridge, UK

## Abstract

Self-employed workers account for between 8 and 30 % of participants in the labor markets of OECD countries (Blanchower, Self-employment: more may not be better, 2004). This paper develops and estimates a general equilibrium model of the labor market that accounts for this sizable proportion. The model incorporates self-employed workers, some of whom hire paid employees in the market. Employment rates and earnings distributions are determined endogenously and are estimated to match their empirical counterparts. The model is estimated using the British Household Panel Survey (BHPS). The model is able to estimate nonpecuniary amenities associated with employment in different labor market states, accounting for both different employment dynamics within state and the misreporting of earnings by self-employed workers. Structural parameter estimates are then used to assess the impact of an increase in the generosity of unemployment benefits on the aggregate employment rate. Findings suggest that modeling the self-employed, some of whom hire paid employees implies that small increases in unemployment benefits leads to an expansion in aggregate employment.

**JEL Classification:** J21, J24, J28, J64

**Keywords:** Self-employment, Job search, Firm growth

## 1 Introduction

The proportion of total employment made up by the self-employed in the UK rose steadily over the period 2000–2004. The number of self-employed increased by 8.9 % compared with an increase of 0.1 % of paid employees. This growth was across gender, region, and industry (Lindsay and Macauley 2004). Over this period, the average self-employment rate was at 11.5 % of total employment and this large proportion is not unique to the UK: across all OECD countries this proportion varied from 8 to 30 % (Blanchflower 2004). Overall employment, wage determination and dynamics across the two sectors are clearly intrinsically linked, especially when one also considers that in the UK one third of all the self-employed hire at least one paid employee (Moralee 1998). Considering the size and importance of self-employment, literature that incorporates it into a model of the labor market is relatively sparse.

This paper develops a search-theoretic general equilibrium model of the labor market that incorporates self-employed individuals. The self-employed entrepreneurs are treated in a Schumpeterian way, as a source of innovation (Schumpeter 1934)<sup>1</sup>. It is distinct from other labor market models of self-employment in that after innovation, the self-employed

agent can work on their own, as an own-account worker, or begin hiring workers from a labor market with search frictions. The importance of incorporating the self-employed into an equilibrium model of the labor market is seen by looking at simulations of labor market policy. Conventional wisdom would suggest that a rise in unemployment benefits would have an adverse effect on aggregate unemployment—workers require better job offers or better *ideas* to exit unemployment for paid or self-employment respectively. Introducing the self-employed as a source of job creation introduces a counterweight to this straightforward mechanism. If only sufficiently good ideas create employment opportunities, making the self-employed more fastidious about which ideas to act on could create employment opportunities for other agents. This paper finds that for small increases in the generosity of unemployment benefits, aggregate employment increases. This is used as an illustrative example to show the importance of considering the self-employed when implementing active labor market policy.

In the model, there are two types of agents, large private sector firms who vary in their productivity and ex ante homogeneous workers who are exposed to innovative ideas which arrive at an exogenous Poisson rate that is dependent on their labor market status. When an agent gets an idea, its quality is drawn from a known distribution and the agent decides whether to act on it. If they choose to start a business, then depending on the quality of the draw, they may commence attempting to hire workers as paid employees. Thus, paid employees are either hired by large firms or self-employed recruiters. All workers are finitely lived and when in paid employment are exposed to an exogenous probability that they lose their job; all wage offers and job arrival rates to paid employment are determined endogenously. This rich setting allows for a multitude of avenues in which the two sectors are interlinked.

The model is structurally estimated using the British Household Panel Survey (BHPS). The identification strategy allows primitive productivity distributions and hiring behavior of both types of firm to be uncovered. The distribution of productivity amongst large firms is uncovered, as in Bontemps et al. (2000) by an inversion of the wage offer distribution; the productivity amongst the self-employed is obtained directly from their earnings; and hiring behavior is estimated in order to match the distribution of firm size amongst large privately owned companies and small firms owned by self-employed recruiters. Estimates suggest that the productivities of the two types of firms are very similar; however, self-employed owned firms are responsible for a much smaller share of employment because they face much greater frictions in hiring paid employees. It is necessary to uncover output and hiring through the guise of an equilibrium model as there exist large limitations of data on the self-employed. Data are specifically limited in information regarding the recruiting self-employed. The BHPS is useful as it distinguishes the self-employed between those who hire paid employees and those who do not. However, it does not identify the paid employees who are hired by recruiting self-employed nor does it provide information on output. Therefore, in order to infer rates of hiring and production, the paper leans heavily on a model that provides a great deal of structure to the data.

There are models that embed self-employment into a labor market equilibrium. However, there are none, to the author's knowledge, that allow for the two distinct types of self-employment discussed. Kumar and Schuetze (2007) develop a model of the labor market that incorporates self-employment and assess the effect of changes to

unemployment insurance and the minimum wage on labor market equilibrium. Self-employed hire paid employees as in this paper, but unlike this model, there is no wage dispersion within sector nor are there any movements across sectors. Thus, they restrict a direct interaction between the two sectors. Narita (2014) and Margolis et al. (2014) allow for wage dispersion within sector and structurally estimate the parameters of their models using data from Brazil and Malaysia, respectively. However, unlike Kumar and Schuetze (2007) and this paper, the self-employed are restricted to being own-account workers (they are restricted from hiring). Also, although they allow for more mobility than Kumar and Schuetze (2007), they still omit any direct transitions between paid and self-employment.

Although not all concern themselves with self-employment, perhaps the most similar papers methodologically are Meghir et al. (2015), Bradley et al. (2015), and Millán (2012). All introduce another sector into an equilibrium labor market setting allowing for a great deal of mobility between and across sectors, be it an informal, public, or self-employed sector. Decisions made by workers depend not only on the wage they are offered but also future prospects associated with the sector. Millán (2012) suggests that self-employment is used as a route out of unemployment. This paper does not consider this hypothesis, but workers may be encouraged to become self-employed as they face the opportunity of starting a business which can grow. Only Kumar and Schuetze (2007) and this paper entertain the idea of the self-employed as employers, and only this paper distinguishes between the self-employed as own-account workers and recruiters.

The consensus in the empirical literature is that there exists a substantial paid employment premium over being self-employed. Hamilton (2000) finds that taking into account within sector earnings growth and without distinguishing between own-account self-employed and recruiters, there is a 35 % differential between median earnings of a self-employed individual and a paid employee with 10 years of experience in the USA. To rationalize workers' career choices, "results suggest that the nonpecuniary benefits of self-employment are substantial". Critiquing the shortcomings of the existing literature Hamilton (2000) goes on to say "results presented here are of a reduced form [...] structural estimates of the compensating wage differential, for example, would require [...] the probability of observing particular employment and earnings sequences". To my knowledge, this is the first paper that attempts to structurally estimate the nonpecuniary benefits of self-employment while taking into account different employment options, including starting and growing one's own business. A priori, it is not clear which sector has the preferable employment and earnings profile. On the one hand, those in paid employment are better exposed to other jobs in paid employment and can climb the job ladder thusly. However, the self-employed are far more likely to become recruiters. If their firm successfully grows, so will their earnings. Consistent with the empirical literature, this paper too finds a large positive amenity associated with self-employment.

The rest of the paper is structured as follows. Section 2 presents and solves the equilibrium model. Section 3 presents the data, which corrects the earnings of the self-employed, attempts to validate assumptions made in the model, and explains the moments used for identification. Section 4 outlines the estimation protocol, the results, and the fit of the model. Section 5 runs counterfactual policy simulations to assess the effects of increasing unemployment benefit on aggregate employment, and finally, Section 6 concludes.

## 2 The model

### 2.1 The environment

Time is continuous, and at any point, there exists a unit mass of ex ante homogeneous workers and a mass  $N$  of private sector firms who are heterogeneous in their level of productivity. Workers and firms are risk-neutral and discount the future at a rate  $r > 0$ . Workers can be in one of three broad states: unemployment, paid employment, or self-employment. They leave the labor market at an exogenous Poisson rate  $\mu$  which is independent of their labor market state; they are replaced with new agents who are born into unemployment. Firms are infinitely lived<sup>2</sup>.

Paid employment comes from two sources, large private sector firms and self-employed individuals. Job offers arrive to unemployed agents from firms at a rate  $\lambda_0$  and paid employees in large firms at a rate  $\lambda_1 = \kappa\lambda_0$ .  $\kappa$  is the relative search intensity of an employed worker compared to an unemployed one. Wages are drawn from a known distribution  $F(\cdot)$ . Job offers can also arrive from self-employed recruiters at a rate  $\lambda_0^s$  to the unemployed. For tractability, it is assumed that the recruiters direct all their search intensity to the unemployed, attempts will be made to justify said assumption in Section 3. These wages are drawn from a known distribution  $F^s(\cdot)$ . All job offer arrival rates and wage offer distributions are endogenous objects. A match is destroyed at a Poisson rate  $\delta$ . Both firm and worker continue to exist, but the worker is reallocated to unemployment. If the worker exits the labor market, at a rate  $\mu$ , the firm will continue to exist, but with one less worker. Agents are also exposed to the possibility of having an innovative idea. The rate agents get ideas follows a Poisson process which is dependent on one's labor market state. Ideas arrive to unemployed agents at a rate  $\eta_0$  and to paid employees in large firms at a rate  $\eta_1$ . The quality (productivity) of the idea is drawn from an exogenous distribution  $\Gamma(\cdot)$ . Depending on the draw, this will allow workers to cross the market from unemployment or paid employment and become self-employed. The reservation productivity of paid employees will depend on their current wage.

Self-employment spells begin as an agent working on his own producing output according to the draw from  $\Gamma(\cdot)$ . There is a search friction in the matching process: hiring workers takes time. One can think of this as time required for vetting, preliminary training etc. So, as will transpire, some of the self-employed will wish to hire but are time restricted and therefore hire at a suboptimal rate.

It is assumed that the self-employed hire at a Poisson rate  $h\ell$ , where  $h$  is exogenous and constant for all firms (independent of size and productivity) and  $\ell$  is the size of the firm, can take any positive integer value. Thus, by construction, the growth of a firm adheres to *Gibrat's law*<sup>3</sup>. The rate of total hiring is proportional to the firm size as the hiring process requires a certain amount of work. Larger firms can share this workload over a greater number of workers and hence the rate of hiring increases proportionally with the size. In a way akin to Coles and Mortensen (2011), the size of the firm  $\ell$  will follow an endogenously determined Markov process.

Finally, the model is derived in a steady-state. The stocks of agents in each of the three states are constant over time as are the distributions of firm size, productivity amongst the self-employed, and the distribution of wages amongst the paid employees. To solve the model, one needs to (i) derive the value functions of the three states, (ii) derive reservation levels for which workers change states, (iii) calculate the size of each state and derive the ergodic distributions of wages and productivity

within states, and (iv) derive the profit maximizing wage policy of large private sector firms.

## 2.2 Value functions

Workers can be in one of five states: paid employment in a large firm; unemployment; paid employment, employed by a self-employed recruiter; own-account self-employment (working alone); and recruiting self-employment (hiring paid employees). In all states, workers are maximizing their lifetime income discounted at a rate  $r$ . The following subsections derive the lifetime values for each state.

### 2.2.1 Paid employees

A paid employee working for a large firm has two sources of revenue flow from employment, a basic wage and a nonpecuniary amenity, which ex ante can be positive or negative; it is measured relative to being a self-employed worker. As well as the revenue flow of income and amenity, the value function of a paid employee has the option value of transiting into other states, namely the option value of unemployment, higher value paid employment (employed by firm or a self-employed agent), and self-employment. At any point, there is a possibility that the agent exits the labor force. All future revenue flows are discounted at a rate  $r$ .

