RIETI Discussion Paper Series 09-E-002

Are Contingent Jobs Dead Ends or Stepping Stones to Regular Jobs? Evidence from a Structural Estimation

Julen ESTEBAN-PRETEL
the University of Tokyo

NAKAJIMA Ryo
University of Tsukuba

TANAKA Ryuichi
Tokyo Institute of Technology

Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

Are Contingent Jobs Dead Ends or Stepping Stones to Regular Jobs? Evidence from a Structural Estimation *

Julen Esteban-Pretel
University of Tokyo

Ryo Nakajima†
University of Tsukuba

Ryuichi Tanaka
Tokyo Institute of Technology

January 16, 2009

Abstract

The proportion of part-time, dispatch, and temporary workers has increased in many developed economies in recent years. These workers receive lower average wages and benefits, and are subject to lower employment stability. This paper analyzes the effects of initially taking such jobs on the employment careers of young workers. We build an on-and-off-the-job search model, using Japanese data to perform a structural estimation of the model parameters and simulate career paths, in order to study the effects of the initial choice of employment on the probability of having a regular job in the future and on the welfare of the worker. We find that although contingent jobs are neither stepping stones towards regular employment nor dead-ends, starting a career in a contingent job has a lasting effect on the welfare of the individual in Japan.

Keywords: contingent employment, job transition, on-and-off-the-job search structural estimation.

JEL Classification: I30, J21, J63, J64.

*This research is a part of the project "Empirical Analysis of Japan's Labor Market" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). The opinions expressed and arguments employed in this paper are the sole responsibility of the author and do not necessarily reflect those of RIETI or the Ministry of Economy, Trade and Industry of Japan. We are grateful to Hidehiko Ichimura, Daiji Kawaguchi, Katsuya Takii and the seminar participants at GRIPS, Kansai Labor Workshop, Kyushu University and RIETI for their helpful comments. The financial support of JSPS is acknowledged. Of course, all remaining errors are solely our own.

†Corresponding author: Doctoral Program in Economics, Graduate School of Humanities and Social Sciences, University of Tsukuba, 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8571, Japan; E-mail: nakajima@dpipe.tsukuba.ac.jp; Tel: +81-29-853-5064; Fax:+81-29-853-7440

1 Introduction

In recent years, part-time and temporary employment have increased dramatically in many developed economies. Countries such as France, Japan, and Spain, traditionally characterized by full-time, long-term contracts and high worker protection, have seen the proportion of contingent employment[1] rise by 50 percent or more in the last two decades[2]. One of the reasons for this upward trend in contingent employment is the introduction of legislation to flexibilize the labor market and reduce costs to employers. European countries have started to rely to a greater extent on temporary or fix-term workers, whereas in Japan it is part-time employment that has seen the greatest increase.

Of primary concern with regard to contingent employment are the differences in wages, fringe benefits and protection that workers in these types of jobs obtain with respect to regular employees. This concern is more serious in a country such as Japan where, despite the fact that 30 percent of part-time employees work on average the same amount of hours as full-time ones, they have traditionally performed different tasks and are not considered equally within the firm. In most cases, this translates into the part-time workers being excluded from important fringe benefits such as training and medical coverage[3], and in being less protected from dismissal. Although it has been argued that part-time jobs may help unemployed workers to get started in the job market, and that these can later have a stepping-stone effect toward finding a regular job, becoming a contingent employee may make it more difficult to find a better (regular) job in the future should taking such a job send a negative signal about the unobserved skills of the worker. The "stigma" effect associated with these jobs may have lasting effects on the employment prospects of a worker, and thus, young individuals taking such jobs could be affected if this "stigma" is substantial.

This paper studies the effects of contingent employment on the transition probabilities into regular employment for young, recently graduated male workers in Japan, and analyzes the implications of the initial job on the welfare of the worker. We formulate a three state on-and-off-the-job search model, where workers can either be unemployed, employed in a contingent job or employed in a regular job. Then, using data on young male workers from the 2002 Employment Status Survey (ESS), we estimate the structural parameters of the model. These parameter estimates are used to perform simulations of the career paths of workers, starting from the different states in order to study whether we observe the stepping stone effect for young workers who experience periods of contingent employment at the beginning of their careers. Moreover we analyze, using career path simulations, the implications on

[1]We use the term contingent worker to refer to those employees who are hired part-time, on a temporary contract or as dispatched workers.

[2]See Table 1.1 of Houseman and Osawa (2003) for details.

[3]In the case of health insurance, the employer does not provide coverage if the annual income of the worker is below 1.3 million yen and he or she is married to a full-time worker.

employment stability, welfare, and inequality of the initial employment state of the young worker. We also perform counterfactual experiments to study the effects of policy changes on the transition probabilities and welfares of the workers.

The main finding of the paper is that contingent jobs are neither dead ends nor stepping stones towards regular employment, although the welfare effects of starting an employment career in a contingent job are important. We find that in the short run, less that 10 years after graduation, the conditional probability of being employed in a regular job is lowest for individuals who start their careers in contingent jobs, lower even than for those initially unemployed, which in turn is lower than for those initially employed in a regular job. This indicates that contingent jobs do not seem to serve as stepping stones toward regular jobs. However, after 20 years, the initial employment status ceases to have an effect on the probability of having a regular job, which implies that contingent jobs are also not dead-ends. This long-run unimportance of the initial state in terms of the the probability of finding employment in a regular job does not carry over when welfare is analyzed. We find that a worker's welfare for the first 40 years of his life is lower if he begins in a contingent job than if he starts in a regular job, but higher than if he is initially unemployed. This implies that although on average all workers have the same probability of having a regular job after 20 years of being in the labor market, the effects on their welfare are important if they start in a contingent job, and even more so if they initially do not have a job.

The counterfactual experiments show that policies which increase the probability of contingent workers to move into regular positions increase the fraction of regular employment in the economy in the long run, increase welfare, and reduce inequality. Also interesting are policies that reduce the destruction of contingent jobs, since while they make this type of employment more persistent, and hence reduce the chances of workers moving into regular employment, they also improve the welfare of the workers who start their careers in contingent jobs.

Our paper is related to two strands of labor literature. First, it connects to papers studying the employment choices of young workers. In terms of methodology, the paper closest to this one is Flinn (2002). He uses a job search model and data on the employment choices of young workers in Italy and the U.S. to study the differences in cross-sectional wages and life-time welfare distributions of workers in these two countries. Although the question addressed in his paper is different from ours, the estimation procedure here closely follows that used by Flinn. Two other related papers which study the employment choices of young workers are by Genda and Kurosawa (2001), which uses Japanese data, and Gaston and Timcke (1999), which performs the analysis for the Australian economy. In the former study, the authors carry out an empirical evaluation of the effects of labor market conditions at the time of the first employment on the career prospects of workers. They find that young workers who enter the labor market during a downturn have a lower probability of finding full-time jobs and take lower quality jobs, which in turn increases their probability of leaving that employer in the future. The latter study the transition from casual to full-time jobs for young workers and the factors which affect that transition. They find that unemployment and non-fringe benefits, such as on-the-job training, affect these transitions in the short-term but not in the long run, and conclude that casual employment can be considered neither a stepping stone nor a dead-end with regard to full-time employment. However, these two papers are largely descriptive, and propose no model to account for the career choices of the workers. One paper which has a model and performs parametric estimation, although does not focus on young workers, is Bover and Gomez (2004). The authors investigate the variables which affect the choice of the worker between temporary and permanent jobs when exiting unemployment. They conclude that unemployment benefits are the crucial factor affecting the exit rate towards temporary jobs, whereas business cycle variables, such as GDP growth, are those affecting the exit rate toward permanent jobs. Our paper contributes to the literature on young workers' employment choices by analyzing a three state job search model (Flinn, 2002, uses unemployment and employment as the only possible states) and performing parametric estimation of the transition probabilities between these states to understand the extent to which contingent employment is a good choice for young workers.

The second strand of literature to which this paper relates is that which studies the effects on career opportunities for workers who take contingent jobs. One paper closely related to ours is Kondo (2007). She studies the effect on young Japanese workers of not getting a regular job at the start of their career. She finds that the probability for a worker to have regular job is 40 to 50% lower if he did not initially obtain such type of job. This result is not necessarily in contradiction with our findings, since in her sample workers are interviewed, on average, 10 years after graduation. We also obtain that after 10 years from graduation the probability of having a regular job for a worker who did not have it at the beginning is lower, although only around 20% lower. The differences form both sets of results are probably due to the different data and methodologies used in the analyses.[4] Several other papers have analyzed the potential stepping stone effect of contingent work for European countries, although the conclusions of these studies are mixed. Most of these papers consider only temporary employment, since fix-term jobs have been the ones increasing most rapidly in Europe, and are therefore the biggest challenge to understand. Four representative examples of this literature are as follows. Booth, Francesconi and Frank (2002) find that for the U.K., there is little evidence of the stepping stone effect for male workers, while there is some evidence for female workers. Casquel and Cunyat (2005), using Spanish data, show that some workers are able to use temporary employment as a stepping stone toward regular employment, whereas for others (young workers, women, and workers

[4] In another related paper, Genda (2008) estimates the probability that a contingent worker get a regular job by the probit model, and finds that job tenure of the previous contingent job is one of the most important determinants of getting a regular job.

with "bad employment history"), temporary jobs are more likely to lead to dead-ends. Kvasnicka (2005) finds that in Germany, performing temporary work out of unemployment neither harms nor increases the chances of finding employment outside of the temporary help agencies in the short-run. More optimistic findings are shown in Zijl, van den Berg and Heyma (2004), who find evidence of the stepping stone effect for Dutch workers in temporary jobs. One paper which stands out is Autor and Houseman (2008), since they utilize U.S. experimental data and study the chances of disadvantaged workers who are employed by temporary help agencies to obtain regular jobs in the future. They find that workers who transition via these types of agencies have lower chances of being hired in regular jobs than those who find a job by direct hire. This is seen as evidence for temporary agency jobs being dead-ends for disadvantaged workers. Most of the previous papers use hazard models to non-parametrically estimate the transition probabilities between the different employment states. In our paper, we perform parametric estimation of a job search model, which allows us to study potential career paths for workers both in the short- and long run, and also enables us to perform welfare analysis.

The remainder of the paper is organized as follows: Section 2 explains the behavioral model; Section 3 describes the data; Section 4 shows the econometric procedure used to estimate the parameters of the model; Section 5 shows the results of the estimation; Section 6 shows the counterfactual policy simulations; and Section 7 concludes.

2 The Behavioral Model

In this section, we present the behavioral model that will be used to perform the empirical analysis and simulations below.

The model is a continuous time job search model. We assume that individuals are infinitely lived, have linear utility, and discount the future at rate ρ. At any point in time, they are in one of the three following states: unemployed, state denoted by u; employed in a regular job, state denoted by r; or employed at a contingent job, state denoted by c. In each of the three states, the worker obtains a flow value of b_k where $k \in \{u, r, c\}$. The unemployment benefit or the flow value of leisure is represented by b_u The fringe benefits are represented by b_r and b_c for regular and contingent jobs, respectively. [5] The offers from regular and contingent jobs reach individuals according to a Poisson process with arrival rate λ_k^l, where $k \in \{u, r, c\}$ represents the current state of the worker and $l \in \{r, c\}$ is the type of job offer. The wage associated with a job offer is independent of its arrival rate, and is i.i.d. from a distribution F^k. Employment relationships are terminated when the employee chooses to move to another job or for exogenous reasons. Exogenous separations occur according to a Poisson process with arrival rate η_k, where $k \in \{r, c\}$ is the state of the worker before he is terminated.

[5] The fringe benefits can include compensations such as medical benefits, better facilities and paid training.

We now present the decision problem of a worker as a recursive problem in which he chooses whether or not to accept an offered job in order to maximize his lifetime utility. The values of the recursive problem are denoted by V^u, $V^r(w)$, and $V^c(w)$ for a worker to be unemployed, employed at a regular job, and at a contingent job at wage w, respectively.

2.1 Value Functions

The value of unemployment for a worker is the composite of the following terms. In every infinitesimally small period of time Δt, a worker receives a benefit of b_u. He receives an offer of a regular job with arrival rate λ_u^r, in which case he must decide whether to take the job at the offered wage or to remain employed. In that interval of time he may also receive a contingent job offer, which occurs with arrival rate λ_u^c and must then decide if it is more profitable to take it or remain unemployed. If the worker does not receive any offers, or rejects the offers he gets, he remains unemployed. We can write the value for a worker to be unemployed as:

$$\begin{aligned} V^u &= (1+\rho\Delta t)^{-1}\{b_u\Delta t + (1-\lambda_u^r\Delta t - \lambda_u^c\Delta t)\,V^u \\ &\quad +\lambda_u^r\Delta t\int \max\{V^u, V^r(w^r)\}dF^r(w^r) \\ &\quad +\lambda_u^c\Delta t\int \max\{V^u, V^c(w^c)\}dF^c(w^c) + o(\Delta t)\}, \end{aligned} \tag{1}$$

where $(1+\rho\Delta t)^{-1}$ is the infinitesimal discount factor associated with time period Δt. $\lambda_u^j\Delta t$ for $j\in\{r,c\}$ is the approximate probability of receiving an offer from a regular or a contingent job respectively in the interval Δt. Similarly, $(1-\lambda_u^r\Delta t-\lambda_u^c\Delta t)$ is the approximate probability of remaining unemployed after Δt. $o(\Delta t)$ is an error term, which has the property that $\lim_{\Delta t\to 0}\frac{o(\Delta t)}{\Delta t}=0$.

We can write the value for a worker to be employed at a regular job with wage w in a similar way. The difference in the value of the unemployment status is that in every interval of time Δt, a regular employee receives benefit b_r and wage w and faces a probability of losing the job of η_r.

$$\begin{aligned} V^r(w) &= (1+\rho\Delta t)^{-1}\{b_r\Delta t + w\Delta t + \eta_r\Delta t V^u \\ &\quad +(1-\eta_r\Delta t - \lambda_r^r\Delta t - \lambda_r^c\Delta t)\,V^r(w) \\ &\quad +\lambda_r^r\Delta t\int \max\{V^r(w), V^r(w^r)\}dF^r(w^r) \\ &\quad +\lambda_r^c\Delta t\int \max\{V^r(w), V^c(w^c)\}dF^c(w^c) + o(\Delta t)\}. \end{aligned} \tag{2}$$

Finally, the value of being employed at a contingent job at wage w is:

$$\begin{aligned} V^c(w) &= (1+\rho\Delta t)^{-1}\{b_c\Delta t + w\Delta t + \eta_c\Delta t V^u \\ &\quad + (1-\eta_c\Delta t - \lambda_c^r\Delta t - \lambda_c^c\Delta t)\, V^c(w) \\ &\quad + \lambda_c^r\Delta t \int \max\{V^c(w), V^r(w^r)\} dF^r(w^r) \\ &\quad + \lambda_c^c\Delta t \int \max\{V^c(w), V^c(w^c)\} dF^c(w^c) + o(\Delta t)\} \end{aligned} \tag{3}$$