The value function for a paid employee earning a wage  $w$  in a large firm is given by<sup>4</sup>

$$(r + \mu)W(w) = w + a + \delta [U - W(w)] + \lambda_1 \mathbb{E}_F \max [W(x) - W(w), 0] + \eta_1 \mathbb{E}_\Gamma \max [S(z) - W(w), 0]. \quad (1)$$

An individual exits to unemployment at a rate  $\delta$  when the match dissolves. They receive other job offers from firms at a rate  $\lambda_1$ , and they have innovative ideas at a rate  $\eta_1$ . After an idea, a private sector employee either stays employed or becomes self-employed which gives value  $S(z)$ , where  $z$  is the productivity draw. If the idea is sufficiently good, they leave to self-employment and, initially, work individually as an own-account worker. The value for unemployment is given by Eq. (2), where  $b$  is the flow value of unemployment, relative to self-employment.

$$(r + \mu)U = b + \lambda_0 \mathbb{E}_F \max [W(x) - U, 0] + \eta_0 \mathbb{E}_\Gamma \max [S(z) - U, 0] + \lambda_0^s \mathbb{E}_{F^s} \max [W^s(x) - U, 0] \quad (2)$$

If a paid employee is hired by a self-employed agent, they face very different opportunities than if they are employed by a large firm. The validity of these assumptions will be addressed when looking at the data in Section 3. Firstly, as the firm is relatively small, it is assumed that the actions of the workers are observable to the self-employed recruiter. Thus, the worker spends no time searching for other jobs or thinking about innovative ideas. Therefore, in addition to exiting the labor market, the only other potential transition facing the employee is to unemployment. This occurs more frequently than in paid employment in a large private sector firm because not only does the worker contend with the possibility the match is destroyed, at rate  $\delta$ , but the recruiter employing him can also exit the labor market, at a rate  $\mu$ . Like a paid employee in a large firm, the amenity

associated with paid employment in a self-employed owned firm is given by the flow benefit  $a$ . The value function for a paid employee earning  $w$  working for a self-employed recruiter is

$$(r + \mu)W^s(w) = w + a + (\mu + \delta) [U - W^s(w)] \quad (3)$$

### 2.2.2 The self-employed

All self-employed agents start their self-employment spell as an own-account worker. That is, they employ only themselves. If an agent aims to recruit workers, they will arrive at a rate  $h\ell$ , where  $\ell$  is the integer number of employees they already have. Workers will quit the firm at a rate  $(\mu + \delta)$ , exiting to either unemployment or out of the labor force. Thus, the number of workers a recruiter employs will follow a Markov process. The only search friction that exists for a recruiter is that hiring is a time-consuming process. For a firm with an integer number of employees  $\ell$ , it takes an estimated time  $1/h\ell$  to hire a worker. Unlike the Burdett and Mortensen (1998) model, a higher wage neither increases the rate of recruitment nor retention. For ease of exposition, it is assumed that the self-employed never exit self-employment for unemployment.

To begin with, attention is restricted to those self-employed whose interest it is to recruit workers. Then, at the start of a self-employment spell, an agent produces as an own-account worker, his output is equal to his productivity, and at a rate  $h$  will hire a worker. The value function for a newly self-employed agent of productivity  $y$  who intends to recruit is given in Eq. (4), where  $R(y, 1)$  is the value of a self-employed individual with productivity  $y$  and one employee.

$$(r + \mu)S_R(y) = y + h [R(y, 1) - S_R(y)] \quad (4)$$

If a self-employed agent does find a worker, he foregoes his own output in order to manage the newly created firm. However, the production per worker is now  $p(y)$ , where  $p(y)$  is the productivity per worker of a self-employed owned firm of quality  $y$ . Any worker employed produces output using the self-employed agent's technology,  $p(y)$ , and receives an endogenously determined wage,  $w^*(y, \ell)$ . Some regularity conditions are imposed on  $p(y)$ , they are, that  $p'(y) > 0$  for all  $y$  and  $p''(y) \geq 0$  for all  $y$ . Intuitively, one might expect  $p(y) > y$ , as the firm has management oversight; however, a priori, only the two regularity assumptions are assumed. The value function for a recruiter is given by  $R(y, \ell)$ , where  $y$  is the productivity and  $\ell$  is the integer number of workers employed. The right hand side of Eq. (5) contains: the profit flow, production net of wages; plus the probability of expanding employment by one employee multiplied by the option value of that occurrence; plus the equivalent for losing one worker, either the worker returns to unemployment or leaves the labor force entirely. Equation (5) is expressed for one worker and an arbitrary amount of workers  $\ell$ .

$$\begin{aligned} (r + \mu)R(y, 1) &= p(y) - w^*(y, 1) + h [R(y, 2) - R(y, 1)] \\ &\quad + (\mu + \delta) [R(y, 0) - R(y, 1)] \\ &\quad \vdots \\ (r + \mu)R(y, \ell) &= (p(y) - w^*(y, \ell))\ell + h\ell [R(y, \ell + 1) - R(y, \ell)] \\ &\quad + (\mu + \delta)\ell [R(y, \ell - 1) - R(y, \ell)] \end{aligned} \quad (5)$$

The recruiter sets  $w^*(y, \ell)$  to maximize his present discounted value. Since posting vacancies and hiring are costless (other than time), there is no incentive to pay higher wages to attract workers. In addition, the recruiter does not need to pay a higher wage to retain his workers, as there is no retention motive. He observes them not looking for other jobs or innovative ideas. He cannot, however, pay them nothing; he must pay a sufficiently high wage such that some workers want to be employed. The recruiter exclusively hires from the pool of unemployed; therefore, the wage paid makes workers indifferent between remaining unemployed and accepting an offer. That is,  $W^s(w^*(y, \ell)) \geq U$ ; a worker must be at least as well off in paid employment than they are in unemployment. The recruiter's problem is

$$\max_{w^*(y, \ell)} R(y, \ell) \quad \text{subject to} \quad W^s(w^*(y, \ell)) \geq U \quad (6)$$

Since the value of recruitment decreases with the wage rate, the optimal wage is independent of  $y$  and  $\ell$  and solves the equality  $W^s(w^*) = U$ . An explicit solution will be given for  $w^*$  in the next section.

Further, it is assumed that if a recruiter has just one employee, if that employee leaves the firm, the firm and the self-employed individual cease to exist (retires) then  $R(y, 0) = 0$ . This assumption is made purely for tractability reasons, as given this, the difference equations specified in Eq. (5) are linear in  $\ell$ . With this in mind, it is trivial to see that

$$R(y, \ell) = \ell R(y, 1). \quad (7)$$

Thus Eq. (5) simplifies to

$$R(y, 1) = \frac{p(y) - w^*}{r + 2\mu + \delta - h}. \quad (8)$$

Substituting Eq. (5) into Eq. (4), one can obtain the recursive solution for the value of becoming self-employed and intending to recruit with productivity  $y$

$$S_R(y) = \frac{y(r + 2\mu + \delta) + h(p(y) - y) - hw^*}{(r + 2\mu + \delta - h)(r + \mu + h)}. \quad (9)$$

However, it is not clear whether it is in the interest of the self-employed to hire any paid employees. They have to forfeit their own production and increase their chance of going under. Looking at Eq. (9), as the wage they pay their employees increases, the value of self-employment with the intention to recruit decreases. The value function of a self-employed individual of productivity  $y$  to remain an own-account worker will be just their output, which they get indefinitely (unless they exit the labor market).

$$(r + \mu)S_O(y) = y \quad (10)$$

For an individual not to set about recruiting, the option value of having one worker must be negative. A self-employed individual with productivity  $y$  will aim to recruit workers if  $S_R(y) \geq S_O(y)$ . Given certain regularity conditions, to follow, there exists a threshold productivity level  $\psi_1$ , above which self-employed agents intend to recruit.

So the value function for a newly self-employed agent is given by Eq. (11). For  $y \in [\psi_0, \psi_1)$ , an agent will make no attempt at hiring and their value function is simply the present discounted value of constant production of amount  $y$ . If  $y \geq \psi_1$ , they will initially be an own-account worker, but will hire someone with Poisson rate  $h$  and produce an amount  $p(y)$ , paying  $w^*$  to their employee. Thus, the expression is increasing in  $h$  and in

$p(y)$  and decreasing in  $w^*$ . It is also decreasing in  $r$ ,  $\mu$ , and  $\delta$ , the discount rate, and the Poisson rates of agents leaving the labor market and job destruction, respectively.

$$S(y) = \begin{cases} \frac{y}{r+\mu} & \text{if } y < \psi_1 \\ \frac{y(r+2\mu+\delta)+h(p(y)-y)-hw^*}{(r+2\mu+\delta-h)(r+\mu+h)} & \text{if } y \geq \psi_1 \end{cases} \quad (11)$$

### 2.3 Reservation strategies

A worker's strategy can be characterized by a set of reservation values which depends on his current labor market state.

The wage a paid employee receives in a large firm that makes him indifferent between becoming self-employed of productivity  $y$  and remaining in paid employment is defined as  $\phi(y)$ . Similarly, the productivity a paid employee will require to enter self-employment is  $\psi(w)$ . These functions solve the equalities:

$$S(\psi(w)) = W(w) \quad S(y) = W(\phi(y)) \quad (12)$$

Using the above two equations:

$$S(y) = W(\phi(y)) = S(\psi(\phi(y))) \quad (13)$$

Thus,

$$y = \psi(\phi(y)) \quad (14)$$

Hence, given monotonicity of the value functions,  $\psi$  and  $\phi$  are reciprocals of one another. Similarly, the wage that makes an unemployed agent indifferent between continuing unemployment and being employed by a large firm is  $\phi_0$  and solves the equality  $W(\phi_0) = U$ , and the productivity of an idea that makes an unemployed agent indifferent between continuing unemployment and being self-employed at that productivity is  $\psi_0$  and solves the equality  $S(\psi_0) = U$ . Clearly,  $\psi(\phi_0) = \psi_0$ .

To find solutions for these reservation strategies, it is convenient to begin by simplifying the value functions for workers: Firstly,  $W^s(w) = W^s$ , as  $F^s(\cdot)$  is a degenerate distribution at  $w^*$ . Since  $W^s = U$ , a worker never gets a positive option value from being employed by a self-employed recruiter; therefore, it drops out of the Bellman equations. Finally, in calculating the expectation, one can integrate by parts, where the overscore on the distribution represents the survival function, for example,  $\bar{F}(\cdot) = 1 - F(\cdot)$ . Thus Eqs. (1), (2), and (3) simplify to (15), (16), and (17), respectively.

$$(r + \mu)W(w) = w + a + \delta [U - W(w)] + \lambda_1 \int_w [W'(x)\bar{F}(x)] dx + \eta_1 \int_{\psi(w)} [S'(z)\bar{\Gamma}(z)] dz \quad (15)$$

$$(r + \mu)U = b + \lambda_0 \int_{\phi_0} [W'(x)\bar{F}(x)] dx + \eta_0 \int_{\psi_0} [S'(z)\bar{\Gamma}(z)] dz \quad (16)$$

$$W^s = \frac{w^* + a}{r + \mu} \quad (17)$$

Solutions for  $\psi_0$ ,  $\phi_0$ , and an ODE defining  $\phi(y)$  are provided in "Solving for reservation strategies" in the Appendix.