Collecting terms in (1)-(3), dividing throughout by Δt and taking the limit when $\Delta t \to 0$, we obtain the continuous time formulation for the value of unemployment:

$$\begin{aligned} \rho V^u &= b_u + \lambda_u^r \int \max\{0, V^r(w^r) - V^u\} dF^r(w^r) \\ &\quad + \lambda_u^c \int \max\{0, V^c(w^c) - V^u\} dF^c(w^c). \end{aligned} \tag{4}$$

The value functions of regular and contingent jobs are given by

$$\begin{aligned} \rho V^r(w) &= b_r + w + \eta_r\left(V^u - V^r(w)\right) \\ &\quad + \lambda_r^r \int \max\{0, V^r(w^r) - V^r(w)\} dF^r(w^r) \\ &\quad + \lambda_r^c \int \max\{0, V^c(w^c) - V^r(w)\} dF^c(w^c), \end{aligned} \tag{5}$$

$$\begin{aligned} \rho V^c(w) &= b_c + w + \eta_c\left(V^u - V^c(w)\right) \\ &\quad + \lambda_c^r \int \max\{0, V^r(w^r) - V^c(w)\} dF^r(w^r) \\ &\quad + \lambda_c^c \int \max\{0, V^c(w^c) - V^c(w)\} dF^c(w^c). \end{aligned} \tag{6}$$

2.2 Reservation Wages

We now proceed to show that this model satisfies the reservation wage property. This property implies that a worker will accept a job if the wage offered is above a certain threshold, called the reservation wage.

Proposition 1 *There exists a unique reservation wage w^{k*} such that $V^k(w) \geq V^u$ if and only if $w \geq w^{k*}$ for $k \in \{r, c\}$*

This proposition directly follows the following lemma.

Lemma 2 *$V^r(w)$ and $V^c(w)$ are strictly increasing in w.*

Proof. To show that $V^r(w)$ is strictly increasing in w we proceed by contradiction. Suppose that $V^r(w)$ is non-increasing in w. Then, the right-hand side of (5) should also be non increasing. However, we can clearly see that if $V^r(w)$ is non-increasing in w, the right-hand side of (5) is in fact strictly increasing in w, which is a contradiction. The monotonically increasing property of $V^c(w)$ can be established in the same manner by using equation (6). ■

Given the reservation wages, w^{r*} and w^{c*}, the decision rule for an unemployed individual is simple. If a worker is unemployed and is offered a regular job at wage w, the decision rule is to accept the job only when $w \geq w^{r*}$. Similarly, if he receives a contingent job offer, the decision rule is to accept if the offered wage is above the contingent job reservation wage w^{c*}.

We can also characterize the reservation wage rule for employed individuals. To do so, we need to distinguish between the cases in which the job switch is within the same type of job and when the change is between different job types. For the within-type-job-change, if a worker employed at a regular job at wage w receives a regular job offer of wage x, he accepts the offer and changes the job if and only if $V^r(x) \geq V^r(w)$. The monotonic increasing property of the value function, as shown by the previous lemma, implies that the job change occurs if and only if $x \geq w$. Thus, the reservation wage of an employed individual is exactly equal to the wage of the current job. The same reservation wage rule is applied for a worker who is employed by a contingent job at wage w and receives a contingent job offer of wage x. That is, he changes jobs if and only if $x \geq w$.

For between-type-job-changes, we can show that the reservation wage for employed individuals is given by a function of the current job wage. The following propositions insure the existence and monotonically increasing property of the reservation wage functions.

Proposition 3 *For given w, there exists a unique reservation wage $w^{r*}(w)$ such that $V^r(x) \geq V^c(w)$ if and only if $x \geq w^{r*}(w)$. Similarly, there exists a unique reservation wage $w^{c*}(w)$ such that $V^c(x) \geq V^r(w)$ if and only if $x \geq w^{c*}(w)$.*

Proof. The proof is obvious from the previous lemma. ■

Proposition 4 *$w^{r*}(w)$ and $w^{c*}(w)$ are increasing in w.*

Proof. To prove that $w^{r*}(w)$ is strictly increasing in w we proceed by contradiction. Suppose that $w^{r*}(w)$ is non-increasing in w. Then there exist w and w' such that $w > w'$ and $w^{r*}(w) \leq w^{r*}(w')$. Since $V^r(w)$ and $V^c(w)$ are strictly increasing in w, $V^c(w) > V^c(w')$ and $V^r(w^{r*}(w)) \leq V^r(w^{r*}(w'))$. However, these inequalities contradict the conditions that the reservation wage functions should satisfy, which are $V^c(w) = V^r(w^{r*}(w))$ and $V^c(w') = V^r(w^{r*}(w'))$. Therefore, $w^{r*}(w)$ is strictly increasing in w. Applying a similar argument, we can prove that $w^{c*}(w)$ is increasing in w. ■

Using the reservation wage functions, $w^{r*}(w)$ and $w^{c*}(w)$, the worker's decision rule for job-to-job transition is characterized as follows. If a worker currently employed by a regular job at wage w receives a contingent job offer of wage x, the decision rule is to accept the contingent job if and only if $x \geq w^{c*}(w)$. Similarly, if a worker is currently employed by a contingent job at wage w and receives a regular job offer of wage x, he moves to the regular job if and only if $x \geq w^{r*}(w)$.

2.3 The System of Integral Equations

We now show that the recursive search model presented above can be reduced to a system of two integral equations.

Under the reservation wage property, the value functions of the model must satisfy the following four conditions:

$$V^u = V^r(w^{r*}) \tag{7}$$

$$V^u = V^c(w^{c*}) \tag{8}$$

$$V^c(w) = V^r(w^{r*}(w)) \tag{9}$$

$$V^r(w) = V^c(w^{c*}(w)) \tag{10}$$

Applying these relations to the value functions (4)-(6), we obtain the following characterizations of the value functions:

$$V^u = \frac{b_u + \lambda_u^r \int_{w^{r*}} V^r(w^r)dF^r(w^r) + \lambda_u^c \int_{w^{c*}} V^c(w^c)dF^c(w^c)}{\rho + \lambda_u^r(1 - F^r(w^{r*})) + \lambda_u^c(1 - F^c(w^{c*}))} \tag{11}$$

$$V^r(w) = \frac{b_r + w + \eta_r V^u + \lambda_r^r \left[\int_w V^r(w^r)dF^r(w^r)\right] + \lambda_r^c \left[\int_{w^{c*}(w)} V^c(w^c)dF^c(w^c)\right]}{\rho + \eta_r + \lambda_r^r(1 - F^r(w)) + \lambda_r^c(1 - F^c(w^{c*}(w)))} \tag{12}$$

$$V^c(w) = \frac{b_c + w + \eta_c V^u + \lambda_c^r \left[\int_{w^{r*}(w)} V^r(w^r)dF^r(w^r)\right] + \lambda_c^c \left[\int_w V^c(w^c)dF^c(w^c)\right]}{\rho + \eta_c + \lambda_c^r(1 - F^r(w^{r*}(w))) + \lambda_c^c(1 - F^c(w))} \tag{13}$$

Using conditions (7) and (8), we can reduce the value functions into a system of two integral equations for $V^r(w)$ and $V^c(w)$[6]:

$$V^r(w) = \frac{b_r + w + \eta_r \left[\frac{b_r + w^{r*} + \lambda_r^r A(w^{r*}) + \lambda_r^c B(w^{c*})}{\rho + \lambda_r^r(1 - F^r(w^{r*})) + \lambda_r^c(1 - F^c(w^{c*}))}\right] + \lambda_r^r A(w) + \lambda_r^c B(w^{c*}(w))}{\rho + \eta_r + \lambda_r^r \tilde{F}^r(w) + \lambda_r^c \tilde{F}^c(w^{c*}(w))}, \tag{14}$$

$$V^c(w) = \frac{b_c + w + \eta_c \left[\frac{b_c + w^{c*} + \lambda_c^r A(w^{r*}) + \lambda_c^c B(w^{c*})}{\rho + \lambda_c^r(1 - F^r(w^{r*})) + \lambda_c^c(1 - F^c(w^{c*}))}\right] + \lambda_c^r A(w^{r*}(w)) + \lambda_c^c B(w)}{\rho + \eta_c + \lambda_c^r \tilde{F}^r(w^{r*}(w)) + \lambda_c^c \tilde{F}^c(w)}. \tag{15}$$

where we define $A(w) \equiv \int_w V^r(x)dF^r(x)$ and $B(w) \equiv \int_w V^c(x)dF^c(x)$. We see that the value functions $V^r(w)$ and $V^c(w)$, which appear both as arguments in the previous integrals and as dependent variables, are the solutions of the system of the integral equations (14) and (15).

Once the value functions $V^r(w)$ and $V^c(w)$ are determined, the conditions (9) and (10) imply that the reservation wage functions $w^{r*}(w)$ and $w^{c*}(w)$, can be obtained as the inverse functions of the value functions as shown by the following conditions:

$$w^{r*}(w) = [V^r]^{-1}(V^c(w))$$
$$w^{c*}(w) = [V^c]^{-1}(V^r(w)).$$

The reservation wage functions that characterize the employed workers decision rule are functions of the following structural parameters:

- Unemployment benefit and employment (fringe) benefits: (b_u, b_r, b_c).
- Job separation rates: (η_r, η_c).
- Job arrival rates: $(\lambda_u^r, \lambda_u^c, \lambda_r^r, \lambda_r^c, \lambda_c^r, \lambda_c^c)$,
- The distributions of wage: (F^r, F^c).

These structural parameters are estimated from the data, as we shall see below.

[6]If the conditions (7) and (8) are applied to (12) and (13), we can find the alternative representations of V^u:

$$V^u = V^r(w^{r*}) = \frac{b_r + w^{r*} + \lambda_r^r \left[\int_{w^{r*}} V^r(w^r)dF^r(w^r)\right] + \lambda_r^c \left[\int_{w^{c*}} V^c(w^c)dF^c(w^c)\right]}{\rho + \lambda_r^r(1 - F^r(w^{r*})) + \lambda_r^c(1 - F^c(w^{c*}))},$$

$$V^u = V^c(w^{c*}) = \frac{b_c + w^{c*} + \lambda_c^r \left[\int_{w^{r*}} V^r(w^r)dF^r(w^r)\right] + \lambda_c^c \left[\int_{w^{c*}} V^c(w^c)dF^c(w^c)\right]}{\rho + \lambda_c^r(1 - F^r(w^{r*})) + \lambda_c^c(1 - F^c(w^{c*}))}.$$

Substituting these equations into equations (12) and (13) yield equations (14) and (15).

3 Data Description

The data used in our empirical analysis are taken from the 2002 Employment Status Survey (ESS). The survey was conducted on household members 15 years of age or older in about 440,000 households that were randomly selected based on the 2000 Japanese Census. The survey interviews provide information on the individuals' employment status and characteristics as of October 1st, 2002.

In the survey, respondents were asked to describe their labor force status as well as their background characteristics. If respondents were employed at the time of the interview, they reported detailed information on their current employment, such as employment duration, employment type, business contents, work contents, days worked, working hours per week and annual labor income. On the other hand, if respondents were not working at the time of the interview, they were asked to provide information such as non-employment duration, whether or not they wished to work, and the reason for wanting to work. For those who had prior working experience, selected information of the previous employment, such as the employment duration, employment type, business contents, and work contents were reported.

The subsamples used for the empirical work are selected in the following manner. We focus on young male workers between the ages of 15 and 25 who graduated no more than three years prior to the interview date. We estimate the individual school graduation date by birthdate and education level. [7] We also restrict all sampled individuals to be active labor force participants[8]. At the same time, employed persons who worked fewer than 25 days in a year are also excluded. In order to control for homogeneity of the sample, we also exclude individuals who were self-employed, individuals who were employed by a government corporation, and individuals who were, or had been, employed by small private companies with less than five employees. Given the sample selection criteria presented above, and after excluding individuals who reported missing or inconsistent labor market experience, the final sample size is 6257.

The selected individuals are categorized as either employed or unemployed [9]. Employed individuals are further subgrouped into two different types of employment, regular and contingent. Regular employment encompasses those jobs with full-time and long-term contracts, while contingent employment is defined as all other employment other than regular employment. Hence, part-time workers, contract company workers, and dispatched workers are all categorized under contingent employment.

In the estimation that follows, the definition of a duration spell plays an important role, so let us

[7]For example, a person with a high school degree who was born on February 1980 is assumed to have graduated from school in March 1999. We assume no school holdovers.

[8]Individuals who were enrolled in school and individuals who had no willingness to work at the time of interview are excluded from the sample.

[9]In this study, we do not attempt to distinguish between unemployment and non-employment. We classify as unemployed any individual who is willing to take a job if one is offered, whether or not he is actively looking for a job.

define it explicitly. We define the duration of an employment spell to be the number of months elapsing from the first calendar month of employment with a particular employer until the calendar month in which the worker left the job. Similarly, we define the duration of an unemployment spell as the number of calendar months during which the individual was not employed. Unemployment spells can be imputed in two different ways: first, using the number of months between the estimated calendar month of school graduation and the first calendar month of the first job; and second, using the number of months from the calendar month in which the worker left a job until the calendar month when he started a new job. However, due to limited data availability, we can only imprecisely estimate a worker's exact school graduation date, and thus the unemployment spell immediately after school graduation is subject to substantial measurement error. Therefore, in the empirical analysis that follows, we only utilize the second type of unemployment spell, that is, the unemployment spell between two employment spells. As we will explain in the next section, we can obtain consistent estimates of the structural parameters of the model even if we restrict the unemployment spell to that of the second type.

Given the definition of employment and unemployment spells presented above, we construct the labor market histories for the sampled individuals. In what follows, we refer to a worker's employment and unemployment spell sequences as his labor market profile, or simply his profile. In general, a profile is a combination of unemployment and employment spells, although two consecutive employment spells with different employers are also considered as a profile, since job-to-job transitions are allowed in our model. It should be noted that, because of the retrospective nature of the survey, we cannot observe the labor market history prior to the last employment spell. Therefore, in the empirical work conducted below, we assume that the young males sampled have no working experience prior to the reported spells. In other words, it is assumed that job turnover, if occurring, is at most one for the sampled individuals[10]. Given this assumption, the maximum number of spells for a profile is three, which occurs when an worker experiences two employment spells with one unemployment spell in the interim.

All possible labor market profiles are presented below with notations, R, C and U, which represent regular employment spells, contingent employment spells, and unemployment spells, respectively:

- Two-employment profiles: (RUR, RUC, CUC, CUR, RR, RC, CC, CR)
- One-employment profiles: (RU, CU, R, C)
- Unemployment profile: (U)

The distribution of each profile is presented in Table 1.

[10]This assumption does not seem improbable for the young male workers that we study here, as they had recently graduated from schools and started their job careers.

Using the information from the 2002 ESS, we can estimate the wages of the currently employed individuals, which can be imputed by the annual labor income divided by the total working hours in a year. However, we cannot observe any information on the wages of the previous jobs and thus the individuals' complete wage profiles are not available.