## 2.4 Steady state

The model is derived in a steady state. A steady state is defined as a constant share of agents in each labor market state, and the distributions of wages amongst the paid employees, the productivity amongst the self-employed, and the distribution of employment size amongst recruiters are all stationary.

The steady state is defined by Eqs. (18) through (21). The sum of all agents in the economy is equal to unity. The flow into unemployment equals the outflow. The flow out of self-employment (paid employment) below a productivity  $y$  (wage  $w$ ) is equal to the flow into self-employment (paid employment) below a productivity  $y$  (wage  $w$ ). As well as this, the distribution of labor force size amongst self-employed recruiters that is dictated by a Markov process has reached its ergodic distribution. This last object is denoted as  $\Sigma(\ell)$ ; it is the measure of recruiters and those who intend to recruit hiring an integer  $\ell$  workers.

An agent can be in one of three broad states and the sum of agents is equal to unity: they can be unemployed, in paid employment (by a large firm or a self-employed recruiter), or in self-employment. One could also differentiate further, defining those self-employed, recruiting or with the intention to recruit and those who are own-account workers.

$$N_u + (N_e^f + N_e^s) + N_s = 1 \quad (18)$$

The flow out of unemployment, the left hand side of Eq. (19) is made up of four flows. In the order they are expressed, they are those leaving to paid employment in large firms, those leaving to be self-employed, those exiting the labor force, and those becoming employed by small self-employed recruiters. This last flow is equal to  $hN_e^s + h\Sigma(0)$ . The rate at which the self-employed hire is equal to  $h$  multiplied by the total number of agents engaged in hiring, that is, all the employees ( $N_e^s$ ) plus the recruiters who are yet to hire any workers ( $\Sigma(0)$ ). The flow exclusively comes out of unemployment as anyone in paid employment (by large firms) would reject any offer. The flow into unemployment is comprised of all the new entrants into the labor market plus the paid employees whose employers have exited the labor market. New entrants arrive at a rate that exceeds  $\mu$ . Firms of size one have a double coincidence of exiting the labor market, as discussed earlier. They also leave the labor market if they lose their final worker.

$$N_u (\lambda_0 + \eta_0 \bar{\Gamma}(\psi_0) + \mu) + h(N_e^s + \Sigma(0)) = \mu + (\mu + \delta)\Sigma(1) + (\mu + \delta)N_e^s + \delta N_e^f \quad (19)$$

Equation (20) equalizes the flow in and out of self-employment below a productivity level  $y$ .  $\Gamma(\cdot)$  is the primitive distribution from which agents draw the productivity of their ideas from and  $\Gamma_s(\cdot)$  is the distribution of productivity amongst the self-employed, an endogenous object. The inflow into self-employment, the right hand side of Eq. (20), comes from those from unemployment and those in paid employment (by large firms). If the latter are earning a wage  $w$ , they must get an innovative idea of greater than productivity  $\psi(w)$ . The self-employed only exit their state to being out of the labor force. This exit rate increases once they start recruiting. Hence, the equation is split either side of  $y = \psi_1$ . If they have one employee, they will cease to exist if that one employee exits

the labor force. The proportion of self-employed recruiters with exactly one employee is given by  $\frac{\Sigma(1)}{N_s \bar{\Gamma}_s(\psi_1)}$ .

For  $y < \psi_1$ :

$$N_s \Gamma_s(y) \mu = N_u \eta_0 [\Gamma(y) - \Gamma(\psi_0)] + N_e^f \eta_1 \int_{\phi_0}^{\phi(y)} [\Gamma(y) - \Gamma(\psi(x))] dG(x)$$

and for  $y \geq \psi_1$ :

$$\begin{aligned} N_s \Gamma_s(y) \mu + N_s (\Gamma_s(y) - \Gamma_s(\psi_1)) \frac{(\mu + \delta) \Sigma(1)}{N_s \bar{\Gamma}_s(\psi_1)} &= N_u \eta_0 [\Gamma(y) - \Gamma(\psi_0)] \\ &+ N_e^f \eta_1 \int_{\phi_0}^{\phi(y)} [\Gamma(y) - \Gamma(\psi(x))] dG(x) \end{aligned} \quad (20)$$

For those in paid employment in large firms below a wage  $\phi(y)$ , agents exit paid employment to unemployment at a rate  $\delta$ ; they exit the labor force entirely at a rate  $\mu$ . They find higher paid jobs in large firms at a rate  $\lambda_1 \bar{F}(\phi(y))$ . Paid employees can also exit to self-employment; they get an innovative idea at a rate  $\eta_1$  and will act on it depending on their current wage and the quality of the draw they get from  $\Gamma(\cdot)$ . The inflow is entirely from unemployment.

$$\begin{aligned} N_e^f G(\phi(y)) (\mu + \delta + \lambda_1 \bar{F}(\phi(y))) + N_e^f \eta_1 \int_{\psi_0}^y \bar{\Gamma}(z) dG(\phi(z)) \\ = N_u \lambda_0 F(\phi(y)) \end{aligned} \quad (21)$$

Equations (18), (19), (20), and (21) are solved simultaneously for the endogenous objects  $(N_s^s, N_u, N_e^f G(\phi(y)), N_s \Gamma_s(y))$ . These coupled with the distribution  $\Sigma(\ell)$  define the steady state allocation of agents. The solution for these objects is provided in ‘‘Solving for the steady state’’ in the Appendix.

## 2.5 Private sector firms

Private sector firms are large and infinitely lived. The law of large numbers is employed, and they are modeled following Bontemps et al. (2000). This paper considers the familiar equilibrium where firms post wages and commit to those wages for their lifetime. The higher the wage a firm posts, the larger the firm (fewer quits and more hires); this is at the cost of making less profit per worker. As pointed out by Coles (2001), this is not a dynamically consistent wage posting strategy without commitment on wages. To see this, imagine a firm has grown to its steady state size. Rather than posting its optimal wage  $w$ , it will be strictly better off to pay its workers’ reservation wage  $\phi_0$ . Over a period  $dt \rightarrow 0$ , no workers will quit and the firm will make strictly greater profits. Coles (2001) identifies a dynamically consistent equilibrium without relying on commitment. Interestingly, the wages posted by self-employed recruiters are a dynamically consistent strategy, without having to rely on commitment. It is the belief of the author that despite theoretical issues regarding dynamic consistency, the tractability of Bontemps et al. (2000) makes it very well suited to modeling the behavior of large firms.

There exists a continuum of infinitely lived private sector firms of mass  $N$  who are profit maximizers and heterogeneous in their level of productivity,  $y$ , where  $y \sim \Gamma^f(\cdot)$ . Each firm has the same exposure in the labor market, and all receive a mass of contacts at an exogenous Poisson rate  $h^f$ . Unlike self-employed recruiters, it is assumed that large

private sector firms are in some way detached from the labor market and cannot target a specific subgroup of job seekers. Thus, not every contact is associated with a hire. The steady state size of a firm posting a wage  $w$  is therefore a function of the proportion of contacts it makes that accept the offer,  $a(w)$ , and the number who are in employment and quit  $\Delta(w)$ .

$$a(w) = \frac{\lambda_0 N_u + \lambda_1 N_e^f G(w)}{\lambda_0 N_u + \lambda_1 N_e}$$

$$\Delta(w) = \mu + \delta + \lambda_1 \bar{F}(w)$$

The size of a firm posting a wage  $w$  is given by  $\ell_f(w)$  and solves the flow balance equation

$$\ell_f(w) \Delta(w) = h^f a(w) \quad (22)$$

By setting a wage  $w$ , a firm of productivity  $y$  will be of size  $\ell_f(w)$  and thus have total profit given by

$$\max_{w \geq \phi_0} \pi(w; y) = (y - w) \frac{h^f a(w)}{\Delta(w)} \quad (23)$$

The firm's problem is solved by the first order condition

$$y - w = \frac{a(w) \Delta(w)}{a'(w) \Delta(w) - a(w) \Delta'(w)} \quad (24)$$

Thus, if wages are increasing in productivity, then one can infer that  $F(w) = \Gamma^f(y(w))$ , where the relationship  $y(w)$  is given by (24) and  $\Gamma^f(\cdot)$  is the cumulative distribution of productivity amongst large private sector firms. Thus, the distribution of wage offers in the private sector  $F(w)$  can be retrieved if we know the distribution of productivity across firms.

To close the model, labor demand is fully endogenized. The total number of contacts from firms that workers receive is given by

$$M = \lambda_0 N_u + \lambda_1 N_e^f$$

The total jobs offered by firms are the product of the mass of firms and the number of contacts made per firm,  $Nh^f$ . Recall, there is some additional friction (if  $\kappa < 1$ ) in looking for a job while employed by a large firm, relative to being in unemployment,  $\lambda_1 = \kappa \lambda_0$ . The arrival rate of job offers from private sector firms to unemployed agents is then given by

$$\lambda_0 = \frac{Nh^f}{N_u + \kappa N_e^f} \quad (25)$$

## 2.6 Equilibrium characterization

Given exogenous parameters  $r, a, b, \mu, \delta, \eta_0, \eta_1, h, p(y), \kappa, \Gamma(\cdot), \Gamma^f(\cdot), N, h^f$ , an equilibrium is characterized by the following conditions: (i) agents behave optimally in their career decisions, solutions to  $\phi_0, \psi_0, \psi_1$ , and  $\phi(y)$ ; (ii) the economy is in steady-state, the inflow from a given state equals the outflow, solutions to  $\Gamma_s(\cdot), G(\cdot), N_u, N_e^f, N_e^s, N_s$ , and  $\Sigma(\ell)$ ; (iii) large private sector firms and self-employed recruiters offer their workers' wages optimally,  $w(p)$  and  $w^*$ ; and (iv) when firms behavior is aggregated the wage offer distribution  $F(\cdot)$  and the offer arrival rates  $\lambda_0, \lambda_0^s$ , and  $\lambda_1$  are determined. Attention is restricted to a

specific class of equilibria where self-employed individuals exist, some of whom recruit paid employees. To guarantee this equilibrium, the exogenous parameter space needs to be constrained.

**Assumption 1.**  $h \in (0, r + 2\mu + \delta)$

**Assumption 2.** Either  $p'(y) \geq 0$  and  $p''(y) > 0$  or  $p'(y) > \frac{r+2\mu+\delta-h}{r+\mu}$  and  $p''(y) \geq 0$ .

**Proposition.** *Given Assumptions 1 and 2, an equilibrium will exist with self-employed recruiters.*

The intuition is as follows. These assumptions are needed for two reasons. Firstly, Assumption 1 guarantees the nonnegativity of the value function of a recruiter; see (8). Given this, Assumption 2 guarantees that for a sufficiently good idea, a self-employed agent will actively seek to hire workers.

**Proof.** For an individual to consider self-employment as a viable option, it must yield a positive value. A recruiter with one employee has value given by Eq. (8). Clearly, he must make positive profit per worker ( $p(y) > w^*$ )

$$R(y, 1) = \frac{p(y) - w^*}{r + 2\mu + \delta - h}$$

□

The above is positive if Assumption 1 holds. Recruiters will exist if, for some  $y$ , the value of intending to recruit,  $S_R(y)$ , is greater than being an own-account self-employed indefinitely,  $S_O(y)$ .  $S_O(y)$  is a linear function of  $y$ , and  $S_R(y)$  is a linear function of  $p(y)$ . If  $p'(y) \geq 0$  and  $p''(y) > 0$ , the first part of Assumption 2. Clearly, for sufficiently large  $y$ ,  $S_R(y) > S_O(y)$  and given Assumption 1, recruiters will exist in equilibrium.