In the following estimation, we consider two schooling subgroups: college graduates and high school graduates. Their fractions of the total samples are 41.76% and 40.78% respectively. Other school subgroups, such as junior high school graduates and junior college graduates, are not considered because of their relatively small fraction of the samples. The summary statistics for the pooled and the selected schooling subgroups are presented in Table 2.

4 Maximum Likelihood Estimation

In this section, we present a maximum likelihood estimation method that consistently estimates the structural parameters of the behavioral model. We specify the likelihood function for each of the labor market profiles defined above. We then show that some of the structural parameters can be estimated consistently by using a part of the total likelihood, and the remaining parameters are estimated by maximizing the total likelihood function conditional to the partial maximizing likelihood estimates.

4.1 Likelihood Function

In order to specify the likelihood function, we first present the hazard rates from employment and unemployment spells respectively. According to the behavioral model described in section 2, an unemployed person leaves unemployment if he is offered a job and the offered wage is above the reservation wage. Given that two types of jobs, regular and contingent, can be offered, the hazard rate that an individual leaves the unemployment state, denoted by h_u, is provided by

$$h_u = \lambda_u^r \tilde{F}^r(w^{r*}) + \lambda_u^c \tilde{F}^c(w^{c*})$$

where the functions $\tilde{F}_r$ and $\tilde{F}_c$ are the survival functions defined by $\tilde{F}_r = 1 - F_r$ and $\tilde{F}_c = 1 - F_c$ respectively. An employed person may leave the job to move to another job or because he is terminated. Thus, the hazard rate that a regular worker who is paid a wage w leaves his job, denoted by $D_r(w)$, is given by the following equation:

$$D_r(w) = \eta_r + \lambda_r^r \tilde{F}^r(w) + \lambda_r^c \tilde{F}^c(w^{c*}(w)).$$

Similarly, the hazard rate that a contingent worker employed at wage w leaves the job, denoted by $D_c(w)$, is given by

$$D_c(w) = \eta_c + \lambda_c^r \tilde{F}^r(w^{r*}(w)) + \lambda_c^c \tilde{F}^c(w).$$

In the empirical analysis that follows, we assume that wage offers are i.i.d. draws from a truncated log normal distribution. The log normal density with parameter μ_w and σ_w is given by

$$g(w|\mu_w, \sigma_w) = \frac{1}{w\sigma_w\sqrt{2\pi}} \exp\left\{ -\frac{(\ln w - \mu_w)^2}{2\sigma_w^2} \right\}.$$

Following Flinn (2002), we assume that the lower bound of the wage offer distribution is equal to the reservation wage in the data. Thus the wage distribution is truncated from below by the reservation wage. Therefore, the density function of the wage for regular employment is given by

$$f^r(w) = g(w|\mu_w^r, \sigma_w^r)/\tilde{G}(w^{r*}|\mu_w^r, \sigma_w^r),$$

where $\tilde{G}$ is the survivor function of the log-normal distribution. The density function of the wage of the contingent employment, $f^c(w)$, is similarly defined.

In defining the likelihood function, we use the following notation: first, t^u denotes the duration of an unemployment spell, and t_i^r and t_i^c denote the time duration of the i-th regular and contingent employment spells respectively for $i = 1, 2$. It is sufficient to consider two employment spells because we assume that the number of employment spells in a labor market profile is at most two. Second, w_i denotes the wage of the i-th employment spell for $i = 1, 2$. Third, we introduce the following dummy variables such that χ_u is one if there is an unemployment spell; χ_{1r} is one if there is a first regular job spell; χ_{1c} is one if there is a first contingent job spell; χ_{2r} is one if there is a second regular job spell; and χ_{2c} is one if there is a second contingent job spell.

Using the notation presented above, the total likelihood function conditional on the first and second wages, w_1 and w_2, is given by

$$\begin{aligned}
\ell(w_1, w_2) \quad = \quad & \left[\exp\left[-D_r(w_1)t_1^r\right] \frac{f^r(w_1)}{\tilde{F}^r(w^{r*})}\right]^{\chi_{1r}} \left[\exp\left[-D_c(w_1)t_1^c\right] \frac{f^c(w_1)}{\tilde{F}^c(w^{c*})}\right]^{\chi_{1c}} \times \\
& \left[(\eta_r)^{\chi_{1r}}(\eta_c)^{\chi_{1c}} \exp\left[-h_u t_u\right] (\lambda_u^r)^{\chi_{2r}} (\lambda_u^c)^{\chi_{2c}}\right]^{\chi_u} \times \\
& \left[(\lambda_r^r)^{\chi_{1r}(1-\chi_u)} (\lambda_c^r)^{\chi_{1c}(1-\chi_u)} \exp\left[-D_r(w_2)t_2^r\right] f^r(w_2)\right]^{\chi_{2r}} \times \\
& \left[(\lambda_r^c)^{\chi_{1r}(1-\chi_u)} (\lambda_c^c)^{\chi_{1c}(1-\chi_u)} \exp\left[-D_c(w_2)t_2^c\right] f^c(w_2)\right]^{\chi_{2c}}.
\end{aligned}$$

Following the previous literature (e.g., Flinn 2002), we assume that wages in the data are observed with measurement errors. Thus the observed wage of the ith employment spell, denoted by w_i^o, is specified in relation to the hidden true wage w_i and an idiosyncratic error term ϵ such that

$$w_i^o = w_i \epsilon.$$

The error ϵ is assumed to follow a log-normal distribution with density

$$m(\epsilon) = \phi\left((\ln(\epsilon) - \mu_\epsilon)/\sigma_\epsilon\right)/\{\sigma_\epsilon \epsilon\}.$$

Given this error specification, the density function of the observed wage w_i^o is given by $m(w_i^o/w_i)/w_i$. Under the assumption that $\mathrm{E}(w^o|w) = w$, the parameters μ_ϵ and σ_ϵ are restricted so that $\sigma_\epsilon = \sqrt{-2\mu_\epsilon}$. Thus, in the following estimation, only μ_ϵ is considered as a free parameter, while σ_ϵ is determined through the previous restriction.

Since we consider measurement errors in the observed wages, the total likelihood function needs to be integrated over the unobserved wages, w_1 and w_2, and is given by the following equation:

$$\begin{aligned} &\int_{\underline{w}_1} \int_{\underline{w}_2} \ell(w_1, w_2) \times \\ &[m(w_1^o/w_1)/w_1]^{(1-\chi_u)(1-\chi_{2r}-\chi_{2c})} [m(w_2^o/w_2)/w_2]^{(\chi_{2r}+\chi_{2c})} \, dw_2 dw_1 \end{aligned} \tag{16}$$

where the integral bounds $\underline{w}_1$ and $\underline{w}_2$ are defined as

$$\begin{aligned} \underline{w}_1 &\equiv (w^{r*})^{\chi_{1r}} (w^{c*})^{\chi_{1c}} \\ \underline{w}_2 &\equiv \left[(w^{r*})^{\chi_{2r}} (w^{c*})^{\chi_{2c}}\right]^{\chi_u} \left[(w_1)^{(\chi_{1r}\chi_{2r}+\chi_{1c}\chi_{2c})} (w^{c*}(w_1))^{\chi_{1r}\chi_{2c}} (w^{r*}(w_1))^{\chi_{1c}\chi_{2r}}\right]^{1-\chi_u}. \end{aligned}$$

The maximum likelihood estimator maximizes the integrated likelihood function (16). The individual likelihood functions for all labor market profiles, the thirteen cases presented above, are shown in appendix A.

The integral lower bounds $\underline{w}_1$ and $\underline{w}_2$ are determined as follows. First, $\underline{w}_1$ is the reservation wage of the first employment spell, so that it is either w^{r*} or w^{c*}. Second, $\underline{w}_2$ is the minimum threshold wage that an individual uses as criterion to accept a second job offer. There are three possible values for $\underline{w}_2$: first, if the individual is unemployed in the previous spell ($\chi^u = 1$), $\underline{w}_2$ is given by the reservation wage; second, if the individual worked in the same type of job for the first and second employment spells ($\chi_{1r}\chi_{2r} + \chi_{1c}\chi_{2c} = 1$), $\underline{w}_2$ is given by the previous job's wage, and thus $\underline{w}_2 = w_1$; and third, if the individual worked at different types of employment in the first and second spells ($\chi_{1r}\chi_{2c} = 1$ or $\chi_{1c}\chi_{2r} = 1$), $\underline{w}_2$ is given by the reservation wage functions that are defined by equations (9)-(10). For example, if the individual changes his job from contingent to regular, $\underline{w}_2 = w^{r*}(w_1)$.

It should be noted that only the observed wage of the current employment, either w_1^o or w_2^o, appears in the total likelihood function (16). The reason is as follows: As we explained in the data section, the data structure of the 2002 ESS does not allows us to observe the entire wage profile of each respondent. Instead, only the wage of the current employment at the time of the interview is observed. For example, for one-employment profiles, such as RU, the wage of the first regular employment (i.e., previous employment) is not reported. Thus, the density of w_1^o does not appear in the likelihood function ($\chi_u = 1$, $\chi_{1r} = 0$ and $\chi_{2r} = 0$), and the likelihood needs to be integrated over the unobserved wage w_1. On the other hand, for two-employment profiles, such as RUR, the wage of the second regular employment (i.e., current employment) is observed while the first regular employment (i.e., previous employment) is not observed. Therefore, the density of w_2^o is included in the likelihood ($\chi_u = 0$, $\chi_{1r} = 1$ and $\chi_{2r} = 1$), and the likelihood is integrated over the true wages w_1 and w_2, which are unobserved in the data.

4.2 Estimation Procedure

Due to the high computational burden associated with the direct estimation of all the parameters by maximizing the total likelihood function (16), we proceed by using an indirect approach which is computationally more efficient and is implemented through a two-step estimation method, explained below.

The structural parameters of the behavioral model are divided into three groups: $\theta_1 = (\lambda_u^c, \lambda_u^r, F^c, F^r, w^{c*}, w^{r*}, \mu_\epsilon)$, $\theta_2 = (\lambda_c^r, \lambda_c^c, \lambda_r^r, \lambda_r^c, \eta_c, \eta_r)$ and $\theta_3 = (b_r, b_c)$. In the first step, we estimate the first group of parameters, θ_1, by maximizing a part of the total maximum likelihood function. In the second step, conditional on the estimates $\hat{\theta}_1$ from the first step, the remaining parameters, θ_2 and θ_3, are estimated using the total likelihood function. Following previous literature (e.g. Flinn, 2002), we assume that the discount rate is fixed, and $\rho = 0.05$. As discussed below, the unemployment benefit b_u cannot be identified from data.

Step 1: In the first estimation step, we estimate a part of the structural parameters, $\theta_1 = (\lambda_u^c, \lambda_u^r, F^c, F^r, w^{c*}, w^{r*}, \mu_\epsilon)$, by maximizing the following likelihood function:

$$\ell_1(\lambda_u^c, \lambda_u^r, F^c, F^r, w^{c*}, w^{r*}, \mu_\epsilon) = \int_{\underline{w}^*} \exp[-h_u \cdot t_u] \left\{ [\lambda_u^r f^r(w_2)]^{\chi_{2r}} \, [\lambda_u^c f^c(w_2)]^{\chi_{2c}} \, m(w_2^o/w_2)/w_2 \right\} dw_2 \tag{17}$$

where $\underline{w}^* = [w^{r*}]^{\chi_{2r}} [w^{c*}]^{\chi_{2c}}$.

It should be noted that the likelihood function (17) is a part of the total likelihood function (16). Due to the fact that the offer arrival rates during an unemployment spell, λ_u^c, λ_u^r, are related to the total likelihood (16) only through this part of the likelihood function, they are consistently estimated by maximizing the likelihood function (17). As for the wage parameters $(F^c, F^r, w^{c*}, w^{r*}, \mu_\epsilon)$, given the assumption made that no systematic selection on unobserved heterogeneity occurred for the sampled individuals, they are also consistently estimated using any of the employment spells. Therefore, all the parameters in θ_1 are consistently estimated by maximizing the function (17) [11].

It is apparent in the likelihood function (17) that the reservation wages, w^{r*} and w^{c*}, are treated as explicit parameters to be estimated by the data. The maximum likelihood estimators of the reservation wages are given by the minimum accepted wages observed in the samples of job takers. It can be shown[12] that these estimators are strongly consistent, although they are not asymptotically normally distributed [13]. Given their non-normal asymptotic distributions, the asymptotic standard errors are not reported in the estimation results presented below.

Step 2: Given the partial maximum likelihood estimate $\hat{\theta}_1$ from the first estimation step, the remaining structural parameters $\theta_2 = (\lambda_c^r, \lambda_c^c, \lambda_r^r, \lambda_r^c, \eta_c, \eta_r)$ and $\theta_3 = (b_r, b_c)$ are estimated by maximizing the total likelihood function (16).

We need to solve for the value functions, $V^r(w)$ and $V^c(w)$ in the process of the maximum likelihood estimation. To do so, we solve the system of integral equations (14) and (15) for each value of parameter. For notational simplicity, we denote the right hand side values of (14) and (15) as S_1 and S_2, respectively. That is, [14]

$$S_1\left(V^r(w), V^c(w); w^{c*}(w)\right) = \frac{b_r + w + \eta_r\left[\frac{b_r + w^{r*} + \lambda_r^r A(w^{r*}) + \lambda_r^c B(w^{c*})}{\rho + \lambda_r^r + \lambda_r^c}\right] + \lambda_r^r A(w) + \lambda_r^c B(w^{c*}(w))}{\rho + \eta_r + \lambda_r^r \tilde{F}^r(w) + \lambda_r^c \tilde{F}^c(w^{c*}(w))},$$

[11] Note that the estimation based on the partial likelihood may not be efficient due to the sample reduction. The efficiency loss, however, is compensated by the reduction in the computational burden of the maximum likelihood estimation. The system of value functions, $V^c(w)$ and $V^r(w)$, needs to be solved for each iteration of the parameter values in the maximum likelihood estimation. Hence, the computational time is not a trivial issue in this estimation. The estimation based on the partial likelihood function (17) reduces the number of parameters to be searched, and thus reduces computational time substantially.

[12] See Flinn and Heckman (1985) for a detailed discussion.

[13] The estimators do not follow the usual $\sqrt{N}$-asymptotic normal distributions. Flinn and Heckman (1985) show that the $\sqrt{N}$-standardized estimators, such as $\sqrt{N}(\hat{w}^{r*} - w^{r*})$ and $\sqrt{N}(\hat{w}^{c*} - w^{c*})$, have degenerate limiting distributions for many types of wage distributions, including log-normal distribution.