If  $p(y)$  is linear and increasing in  $y$ , the second part of Assumption 2. Then both  $S_O(y)$  and  $S_R(y)$  are linear and increasing in  $y$ . Thus, for sufficiently large  $y$ , recruiter will exist if  $S'_R(y) > S'_O(y)$ . This is the case, given  $p'(y) > \frac{r+2\mu+\delta-h}{r+\mu}$ , contained in Assumption 2. *Q.E.D.*

**Identification result:** *A sufficient condition for the existence of recruiters is  $p'(y) > \frac{r+2\mu+\delta}{r+\mu}$*

For given  $p(y)$ ,  $r$ ,  $\delta$ , and  $\mu$ , one can guarantee the existence of recruiters (given Assumption 1 holds) if

$$p'(y) > \max_{0 < h < r+2\mu+\delta} \left( \frac{r + 2\mu + \delta - h}{r + \mu} \right)$$

$$p'(y) > \frac{r + 2\mu + \delta}{r + \mu}$$

**Class of equilibrium:** the equilibrium is restricted to one where self-employed recruiters exist. Clearly, some self-employed will be own-account workers. These can fall into two categories either they are own-account workers, but because of the frictions present in the market, as yet, they have been unable to hire anyone. Or, they are own-account workers with no intention to recruit. For the latter type of agent to exist,

it is required that  $S_0(y) > S_R(y)$  for some  $y$  and because  $S'_R(y) > S'_O(y)$  for all  $y$ , this is equivalent to  $S_O(\psi_0) > S_R(\psi_0)$ .

Since one cannot restrict attention to either class as neither can be invalidated by data, any simulation of the model needs to repeatedly check in what equilibrium it is in. One where all self-employed intend to recruit or one where some self-employed intend to stay own-account indefinitely.

### 3 Data

The model described is estimated using the British Household Panel Survey (BHPS). It is identified using transition rates across labor market states, the earnings of paid employees and self-employed workers and the distribution of firm size.

Exogenous parameters are estimated by simulated method of moments. The principle of the estimation technique is to find values of the structural parameters that minimize a function of the difference between a chosen set of moments from the data and data simulated with these values of the structural parameters.

Of the exogenous parameters, the only one that is not estimated is the discount rate  $r$ .  $r$  is calibrated as equal to 0.0043, which is the monthly (continuous time) equivalent of a 5 % annual rate. In order to estimate the distribution from which the self-employed draw their level of productivity, a parametric assumption is made. It is assumed that  $\Gamma(\cdot)$  follows a log-normal distribution with the mean and standard deviation of the associated normal given by  $m_y$  and  $s_y$ , and the production function of self-employed recruiters is specified as  $p(y) = \beta y$ . Thus, the rest of this section concerns itself with the estimation of the vector of exogenous parameters  $\theta$ :

$$\theta = \left( h, \eta_0, \eta_1, \mu, \beta, m_y, s_y, a, b, \kappa, \Gamma^f(\cdot), h^f, N \right) \quad (26)$$

The role of the rest of the data section are twofold. Firstly, it aims to inform the reader about the data that is used to estimate the model. Secondly, it attempts to find empirical support for predictions of the model as well as assumptions made for tractability.

#### 3.1 The sample

The data used in the analysis are taken from the BHPS, a longitudinal dataset of British households. Data were first collected in 1991, but attention is restricted to five waves covering the period from 2004 to 2008. The sample comprises of prime age (21–60) white male low-skilled workers. Low-skilled is defined as not having obtained A-level qualifications. These are the highest qualifications available for students aged 18 in the UK, before they enter higher education. A worker is an individual who is never inactive in the period looked at: if out of work, they declare themselves to be actively seeking work. The hourly earnings distribution is adjusted by treating the bottom and top 2.5 % of the distribution as missing, hopefully ridding the sample of erroneously reported earnings.

Data are homogenized to include only low-skilled workers as agents are ex ante homogeneous so it is important that they are of similar skill in the data. Implicitly, it has been assumed that agents are exposed to employment opportunities at the same rate within a labor market state. This seems less of an imposition on the data for low-skilled workers, as many high-skilled professions exist where agents become self-employed more frequently. In addition, the type of self-employment more prevalent amongst high-skilled workers, being made a partner of a firm for example, is less in keeping with the arrival

of an innovative idea, as modeled in this setup. Information on cross-sector differences in employment, hours worked, and earnings are reported in Table 1. Following the methodology of Hurst et al. (2014), using British consumption data, the earnings of the self-employed are adjusted to take into account any misreporting of earnings. This exercise is explained in detail in “Misreporting of earnings by self-employed” in the Appendix. The adjustment is relative to paid employment, one interpretation is therefore how much the self-employed under/over state their income relative to paid employees.

There are a number of points worth noting from Table 1. There is a non-negligible share of recruiters, consistent with the findings of Moralee (1998), with approximately one third of self-employed agents hiring paid employees. The self-employed work longer hours than their paid employee counterparts and the recruiting self-employed work significantly more still. There is a clear ordering in the second moment of the earnings distribution across the three labor market states irrespective of whether one looks at the raw or the adjusted data. The ordering of the first moment is ambiguous however; the preferred specification that will be used in estimation is the adjusted data. Looking at the large implied differences, the importance of correcting for the misreporting of earnings is evident. The ranking of the second moment of the earnings distribution is not targeted in the estimation, but will still be replicated by the theoretical model.

Table 2 shows monthly worker turnover observed in the data. For the most part, the cells of Table 2 are self-explanatory. The diagonal elements are all unobservable with the exception of workers changing jobs within paid employment. Although it is possible that a business goes bust and one instantly starts up a new one, this phenomenon is not possible to observe using the BHPS so transitions within specific self-employment states are not reported. A paid employee in a small firm is anyone who declared their establishment to have fewer than 25 people working in it. Small and large firms are differentiated between because very few recruiters have 25 or more employees, see Table 4. Therefore, the large firms provide a better indication of mobility for the paid employees hired by large private sector firms.

The rates in bold are those that the model is capable of replicating, some of which the estimator will target, and those in plain text are ones that the model is unable to generate. The model is unable to generate a number of classes of transitions. One is any transition into self-employment where the worker instantly becomes a recruiter. This is inconsistent with a labor market with search frictions, as one cannot hire individuals instantaneously.

**Table 1** Composition

	Paid employment	Own-account	Recruiter
Employment share	82.8 %	12.6 %	4.6 %
Mean hours worked per week	40.39	43.34	51.68
Raw data			
Mean earnings	10.69	9.34	11.14
Standard dev. of earnings	3.88	5.81	10.44
Adjusted data			
Median earnings	9.92	13.20	14.93
Mean earnings	10.69	15.71	18.72
Standard dev. of earnings	3.88	9.76	17.55

Employment shares are measured as the proportion employed in that state amongst all those employed. Earnings are measured in £/hour

**Table 2** Transition matrix: 2004–2008

	Unemployment	Small paid employment	Large paid employment	Own-account	Recruiter
Unemployment	–	<b>0.0009</b>	<b>0.0899</b>	<b>0.0066</b>	0.0019
Paid employment (s)	<b>0.0160</b>	<b>0.0160</b>	<b>0.0841</b>	<b>0.0053</b>	0.0027
Paid employment (l)	<b>0.0034</b>	0.0002	<b>0.0079</b>	<b>0.0009</b>	0.0000
Own-account	0.0011	0.0000	0.0056	–	<b>0.0027</b>
Recruiter	0.0007	0.0000	0.0043	0.0108	–

Transition rates are (continuous time) monthly. Rows do not add up to one. The entries along the diagonal are the fractions of individuals *changing jobs* within their current labor market status. Large paid employment is defined as being employed in an establishment that has 25 or more employees. Rates in bold are ones in which the model is able to replicate

The second is a simplifying assumption made for tractability: once in self-employment an individual is allowed little labor market mobility. Of the transitional moments omitted, the rate at which a recruiter transits to being an own-account worker is the most glaring. In a given month there is a 1 % chance of a recruiter losing its workforce and becoming an own-account worker. This model is not able to generate this, in order to keep the result that a recruiter's value is proportional to the number of employees it has (Eq. (7)), which helps the tractability of the model. Also, while 1 % seems large, in fact, from Table 1 just 4.6 % of employed individuals are recruiters so this translates into very few transitions missed. Finally, as in models of wage posting, firm heterogeneity and on the job search like Bontemps et al. (2000), there is a one-to-one relation between firm size and offered wage. In an identical way, this model can therefore not rationalize why a paid employee would take a wage cut and move to a smaller firm.

The parameter  $h$  is identified using the transition rate from own-account self-employed to becoming a recruiter. Since the employee size distribution of self-employed recruiters is governed by  $\mu$ ,  $\delta$  and  $h$ , this could be an alternative source of identification, conditional on  $\mu$  and  $\delta$ . The firm size distribution amongst private sector employers is used to identify  $h^f$ , the volume of vacancies a single firm posts. The self-employed are asked how many people they employ; the paid employees are asked how many other people are in their place of work. These numbers are grouped into size bins and from these it is difficult to infer the size distribution of private sector firms. Therefore, purely for comparison, data on firm size from the Office of National Statistics (ONS) in the midpoint of the sample period is used.

The downside of comparing the size distributions in Tables 3 and 4 is that the size bins do not correspond to one another. The cumulative firm size distribution at employee sizes 9, 49, 99, 499, and 999 can be compared. Firms owned by the self-employed are typically smaller than firms at large, 86.1 % of self-employed run firms have less than 10 employees, compared with 82.02 % of all firms. Similarly, 96.69 % of all self-employed owned firms have fewer than 50 employees compared with 96.4 % of all firms. However, at some point between 49 and 99 employees, the two cumulative distributions intersect, with fewer than 100, 500, and 1000 employees in all firms (self-employed firms) account for 98.41 % (96.69 %), 99.84 % (98.01 %) and 99.96 % (98.68 %), respectively. The greater proportion of self-employed in the upper tail could be associated with a small sample bias that will over-represent the tails of the distribution. For example, if one were to omit the three highest reported firm sizes amongst the self-employed then there would be first order stochastic domination of the distribution of the size of all firms over those self-employed. It is also further evidence of insufficient data that there is an observed hole in

**Table 3** Private sector firm size distribution 2006

Firm size	No. of firms	Percent
Total	2,084,495	100
0–4	1,391,960	66.78
5–9	317,745	15.24
10–19	178,820	8.58
20–49	120,870	5.80
50–99	41,905	2.01
100–249	23,100	1.11
250–499	6740	0.32
500–999	2440	0.12
1000 or more	915	0.04

**Source:** ONS

the size distribution, between 50 and 99 workers. To my knowledge, there is no paper that looks at this relationship in close detail. To do so adequately, one would require a more comprehensive dataset.

Finally, it is worth examining the restrictions imposed on those paid employees hired by self-employed recruiters. It is not sufficient to just look at the mobility patterns of the paid employees in small firms, as perhaps, they are also overwhelmingly hired by private firms. Instead to examine the mobility of these workers we use variation in self-employment rates across industry classification<sup>5</sup>. Using simple weighted least squared regressions, the aim is to examine whether paid employees hired by the self-employed are less or more likely to find alternative employment and consequently earn less or more. The dependent variables in the two regressions are the rate of job mobility in one digit industry classification to any employment state but unemployment and the average log wage of paid employees in one digit industry classification. In each case the explanatory variable is the proportion of employed individuals in a one digit industry who self identify as recruiting self-employed. The logic for this is if you are a paid employee in a recruiting self-employed intensive industry you are more likely to be hired by a self-employed agent than in an industry with relatively few recruiting self-employed. The coefficients are given in Table 5 below, with associated standard errors in the parenthesis.