[14] Note that $F^r(w^{r*}) = 0$ and $F^c(w^{c*}) = 0$ in (14) and (15). This follows from the assumption that the wage distribution is given by the log-normal distribution that is truncated below by the reservation wage.

$$S_2\left(V^r(w), V^c(w); w^{r*}(w)\right) =$$
$$\frac{b_c + w + \eta_c\left[\frac{b_c+w^{c*}+\lambda_c^r A(w^{r*})+\lambda_c^c B(w^{c*})}{\rho+\lambda_c^r+\lambda_c^c}\right] + \lambda_c^r A(w^{r*}(w)) + \lambda_c^c B(w)}{\rho + \eta_c + \lambda_c^r \tilde{F}^r(w^{r*}(w)) + \lambda_c^c \tilde{F}^c(w)}.$$

Then the value functions are the solutions of the following system equations:

$$\begin{cases} V^c(w^{c*}(w)) = V^r(w) \\ V^r(w^{r*}(w)) = V^c(w) \\ V^r(w) = S_1\left(V^r(w), V^c(w); w^{c*}(w)\right) \\ V^c(w) = S_2\left(V^r(w), V^c(w); w^{r*}(w)\right). \end{cases}$$

This system of integral equations does not have an analytical solution. Therefore, we use a recursive numerical method to solve the system of equations. The detailed procedure to solve for the value functions is presented in appendix B.

In the previous system of integral equations, the value functions, $V^r(w)$ and $V^c(w)$ do not depend on the unemployment benefit b_u. Therefore, the reservation wage functions, $w^{r*}(w)$ and $w^{c*}(w)$, which are the inverse of the value functions, are not functions of b_u, and neither is the total likelihood (16). This implies that b_u cannot be identified from data, since it does not affect the likelihood value. Thus, we normalize $b_u = 0$ in the estimation that follows[15].

Following Flinn (2002), we determine the benefit parameters $\theta_3 = (b_r, b_c)$ endogenously by using the previous system of equations. To solve for θ_3, we utilize the fact that the value functions $V^r(w)$ and $V^c(w)$ satisfy the reservation wage conditions given by (7) and (8). Hence, for given values of the structural parameters (θ_1, θ_2) and assumed values for ρ and b_u, we find point estimates for $\theta_3 = (b_r, b_c)$ using the following equations:

$$V^r(w^{r*}) = V^u \iff$$
$$\frac{b_r + w^{r*} + \lambda_r^r A(w^{r*}) + \lambda_r^c B(w^{c*})}{\rho + \lambda_r^r + \lambda_r^c} = \frac{b_u + \lambda_u^r A(w^{r*}) + \lambda_u^c B(w^{c*})}{\rho + \lambda_u^r + \lambda_u^c}.$$

$$V^c(w^{c*}) = V^u \iff$$
$$\frac{b_c + w^{c*} + \lambda_c^r A(w^{r*}) + \lambda_c^c B(w^{c*})}{\rho + \lambda_c^r + \lambda_c^c} = \frac{b_u + \lambda_u^r A(w^{r*}) + \lambda_u^c B(w^{c*})}{\rho + \lambda_u^r + \lambda_u^c}.$$

Although these equations are non-linear on b_r and b_c, the solutions can be found conditional on θ_1, θ_2,

[15] Setting b_u to a higher value does not affect the estimated parameters, except for the values of b_r and b_c, which increase almost proportionally with b_u.

ρ and b_u. Therefore, the total likelihood function (16) is maximized only with respect to θ_2, not with respect to θ_3. Once the value of θ_2 is determined by the maximum likelihood process, we solve for θ_3 using the equations presented above, given the values of θ_1, θ_2, ρ and b_u.

5 Results

In this section, we present three sets of results. The first two help us answer the main question of this paper, that is, *"are contingent jobs stepping stones toward better future jobs or dead-ends for young workers in Japan?"*; the third question addresses issues of welfare and inequality. We first analyze the estimated parameters obtained using the method explained in the previous section. In particular, we are interested in the probabilities of the job-to-job transitions. We then study the probability that a given worker would find himself in a regular job after a certain number of periods as a function of the initial job taken. Finally, we examine the implications on welfare and inequality depending on the initial employment state of the worker.

5.1 Parameter Estimates

Table 3 presents the estimates of $\hat{\theta}_1$ based on the partial maximum likelihood estimation. The first column presents the estimates for the aggregated samples, and the second and third columns show the estimates for the disaggregated samples by schooling level.

The wage distributions are summarized by the estimates of $(\hat{\mu}_k, \hat{\sigma}_k, \hat{w}^{k*})$ for $k = r, c$. For the aggregated samples, the mean wage for regular jobs is $\mathrm{E}(w^r) = 1502.3$ yen, and the mean wage for contingent jobs is $\mathrm{E}(w^c) = 1400.6$ yen [16]. Thus, we find that in general, regular workers earn about 7 percent higher wages on average than continent workers. Furthermore, looking at the disaggregated samples by education, we can see that the mean wages of both regular and contingent jobs increase with educational level. Specifically, the college diploma is shown to increase the regular job wage from 1369.0 yen to 1771.3 yen, and it increases the contingent job wage from 1308.6 yen to 1801.3 yen. Interestingly, we find that for college graduates, the regular job wage is higher than the contingent job wage. The estimates also show that the wage distribution is more dispersed for contingent workers ($\mathrm{Var}(w^c) = 4.8315 \times 10^5$) than for regular workers ($\mathrm{Var}(w^r) = 2.9100 \times 10^5$), even though the mean value of the former is smaller than that of the latter. We can also see that the wage variation between regular and contingent workers is more significant for college graduates ($\mathrm{Var}(w^c) - \mathrm{Var}(w^r) =$

[16] Given that the wage distribution is truncated normal, the mean of type k job is computed by

$$\mathrm{E}(w^k) = \int_{w^{k*}} wg(w|\hat{\mu}^k, \hat{\sigma}^k)/\tilde{G}(w^{k*}|\hat{\mu}^k, \hat{\sigma}^k)dw.$$

2.6619×10^5) than for high school graduates ($\mathrm{Var}(w^c) - \mathrm{Var}(w^r) = 1.8137 \times 10^5$).

Turning our attention to the job offers to unemployed persons, we can observe that the estimated probability of leaving unemployment is $\hat{\lambda}_u^r + \hat{\lambda}_u^c = 0.0537$, which implies that unemployment spells last $1/0.0537 \simeq 18.6$ months on average. The estimates of $(\hat{\lambda}_u^r, \hat{\lambda}_u^c)$ for the aggregated samples imply that unemployed persons are offered regular jobs more often than contingent jobs. However, after controlling for education heterogeneity by disaggregating the sample, a different pattern emerges. For the disaggregated sample estimates it is found that college graduates experienced about a half-year longer unemployment duration ($1/0.044 \simeq 22.8$ months) than high school graduates ($1/0.062 \simeq 16.1$ months). We also find that high school graduates were offered regular and contingent jobs at about the same frequency, while college graduates were offered regular jobs almost three times more often than contingent jobs.

In terms of the estimated reservation wages, we find that they increase with schooling level for both regular and contingent jobs. This is consistent with the estimated wage distributions in which college graduates have a higher mean wage than high school graduates. The estimates also show that high school graduates have lower reservation wages to accept a regular job than to accept a contingent job. These estimates seem to suggest that high school graduates are willing to take a regular job even if its wage is lower than the minimum acceptable wage for a contingent job. This could be for two reasons. First, as we explain below in more detail, the fringe benefits for regular jobs are higher than for contingent jobs. Second, the future present value of a regular job is higher, since job destruction rates are lower. On the other hand, college graduates have a higher reservation wage for regular jobs than for contingent jobs, which may reflect the fact that college graduates have relatively high job offer rates when unemployed, and, as shown below, they have very low job offer rates when employed. Therefore, college graduates are often better off remaining unemployed and waiting for a higher wage regular job offer than accepting a regular job with a low wage, which could imply staying at that low wage job for a long time.