All parameters are statistically significant to any conventional significance level. When the explanatory variable equals zero, that implies there are no self-employed recruiters

**Table 4** Self-employed recruiter's size distribution 2004–08

Firm size	No. of firms	Percent
Total	151	100
1–2	65	43.05
3–9	65	43.05
10–24	15	9.93
25–49	1	0.66
50–99	0	0
100–199	1	0.66
200–499	1	0.66
500–999	1	0.66
1000 or more	2	1.32

**Source:** BHPS



**Table 5** Weighted regressions by industry classification

	Nonunemployment exit rate	Mean (log) wage
Coefficient	−0.033 (0.0008)	−0.84 (0.016)
Constant	0.014 (0.00005)	2.4 (0.00092)

The explanatory variable in each case is the proportion of self-employed recruiters. Dependent variables are the heading of each column. Regression weighted by the paid employment level in each one digit industry, standard errors given in parenthesis

in that one-digit industry classification. If the explanatory variable equals one, it means that all employed individuals in that industry are self-employed recruiters. If one interprets the former as an instance where the probability a paid employee is hired by a self-employed recruiter is zero and in the latter the probability equals one, the coefficients have a straightforward interpretation. Simply, they are the difference in mobility and earnings associated with paid employment, given one is hired by a private sector firm or a self-employed recruiter. With this in mind, the results are supportive of the restrictions imposed on model. Using the parameter estimates from the mobility regression, a linear projection would imply that a paid employee would have zero chance, assuming non-negative probability, of exiting his current job for any other employment state but unemployment. The earnings regression suggests that a paid employee hired by a self-employed recruiter will on average earn 84 % less than one hired by a large private sector firm. This paper does not aim to suggest that these restrictions are true, rather they are not too important and seem to be borne out by the data.

## 4 Estimation

### 4.1 The estimation protocol

The model is estimated using a simulated generalized method of moments estimator (SGMM). The estimation is performed over a number of steps. The reason a multi-stepped estimation is implemented is because it makes clear the source of identification for the moments. It also aids the estimation in not allowing the constraints put on the parameter space in Section 2.6 to be violated.

Some endogenous parameters can be computed without solving the model. The endogenous parameters  $F(\cdot)$ ,  $\psi_0$  and  $\phi_0$  are all fixed according to their empirical counterparts.  $F(\cdot)$  is directly observable as the wage distribution for those who have transited into paid employment straight from unemployment. Note, strictly speaking this also includes those hired by the self-employed. To keep these to a minimum, only those transiting to a sector with less than 10 % of employment made up by recruiters are considered<sup>6</sup>.  $\psi_0$  and  $\phi_0$  are the minimum observed earnings amongst the own-account self-employed and  $\phi_0$  the infimum of the support of  $F(\cdot)$ . The nonpecuniary amenities  $a$  and  $b$  are treated as free parameters so to equalize Eqs. (29) and (30) in the Appendix. After all other parameters are estimated,  $\Gamma^f(\cdot)$  is computed so  $F(\cdot)$  is rationalized according to Eq. (24).

It proves simpler to also treat  $\lambda_0$  and  $\lambda_1$  as temporary exogenous parameters in the estimation and uncover the underlying exogenous parameters ex post of estimation. The total contacts that large private sector firms make is driven by a combination of the number of firms  $N$  and the contact per firm  $h^f$ . While matching the number of contacts the relative size of these two objects will be set to best match the aggregate firm size distribution—including the self-employed recruiters.

Estimation works as follows. In the first step, the transition rates are exactly identified and the following vector of parameters are estimated  $(\delta, h, \lambda_0, \lambda_1, \eta_0, \eta_1, \mu)$ . Conditional on these, and the parameters fixed ex ante the self-employed productivity parameters are estimated to match the earnings of own-account and recruiting self-employed, they are  $(\beta, m_y, s_y)$ . These two steps are continuously iterated on until all parameter estimates are stable. Finally, as discussed in the final step after the parameters have converged, endogenous parameters are rationalized by their exogenous primitives and the aggregate firm size distribution is fitted with its empirical counterpart. Each step is described in more detail below.

#### 4.1.1 Stage 1: transition parameters

In this stage, the remaining transition rates, reported in “Transition rate moment conditions” in the Appendix, are matched. They are the monthly rate at which individuals transit from paid employment in a firm of any size to unemployment, unemployment to paid employment, unemployment to becoming an own-account worker, paid employment to another paid employer, paid employment to becoming an own-account worker, and from an own-account self-employed worker to a recruiter.

These moment conditions are reported in “Transition rate moment conditions” in the Appendix and are exactly identified by the parameters  $(\delta, h, \lambda_0, \lambda_1, \eta_0, \eta_1, \mu)$ ;  $\lambda_0$  and  $\lambda_1$  are endogenous to the model and are rationalized in Section 4.1.3.

#### 4.1.2 Stage 2: self-employed earnings

The distribution of earnings among own-account workers is given by the solution to the set of steady state equations in Section 2.4. The earnings distribution and productivity distribution are equivalent as own-account workers earn their output. Earnings for recruiters differ by their level of productivity and their size. The distribution of profits is calculated by summing over the measure at each size distribution,  $\Sigma(\ell)$ . Deciles from the data are matched with deciles from the simulated model using the mean and variance of log productivity,  $m_y$  and  $s_y$  as well as the parameter  $\beta$  which describes the increased profitability associated with hiring. The deciles of the two earnings distributions are fitted using an equally weighted matrix, following the criterion in Altonji and Segal (1996). The authors show that when moment conditions are based on relatively few observations, an equally weighted matrix often performs better than an optimally weighted one.

After the exactly-identified first stage has fitted the transition rates, conditional on these estimates, the over-identified second stage fits the earnings of own-account workers and recruiters according to  $m_y$ ,  $s_y$ , and  $\beta$ . These parameters are updated and stage one is repeated; this is done until all the estimated parameters have converged. It is found that this multi-stepped procedure performs better in fitting the data than a single-step protocol.

#### 4.1.3 Stage 3: ex post calculations

$\kappa$ , the degree to which paid employees are exposed to private sector firms relative to the unemployed, is given by  $\kappa = \frac{\lambda_1}{\lambda_0}$ . The endogenous parameters  $\lambda_0$  and  $\lambda_1$  are estimated in Section 4.1.1.

The distribution of productivity amongst large private sector firms is identified non-parametrically. Equation (24) is computed so the productivity of a firm paying  $w$  is known and given by the relation  $y(w)$ . Then, since this is an increasing function (verified by

repeated simulation)  $F(w) = \Gamma^f(y(w))$ , where  $F(w)$  is nonparametrically estimated before the first step.

Paid employees are asked the size of their employer, put into the same size bins as Table 4. The firm size distribution of self-employed owned firms depends on  $\mu$  and  $h$  and has already been determined. In order to fit the number of workers in firms of certain size or less, the model is only able to adjust the firm size distribution of large private sector firms, who are responsible for the majority of paid employment. This distribution  $\ell_f(w)$  is given by Eq. (22). The only parameter that is left undetermined is  $h^f$ , the number of contacts each firm makes. The distribution of firm size employing paid employees' is matched by minimizing the Kolmogorov-Smirnov distance between the theoretical and empirically observed distribution. Thus, all that is left to estimate is the measure of firms in the economy  $N$ . This is set to equalize the accounting identity, given by Eq. (25).

## 4.2 Results

Table 6 presents the point estimates of the exogenous parameters in the model; bootstrapped standard errors are given in the parentheses. Standard errors are based on 500 resamples of both data sources, taking into account imprecision in the estimation of how much the self-employed underreport earnings.

The transitional parameters given in the first section of Table 6 are all monthly Poisson rates. At first glance,  $h$  seems fairly small; 35 % of new own-account workers who aim to recruit will hire someone in their first 5 years. However, because hires are made at a rate  $h\ell$ , the frequency of hires increases as the number of employees grows, 88 % (98.5 %) of self-employed recruiters with five (ten) workers will hire another worker in the next 5 years.

The rate at which individuals receive innovative ideas is also infrequent. If one spends the majority of their lives in paid employment, it is highly likely that they will never have a single idea. The rate at which agents are exposed to ideas is six times higher when they are unemployed compared to when they are in paid employment. When an idea does arrive, it is drawn from a log-normal distribution with the mean and standard deviation of the natural log of productivity as given by  $m_y$  and  $s_y$  in Table 6.

Amenities presented in the second panel of Table 6 are measured in pounds per hour. The value of leisure  $b$  is negative and large, meaning for low-skilled male workers, there is a large stigma associated with unemployment<sup>7</sup>. The nonpecuniary amenity associated with self-employment is £6.53 per hour. This is commonly referred to in the literature as

**Table 6** Parameter estimates

$\delta$	$\eta_0$	$\eta_1$	$\mu$
0.0034 (0.0003)	0.0066 (0.0002)	0.0011 (0.00002)	0.0107 (0.0002)
$a$	$b$		
-6.53 (0.77)	-11.50 (1.42)		
$h$	$h^f$	$N$	$\kappa$
0.0070 (0.0001)	2.7305 (0.1069)	0.0138 (0.0005)	0.4075 (0.0119)
$m_y$	$s_y$	$\beta$	
2.4796 (0.0927)	0.2689 (0.0242)	1.6740 (0.0315)	

Given in the parentheses are the bootstrapped standard errors

the benefit associated with “being your own boss”. Comparing this with Table 1 reveals this as 49 % (44 %) of the adjusted median wage of own-account workers (recruiters).

The parameter  $\kappa$  has a slightly different interpretation as it ordinarily would have. It is the ratio of job offers that paid employees in large firms receive from other large firms compared with job offers received by unemployed agents from large firms. Since the unemployed are also exposed to job offers from self-employed recruiters which the paid employees are not, the estimate of  $\kappa$  is inflated in comparison with other canonical models. Estimates of  $N$  and  $h^f$  suggest that the mass of private sector firms is equivalent to approximately 1.4 % of the total active members of the labor market and that in total, the firms are in contact with 3.8 % of all agents active in the labor market in a given month.

$\beta$  is the factor by which production increases when an own-account worker begins to recruit. Recall, when an agent recruits, he steps down from production and acts as a managerial overseer. Since  $\beta$  is given by 1.674, a recruiting self-employed will make less after hiring their first worker than they were previously. However, losses are recouped as soon as they hire their second employee.

### 4.3 The fit

To review, moments that were specifically targeted were a selection of transition rates, information on firm sizes, and deciles of the earnings distributions of self-employed workers.

Since the transition rates were exactly identified, the seven rates targeted matched perfectly. There are other possible transitions, but as discussed, these are the transition rates the empirical procedure attempts to fit with the empirical moments. These moments are fitted extremely well, with all identical to four decimal places. Since this step is exactly identified, the quality of the fit is unsurprising. Table 7 reports the fit of the model for deciles of the earnings distribution of own-account workers. The distribution of earnings fits quite well. The minimum, although not included, is perfectly matched by construction, as it is  $\psi_0$ . Table 8 shows the moment conditions fitted for the earning of recruiters. Again, the fit is fairly good, especially when one considers that three parameters ( $m_y, s_y, \beta$ ) have effectively been used to fit the 18 moment conditions listed in Tables 7 and 8.

Table 9 shows the proportion of workers employed in a firm of given size or less, as seen empirically and as predicted by the model. Both are a mixture of those employed in large firms and small self-employed run firms. By the final stage, the size distribution of small self-employed firms has been determined. The parameter  $h^f$ , the number of contacts a large private sector firm makes, determines the size distribution of this class of firm. The

**Table 7** Fit of deciles of the earnings distribution for own-account workers

Decile	Theoretical moment	Empirical moment
10 %	7.3556	6.1111
20 %	9.0002	7.8725
30 %	10.4178	9.7059
40 %	11.8070	11.4559
50 %	13.2911	13.1993
60 %	14.9602	15.6732
70 %	17.0026	18.2312
80 %	19.7461	21.1564
90 %	24.2895	27.0048

**Table 8** Fit of deciles of the earnings distribution for recruiting self-employed

Decile	Theoretical moment	Empirical moment
10 %	1.5080	2.8799
20 %	4.5268	5.6617
30 %	7.9110	8.6274
40 %	9.6422	10.6302
50 %	13.6360	14.9321
60 %	18.0819	18.0795
70 %	22.4150	21.5686
80 %	29.7773	24.2646
90 %	45.3309	48.5292

fit appears quite poor. This is because, as discussed, there is only one parameter  $\mathcal{W}^f$ , fitting eight moment conditions. An alternative specification is to fit the estimated median of the empirical distribution, and this would be fitted perfectly. Fitting the whole distribution is preferred as the median is not directly observable, neither does it exploit all the information contained in the data.

## 5 Counterfactual policy simulation

To illustrate the importance of explicitly modeling the self-employed, specifically as a source of job creation, this section looks at the endogenous employment response as a consequence of a change in unemployment benefit. It turns out, in the simulations, for small increases in unemployment benefit, aggregate employment increases. But underlying this, aggregate employment shift is a large reallocation of workers: growth in paid employees in large private sector firms at the expense of small self-employed owned firms and a shift in the composition of the self-employed, who are now operating with better ideas, but are far less likely to take on workers. This result is in stark contrast to a typical one sector model of the labor market.

In a prototypical single sector model of this kind, the employment response is straightforward. Unemployment benefit increases the value of unemployment which in turn means that workers need higher wages to leave for paid employment. Thus, fewer firms can afford to employ workers and the unemployment exit rate falls, and with a constant employment exit rate, unemployment will unambiguously increase. This mechanism is confused somewhat with the introduction of a second sector, with differing transition rates across sector, the unemployment rate is the solution to a more

**Table 9** Proportion of paid employees in firms of given size or less

Firm size	Theoretical moment	Empirical moment
2	0.0312	0.0387
9	0.0358	0.1916
24	0.1198	0.3288
49	0.3122	0.4765
99	0.5656	0.5883
199	0.9031	0.6956
499	1.0000	0.8429
999	1.0000	0.9234

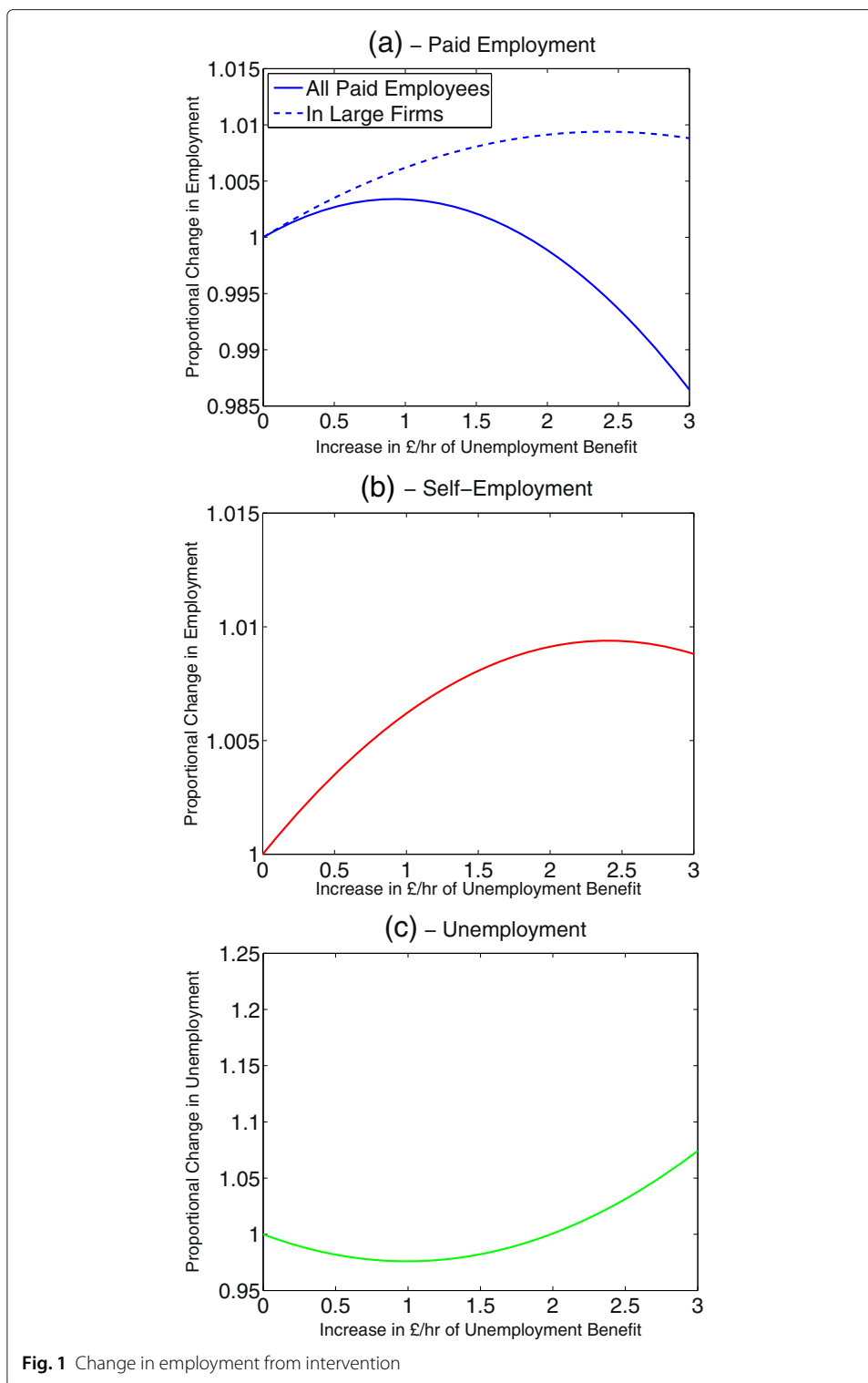
complicated set of flow equations, and the net effect is ambiguous. The inclusion of the recruiting self-employed confounds the issue yet further, as now there is a positive employment externality of one sector on another. On the one hand, if more people are unemployed, with the rate of ideas approximately six times as large in unemployment compared with employment, in aggregate, one would expect more ideas. Thus, agents only act on very good ideas, and perhaps, this leads to more recruiting self-employed and hence more job creation. Conversely, an increase in unemployment benefits will make it more expensive for the recruiting self-employed to hire workers, and therefore perhaps, fewer workers will be hired and self-employment will be less desirable.

The exact specification of the policy is to change the value of  $b$ . Recall, that the estimated value of  $b$  is negative and can be thought of as the stigma associated with unemployment net of any existing unemployment benefit. In these simulations, a series of increases from zero to £3 per hour are considered. To put this in some context, assuming a 40-h week, the maximum increase in benefit considered is equivalent to £120 per week. At the time of writing, a typical over 25-year-old claimant would expect to get £73.10 per week, so the maximum amount considered represents a fairly large expansion in the degree of generosity. The practicalities of the simulation are similar to the estimation, with two exceptions. In the estimation,  $a$  and  $b$  were treated as free parameters, now these are fixed and  $\phi_0$  and  $\psi_0$  are solved explicitly. Similarly, the wage offer distribution  $F(\cdot)$  and the offer arrival rates  $(\lambda_0, \lambda_1)$  are backed out from the productivity distribution of firms and the parameters governing the number of firms and contact rate per firm  $(N, h^f)$ .

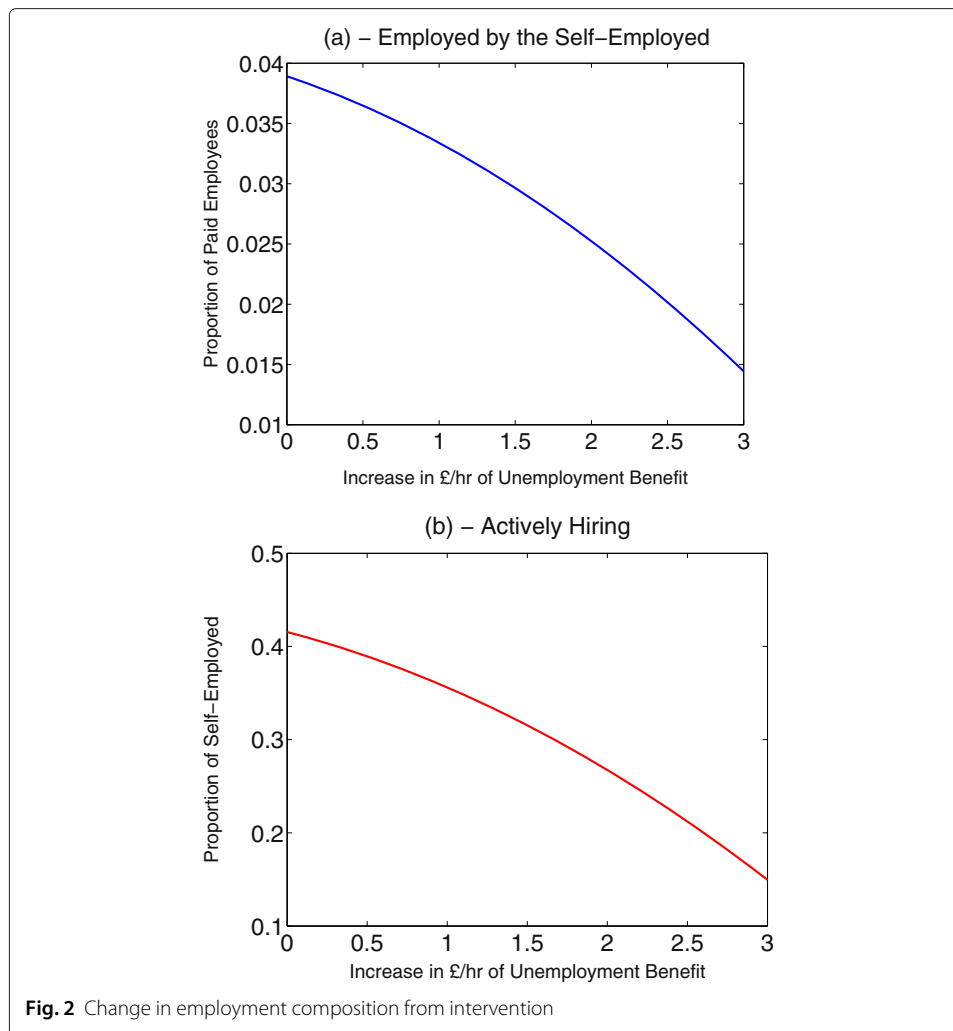
This section focuses on the endogenous employment outcomes associated with increasing unemployment benefit. These rates are calculated as described in Section 2.4. Inspection of Fig. 1 shows that for small increases in unemployment benefit, total employment will increase. This is driven by an expansion of paid employment in large private sector employment (Fig. 1a) and in self-employment (Fig. 1b). For extremely generous levels of unemployment benefits, both these rates begin to decline and this is reflected in the u-shaped unemployment rate (Fig. 1c).

The groups hit hardest by the reforms are the recruiting self-employed and their employees. Fig. 1a shows a clear disparity between how the two types of paid employment respond to the reform. As can be seen in Fig. 2a, this is driven by there being far fewer paid employees in small self-employed firms. Workers from unemployment, in particular, now command a higher wage, as there is a direct increase in the value of a worker's outside option. The self-employed recruiters, who exclusively hire from this pool are disproportionately affected. Therefore, despite having on average better quality ideas, fewer engage in actively hiring workers (Fig. 2b). These effects are large; paid employees hired by self-employed owned firms constitute just less than 4 % of the share of paid employees in the pre-reform economy. This falls to less than 1.5 % after an increase by the equivalent of £3 per hour. Similarly, the proportion of the self-employed willing to hire falls from over 40 % to around 10 % with the same level of intervention.

The fall in the level of recruiting self-employed means that there are more workers to be hired by large firms, both directly through the lack of recruiters and crucially indirectly through the lack of paid employees in small self-employed owned firms. This feedback effect is so large that for small increases in benefit, there is an expansion in aggregate



employment. This is not necessarily the only mechanism that could generate such a phenomenon, but it highlights that ignoring the self-employed who constitute such a large part of the aggregate economy may lead to misjudgments in active labor market policy.



## 6 Conclusions

This paper builds an equilibrium model of the labor market with frictions in which agents endogenously locate on either side of the market, as a paid employee, or a recruiting self-employed individual. The model is able to replicate differential features of the earnings distributions of agents in different labor market states.

Using British data, the model is estimated and the career options of the self-employed are critically assessed. Underreporting of earnings is taken into account as in Hurst et al. (2014), as are future employment and earnings profiles and any nonpecuniary amenity associated with either state. The estimated parameters are used in a counterfactual policy exercise that examines the effects of an increase in the generosity of unemployment benefits. Including the self-employed in the model yields an interesting result, that for low levels of benefit, an increase will be associated with an expansion in aggregate employment. A prediction that stands in stark contrast to typical one sector structural models of the labor market.

There has been a recent surge in the literature that incorporates self-employment into models of the labor market: Narita (2014), Margolis et al. (2014), and Millán (2012). As a result, there is a deeper understanding of the puzzle outlined by Hamilton (2000), of why



agents choose self-employment at all. By distinguishing between own-account workers and recruiters and giving the self-employed the option to develop a firm, this paper goes further still. Improvements to the precision of estimates could be made from increases to the size of the data. However, finding data with the necessary information regarding employment spells and earnings and whether an individual is a recruiter could be a challenge.

Finally, it is worth stating that there are other salient features regarding self-employment that have been overlooked in this analysis. Amongst others, poignant factors include cross employment state heterogeneity of workers, family structure, and financial constraints. Future research aimed at incorporating ex ante worker heterogeneity in order to explain the differences in composition between paid employees and the self-employed could be extremely fruitful. In a meta-analysis of recent research, Parker (2004) (Table 3.3 page 104) suggests overwhelming cross-country evidence that the self-employed are older, better educated, have more labor market experience, and are wealthier than paid employees. Another research agenda that the author believes deserves particular focus is to incorporate asset accumulation into this type of model and examine the impacts of financial constraints on self-employment. There is considerable empirical evidence suggesting that financial constraints play an important role in an individual's decision to become self-employed; for a UK context, see for example Blanchflower and Oswald (1998), Cowling and Mitchell (1997), and Black et al. (1996). Incorporating these features into the theoretical model and estimating the model, if possible, with a larger dataset could prove to be very fruitful future research projects.

## Endnotes

<sup>1</sup>For a comprehensive discussion of the history of economic thought regarding entrepreneurship, see Hébert and Link (1988).

<sup>2</sup>Firms are assumed to be infinitely lived to keep features of the Bontemps et al. (2000) model. Haltiwanger et al. (2013) finds for larger firms in the USA—those with more than 500 employees, there is less than a 1 % chance that a firm will go out of business in a given year. Implying that large firms exist on average for longer than 100 years, considerably longer than a typical agent's tenure in the labor market.

<sup>3</sup>The empirical evidence on Gibrat's law in relation to firm growth is mixed. For a summary of the literature, see Santarelli et al. (2006).

<sup>4</sup> $\mathbb{E}_F \max [W(x) - W(w), 0] = \int \max [W(x) - W(w), 0] dF(x)$  and  
 $\mathbb{E}_\Gamma \max [S(z) - W(w), 0] = \int \max [S(z) - W(w), 0] d\Gamma(z)$ .

<sup>5</sup>Industry classification are based on *The Standard Industrial Classification of economic activities 1992*. They can be found by visiting url: <http://www.ons.gov.uk/ons/guide-method/classifications/archived-standard-classifications/uk-standard-industrial-classification-1992--sic92-/index.html>.

<sup>6</sup>This corresponds to omitting those who gained employment in sectors with one-digit industry code zero, six, or eight. That is, those employed in agriculture, forestry, and fishing; finance, insurance, and real estate; and certain services.

<sup>7</sup>Interestingly, in a similar multi-sector model, Meghir et al. (2015) also estimate a large negative flow value for low-skilled unemployed workers in Sao Paulo, Brazil. Recall that  $b$  is treated as a free parameter to ensure Eq. (32) is satisfied. Thus, one reason the estimate of  $b$  is so small is that the value of self-employment for a given productivity is underestimated because it contains a large option value of paid employment, as argued by Millán (2012).

<sup>8</sup>For a comprehensive assessment of this phenomenon, see Levitt and List (2011).

### Appendix

#### Solving for reservation strategies

Before embarking on the reservation solution, first note the derivatives of the value of paid employment in a large firm and self-employment with respect to their respective arguments are given by

$$W'(w) = [r + \mu + \delta + \lambda_1 \bar{F}(w) + \eta_1 \bar{\Gamma}(\psi(w))]^{-1} \tag{27}$$

$$S'(y) = \begin{cases} (r + \mu)^{-1} & \text{if } y < \psi_1 \\ \frac{(r+2\mu+\delta+h(p'(y)-1))}{(r+2\mu+\delta-h)(r+\mu+h)} & \text{if } y \geq \psi_1 \end{cases} \tag{28}$$

$\phi_0$ ,  $\psi_0$ , and  $w^*$  are calculated by solving the equalities  $W(\phi_0) = U$ ,  $S(\psi_0) = U$ , and  $U = W^s$ . They are given by Eqs. (29), (30), and (31), where  $W'(\cdot)$  and  $S'(\cdot)$  are given by Eqs. (27) and (28).

$$\phi_0 = (b - a) + (\lambda_0 - \lambda_1) \int_{\phi_0} [W'(x)\bar{F}(x)] dx + (\eta_0 - \eta_1) \int_{\psi_0} [S'(z)\bar{\Gamma}(z)] dz \tag{29}$$

Assuming self-employed agents with no intention to recruit exist. Then,  $\psi_0$  is given by the solution to the equality  $S_O(\psi_0) = U$ . The value of  $y$  that solves this equality is labelled  $\tilde{\psi}_0$ .

$$\tilde{\psi}_0 = b + \lambda_0 \int_{\phi_0} [W'(x)\bar{F}(x)] dx + \eta_0 \int_{\psi_0} [S'(z)\bar{\Gamma}(z)] dz \tag{30}$$

As discussed,  $w^*$  is solved for as the solution to  $W^s = U$ . Looking at Eqs. (16), (17), and (30), it is clear that  $w^*$  can be expressed as in Eq. (31).

$$w^* = \tilde{\psi}_0 - a \tag{31}$$

However, if agents always intend to recruit, the solution is given by  $S(\psi_0) = U$ , where  $\psi_0 > \psi_1$ ; this value for  $\psi_0$  is denoted as  $\check{\psi}_0$ . The explicit solution for  $\check{\psi}_0$  depends on the parameterization of  $p(y)$ .

The minimum productivity required for an agent to leave unemployment for self-employment, is thus

$$\psi_0 = \min(\tilde{\psi}_0, \check{\psi}_0) \tag{32}$$

$\phi(y)$  is the solution to the equality  $S(y) = W(\phi(y))$ . Differentiating both sides gives a neat ODE in  $\phi(y)$ ,  $S'(y) = W'(\phi(y))\phi'(y)$ . Substituting in the expressions for the derivatives of the value functions, Eqs. (27) and (28), gives the ODE

$$\phi'(y) = \begin{cases} \frac{(r+\mu+\delta+\lambda_1\bar{F}(\phi(y))+\eta_1\bar{\Gamma}(y))}{(r+\mu)} & \text{if } y < \psi_1 \\ \frac{(r+2\mu+\delta+h(p'(y)-1))(r+\mu+\delta+\lambda_1\bar{F}(\phi(y))+\eta_1\bar{\Gamma}(y))}{(r+2\mu+\delta-h)(r+\mu+h)} & \text{if } y \geq \psi_1 \end{cases} \tag{33}$$

with the initial condition  $\phi(\psi_0) = \phi_0$ .

#### Solving for the steady state

Differentiating Eq. (20) with respect to  $y$  gives

$$N_u \eta_0 \gamma(y) + N_e^f \eta_1 G(\phi(y)) \gamma(y) = \begin{cases} \mu \gamma_s(y) & \text{if } y < \psi_1 \\ \left( \mu + \frac{(\mu+\delta)\Sigma(1)}{N_s \bar{\Gamma}_s(\psi_1)} \right) \gamma_s(y) & \text{if } y \geq \psi_1 \end{cases} \tag{34}$$

where,  $\gamma_s(y) = \frac{d}{dy} \{ \Gamma_s(y) \}$ .

This would be a straightforward ODE in  $N_s\Gamma(y)$  if it was not for the term including  $N_e^f G(\phi(y))$ . However, one can isolate this term from the steady state condition for paid employees, Eq. (21). Summing across Eqs. (21) and (20) gives

$$N_s\mu\Gamma_s(y) + \mathbb{I}_{\{y \geq \psi_1\}} \left( \frac{(\mu + \delta)\Sigma(1)}{N_s\bar{\Gamma}_s(\psi_1)} \right) (\Gamma_s(y) - \Gamma_s(\psi_1)) + \quad (35)$$

$$N_e G(\phi(y)) [\mu + \delta + \lambda_1 \bar{F}(\phi(y)) + \eta_1 \bar{\Gamma}(y)] = N_u [\eta_0 \Gamma(y) - \eta_0 \Gamma(\psi_0) + \lambda_0 F(\phi(y))]$$

where  $\mathbb{I}_{\{y \geq \psi_1\}}$  is an indicator function taking the value one if  $\{y \geq \psi_1\}$  is satisfied and zero otherwise.

The Markov process that determines the size of the firm is determined only by  $\mu$  and  $h$ . Thus, the ergodic distribution of self-employed firm sizes can be computed for a given  $h$  and  $\mu$  by simulating the Markov process for a sufficiently long period. Let  $s(\ell)$  be the distribution of employee numbers amongst the self-employed, then the measure  $\Sigma(\ell)$  is given by the distribution weighted by the number of self-employed who intend to recruit (of productivity greater than  $\psi_1$ ).

$$N_s \bar{\Gamma}(\psi_1).s(\ell) \quad (36)$$

$N_e^s$  the number of paid employees employed by the self-employed is given by the accounting identity in Eq. (37).

$$N_e^s = \sum_{\ell=1}^{\infty} \Sigma(\ell)\ell \quad (37)$$

Solving for the steady state is fairly cumbersome and requires an iterative solution. Initially,  $s(\ell)$  is computed by simulating the Markov process. The outer loop iterates around the measure of potential recruiters  $N_s \bar{\Gamma}(\psi_1)$  and the inner loop around the measure of unemployed. An initial guess is made regarding the number of potential recruiters, from which  $\Sigma(\ell)$  is calculated (Eq. (36)) and so is  $N_e^s$  (Eq. (37)). In the inner loop, the ODE (34) with the initial condition  $\Gamma_s(\psi_0) = 0$  is solved, where  $N_e^f G(\phi(y))$  is imputed using Eq. (35). Then,  $N_u$  is updated according to Eq. (18). Once the procedure converges to a solution, it goes to the outer loop. The number of potential recruiters has only been used to determine the measure  $\Sigma(\ell)$  (and  $N_e^s$ ); it is updated using the solution to the differential equation in the inner loop. The whole process is iterated on until stable.

### Misreporting of earnings by self-employed

The data on income that is relied upon in the estimation are collected from survey data. While there are no clear incentives to lie about one's income in a survey, where there would be, to say, tax authorities, there exists a literature that suggests people answer or behave differently when being studied. This is commonly known as the *Hawthorne effect*<sup>8</sup>. To calculate the degree of misreporting, this subsection follows the methodology proposed by Hurst et al. (2014). Relying on consumption and income data, which are obtained from the Expenditure and Food Survey (EFS), information is obtained about total personal weekly consumption, gross weekly income, the employment status of the individual, and a variety of demographic information. The sample is restricted to males in employment.

There are four identifying assumptions that allows the uncovering of the degree of misreporting. The income and expenditure relationship is governed by the log-linear Engel

curve. Self-employed agents systematically misreport their earnings by a factor  $\kappa_s$ , it need not be assumed that  $\kappa_s \leq 1$  and paid employees provide an unbiased reporting of their income. Finally, both the self-employed and paid employees provide unbiased reports of total expenditure.

Following the notation of Hurst et al. (2014),  $k$  denotes the employment status, ( $k = S$ ) implies an individual is self-employed and ( $k = W$ ) implies he is a paid employee, individual  $i$  has preferences that generate the log-linear Engel curve given by Eq. (38).

$$\log c_{ik} = \alpha + \beta \log y_{ik}^p + \Theta' X_{ik} + \epsilon_{ik} \quad (38)$$

$y_{ik}^p$  is the permanent income of individual  $i$  in labor market state  $k$ ,  $\beta$  is the income elasticity, and  $X_{it}$  is a vector of demographic controls. The vector of controls include a series of 5-year age dummies, a dummy for if the individual is white, a dummy for if the individual is married, and the number of children an individual has. Estimating (38) is problematic as the self-employed are systematically misreporting their income and the income that is observed is an individual's transitory rather than permanent income. It is assumed that reported income is governed by

$$\log y_{iW} = \log y_{iW}^p + \Omega' X_{iW} + v_{iW} \quad (39)$$

$$\log y_{iS} = \log \kappa_s + \log y_{iS}^p + \Omega' X_{iS} + v_{iS} \quad (40)$$

It is assumed that  $v$  has mean zero and is orthogonal to the unobserved determinants of consumption  $\epsilon$ . The self-employed are also systematically misreporting their income by a factor  $\kappa_s$ . Rearranging Eqs. (38), (39), and (40), gives the following, where  $\kappa_s$  can be backed out as  $\kappa_s = \exp(-\gamma/\beta)$ :

$$\log c_{ik} = \alpha + \beta \log y_{ik} + \gamma D_i + \Psi X_{ik} + \xi_{ik} \quad (41)$$

$D_i$  is a dummy variable taking the value one if individual  $i$  is self-employed,  $\Psi = \Theta - \beta' \Omega$  and the unobservable component is given by  $\xi_{ik} = \epsilon_{ik} - \beta v_{ik}$ . Even in the absence of measurement error, transitory income fluctuations governed by Eqs. (39) and (40) introduce attenuation bias in the estimate of  $\beta$  since  $E[\log y_{ik} \xi_{ik}] \neq 0$ . To overcome this problem, the age at which an individual left full-time education is used as an instrument for reported income. Year left education is split into five categories: left before 16, left at 16 or 17, left at 18, left at 19 or 20, and left at 21 or over. Thus, for these instruments to be valid, it is assumed that education affects consumption only through changes in permanent income. The results of  $\kappa_s$ ,  $\gamma$ , and  $\beta$  are given in Table 10, using ordinary least squares and the preferred instrumental variables estimation procedure.

Standard errors in parentheses are obtained using a bootstrap procedure, redrawing the sample with repetition 500 times. The estimation is based upon earnings for the self-employed which have been adjusted by the factor  $\kappa_s$ .

**Table 10** Fraction of under reported income by self-employed

	IV	OLS
$\beta$	0.913 (0.184)	0.368 (0.049)
$\gamma$	0.474 (0.117)	0.264 (0.090)
$\kappa_s$	0.595 (0.065)	0.489 (0.124)

### Transition rate moment conditions

The model is in continuous time, with one unit representing a month. Empirically, we observe the monthly rate; the moment and its theoretical counterpart are given below:

$$\begin{aligned}
 &\text{Unemployment to paid employment: } 1 - \exp(-\lambda_0 - \lambda_0^s) \\
 &\text{Unemployment to own-account self-employment: } 1 - \exp(-\eta_0 \bar{\Gamma}(\psi_0)) \\
 &\text{Paid employment (in large firm) to unemployment: } 1 - \exp(-\delta) \\
 &\text{Paid employment (any) to unemployment: } 1 - \exp\left(-\frac{N_e^s}{N_e^s + N_e^f}(\mu + \delta) - \frac{N_e^f}{N_e^s + N_e^f}\delta\right) \\
 &\text{Paid employment to another paid employer: } 1 - \exp\left(-\frac{N_e^f \lambda_1}{N_e^f + N_e^s} \int_{\phi_0}^{\infty} \bar{F}(x) dG(x)\right) \\
 &\text{Paid employment to own-account self-employment:} \\
 &1 - \exp\left(-\frac{N_e^f \eta_1}{N_e^f + N_e^s} \int_{\psi_0}^{\infty} \bar{\Gamma}(x) dG(\phi(x))\right) \\
 &\text{Own-account self-employment to recruiter: } 1 - \exp(-h \bar{\Gamma}_s(\psi_1))
 \end{aligned}$$

Note,  $\lambda_0^s$ , the arrival rate of job offers to the unemployed from self-employed recruiters is given by

$$\lambda_0^s = \frac{h(N_e^s + \Sigma(0))}{N_u}$$

The numerator is the total number of hires, and the denominator is the mass of potential recipients.

### Competing interests

The IZA Journal of Labor Economics is committed to the IZA Guiding Principles of Research Integrity. The author declares that he has observed these principles.

### Acknowledgements

The author would like to thank two anonymous referees and the editor of this journal, Pierre Cahuc for his insightful and constructive comments. I would also like to thank Jim Albrecht, Etienne Lalé, Fabien Postel-Vinay, Robert Shimer, Héléne Turon, Ludo Visschers, and Susan Vroman as well as seminar audiences at the Barcelona Graduate School of Economics, University of Bristol, The Search and Matching Conference 2014 at the University of Edinburgh, The International Association for Applied Econometrics 2014 at Queen Mary College, University of London, and The Bank of England for all their helpful insights. All errors are, of course, my own. Email: [jb683@cam.ac.uk](mailto:jb683@cam.ac.uk).  
Responsible Editor: Pierre Cahuc

Received: 1 February 2016 Accepted: 10 May 2016

Published online: 10 June 2016

### References

- Altonji JG, Segal LM (1996) Small-sample bias in GMM estimation of covariance structures. *J Bus Econ Stat* 14:353–366
- Black J, de Meza D, Jeffreys D (1996) House prices, the supply of collateral and the enterprise economy. *Econ J* 106(434):60–75
- Blanchflower DG (2004). NBER Working Paper, No. 10286, <https://ideas.repec.org/p/nbr/nberwo/10286.html>
- Blanchflower DG, Oswald AJ (1998) What makes an entrepreneur? *J Labor Econ* 16(1):26–60
- Bontemps C, Robin JM, van den Berg GJ (2000) Equilibrium search with continuous productivity dispersion: theory and nonparametric estimation. *Int Econ Rev* 41(2):305–58
- Bradley J, Postel-Vinay F, Turon H (2015) Public sector wage policy and labor market equilibrium: a structural model. <https://ideas.repec.org/p/spo/wpmain/infohd12441-23tsou9ve8lubqe4pa2of8d7d.html>
- Burdett K, Mortensen DT (1998) Wage differentials, employer size, and unemployment. *Int Econ Rev* 39(2):257–73
- Coles MG (2001) Equilibrium wage dispersion, firm size and growth. *Rev Econ Dyn* 4(1):159–187
- Coles MG, Mortensen DT (2011). NBER Working Paper, No. 17284, <https://ideas.repec.org/p/nbr/nberwo/17284.html>
- Cowling M, Mitchell P (1997) The evolution of u.k. self-employment: a study of government policy and the role of the macroeconomy. *Manch Sch* 65(4):427–442
- Haltiwanger J, Jarmin R, Miranda J (2013) Who creates jobs? Small versus large versus young. *Rev Econ Stat* 95(2):347–361
- Hamilton BH (2000) Does entrepreneurship pay? An empirical analysis of the returns to self-employment. *J Polit Econ* 108(3):604–631
- Hébert RF, Link AN (1988) *The entrepreneur: mainstream views and radical critiques*. Second Edition. Praeger, New York
- Hurst E, Li G, Pugsley B (2014) Are household surveys like tax forms? Evidence from income underreporting of the self employed. *Rev Econ Stat* 96(1):19–33
- Kumar A, Schuetz HJ (2007) Self-employment and labor market policies. <https://ideas.repec.org/p/vic/vicddp/0704.html>

- Levitt SD, List JA (2011) Was there really a Hawthorne effect at the Hawthorne plant? An analysis of the original illumination experiments. *Am Econ J Appl Econ* 3(1):224–238
- Lindsay C, Macauley C (2004) Growth in self-employment in the UK. *Labour Mark Trends* 112:399–404
- Margolis DN, Navarro L, Robalino DA (2014) Unemployment insurance, job search and informal employment. *Social insurance, informality and labor markets: how to protect workers while creating good jobs*. Oxford University Press, Oxford
- Meghir C, Narita R, Robin JM (2015) Wages and informality in developing countries. *Am Econ Rev* 105(4):1509–1546
- Millán A (2012) The choice to become self-employed: acknowledging frictions. [http://www.uab.cat/doc/DOC\\_Ana\\_Millan](http://www.uab.cat/doc/DOC_Ana_Millan)
- Moralee L (1998) Self-employment in the 1990s. *Labour Mark Trends* 106:121–130
- Narita R (2014) Self employment in developing countries: a search-equilibrium approach. <https://ideas.repec.org/p/spa/wpaper/2013wpecon21.html>
- Parker SC (2004) *The economics of self-employment and entrepreneurship*. No. 9780521030632 in Cambridge Books. Cambridge University Press, Cambridge
- Santarelli E, Klomp L, Thurik AR (2006) Entrepreneurship, Growth, and Innovation: the Dynamics of Firms and Industries. In: Santarelli E (ed). *Gibrat's Law: an Overview of the Empirical Literature*. Springer Science, Berlin. pp 41–73
- Schumpeter JA (1934) *The theory of economic development: an inquiry into profits, capital, credit, interest and the business cycle*. Harvard University Press, Cambridge, MA. Reprinted in 1962

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](http://springeropen.com)

---