Table 4 reports the maximum likelihood estimates of $\hat{\theta}_2$ and $\hat{\theta}_3$ for the aggregated samples in the first column, and for the disaggregated estimates by schooling in the second and third columns, respectively.

The estimates of the job separation rates show that, for both the aggregated and disaggregated samples, contingent workers are about five times more likely to lose their jobs than regular workers. We also find that high school graduates are more likely to lose their jobs ($\hat{\eta}_r + \hat{\eta}_c = 0.0313$) than college graduates ($\hat{\eta}_r + \hat{\eta}_c = 0.0185$). The estimated mean job durations for high school graduates are 3.2 years for contingent jobs and 16.1 years for regular jobs. The mean durations for college graduates are 5.4 years for contingent jobs and 28.3 years for regular jobs.

We can also see in Table 4 that the job-offer rates to employed individuals are generally low compared with the magnitudes of the job-offer rates to unemployed persons. According to the estimates for the aggregated samples, the monthly rate of receiving at least one job-to-job offer is $\hat{\lambda}_r^r + \hat{\lambda}_r^c + \hat{\lambda}_c^c + \hat{\lambda}_c^r = 0.0143$, which is about twenty-five percent of the offer rates to unemployed workers. Interestingly, we find that the probability of receiving an offer to move to a new job is slightly higher for regular job holders ($\hat{\lambda}_c^c + \hat{\lambda}_c^r = 0.0074$) than for contingent job holders ($\hat{\lambda}_c^c + \hat{\lambda}_c^r = 0.0069$). When looking at the job offer rates for employed persons by educational level, we find that for college graduates, the highest arrival rate is from regular to contingent employment, while for high school graduates the highest rate is that to move from regular to contingent employment. These two education level findings could indicate that, on the one hand, college graduates search for better jobs most intensively when holding a contingent job, and are more successful in obtaining offers. On the other, for high school graduates to hold a regular job sends a good signal which attracts contingent job offers. However, it should be stressed that these direct job-to-job transitions are not the only channels for workers to move from one job to another. There is another indirect channel through unemployment, since the offer rate for unemployed workers is higher than for employed people. The entire transition pattern from contingent jobs to regular jobs, allowing for these direct and indirect paths, is investigated through a simulation exercise, which is shown in the next subsection.

Finally, as for the employment benefits, the estimates shown in Table 4 seem to be large in magnitude compared to the estimated average hourly wage. This is not implausible, however. Notice that these fringe benefits comprise every type of remuneration for the worker which is not directly reflected in the wages, such as medical benefits, pension payments, or on-the-job training. Moreover, since the model assumes linear utility, it also encompasses the utility the worker gets out of being employed, which could be substantial if the social stigma associated with being unemployed if high. We can also see that the employment benefit for the regular job is larger than the employment benefit for the contingent job, independent of schooling level. Furthermore, employment benefits increase with schooling level, and the difference in benefits between the regular and contingent jobs is larger for college graduates than for high school graduates.

5.2 Stepping Stones vs. Dead Ends

The main question of this paper is whether contingent jobs are dead ends or stepping stones toward regular jobs. Before trying to answer this question, we state the criteria used to assess the results in terms of this question. We consider that contingent jobs are stepping stones towards regular jobs if the probability of finding a regular job increases by taking a contingent job relative to staying unemployed[17]. And, we consider that contingent jobs are dead ends if for the whole length of a person's career, that is, for around 40 years, a worker who starts at a contingent job is less likely to find himself

[17]This criterion is in the spirit of Autor and Houseman (2008), which considers whether a worker is better off taking a temporary help job than not getting a job at all in terms of eventually finding an permanent job.

in a regular job than do workers who are initially unemployed or employed at regular jobs.

To answer our main question, the first place to look is at the transition probabilities estimated in the previous subsection. Tables 3 and 4 show that the probability for a worker to move from a contingent job to a regular job ($\hat{\lambda}_c^r$) is substantially lower than the probability of moving to such a of job from unemployment ($\hat{\lambda}_u^r$). This occurs for the aggregated data and also when splitting the sample by educational level. Hence, this result seems to indicate that contingent jobs do not help the chances of a worker to find a regular job, and therefore are not stepping stones. However, in order to fully understand if contingent jobs can help workers to move to regular jobs, we need to consider not only the direct path from contingent to regular jobs, but also all intermediate paths between these jobs. For that purpose, we use the estimated parameters to simulate all possible employment and unemployment paths for workers, and analyze the chances of being employed at a regular job depending on the initial employment status and at different time horizons. This approach provides us with more information than just the instantaneous probability to move from one state to another, and informs us of whether the effect of an initial contingent job is permanent, or disappears with time, and if it disappears, how many years it takes to do so.

To perform the analysis, we use the parameter estimates and simulate the job histories of one hundred thousand workers starting from each possible employment status: regular job, contingent job, and unemployed. From the simulated job histories, we calculate the probability that starting from each of the three employment states, the worker's employment status is regular, contingent, or unemployed after a certain number of periods. Given the large size of the simulated samples, these probabilities are given by the fraction of workers in each employment state at the end of the studied time frame. The details of the simulation are described in appendix C.

Figure 1 contains the job transition probabilities to regular employment from each initial employment status at different time horizons for the aggregated sample [18]. We can see in the figure that the probability for a worker to be employed at a regular job is higher when the individual is initially unemployed than when he is at a contingent job. This happens at the initial stage, which is consistent with what we have explained above about the instantaneous transitions probabilities, but also at 5 and 10 year horizons. That is, initially taking a contingent job has an effect on the probability of an individual to find regular employment. However, this effect is not permanent, since after 20 years the difference in the probability of being employed at a regular job is almost the same for those initially unemployed and employed at a contingent job. Given that a worker who enters the labor market at 20 and retires at 60 has a career of 40 years, this means that halfway through the career of a worker, the effect of having initially taken a contingent job disappears. Therefore, we find that contingent employment is

[18] The results for the disaggregated data are qualitatively similar. The detail simulation results are available from the authors upon request.

not a dead-end.

In summary, contingent jobs seem to be neither dead-ends nor stepping stones to regular jobs. They are not dead-ends because contingent workers share the same probability of having a regular job in a 15 to 20 year horizon with initially regular workers and unemployed workers. However, they are not stepping stones because for unemployed workers, taking a contingent job does not raise the probability of having a regular job either in the short term orin the long-run.

5.3 Employment Stability, Welfare and Inequality

We have seen in the previous subsection that if the goal of an individual is to have a regular job, in the short-run he is better off not taking a contingent job, but in the medium-run that choice has very little influence. In this subsection we explore other implications of the initial employment status for the worker. In particular, we are interested in analyzing the consequences for different time horizons andinitial states on (i) employment stability, understood as the number of different jobs and unemployment spells experienced by a worker; (ii) welfare, as the present discounted value of utility for a worker; and (iii) inequality across individuals. To perform this analysis we work again with the simulated employment paths generated with the estimated coefficients of the model.

Employment stability can be a major concern for workers in the economy. Although absent in our model, the stigma associated with constant job changes, the loss of skills while unemployed, or the desire to smooth consumption over time, are reasons why workers may prefer an employment career with fewer employment status changes. The results of our model for employment stability can be seen in Figure 2. Figure 2 (a) shows the number of turnovers, or equivalently, the number of different employment and unemployment spells, for a worker who is initially unemployed or employed at a regular or a contingent job. We see that the average number of turnovers increases with time for the three types of initial state, but the growth rate is higher for initially unemployed and contingent employees than for regular employees. This lower turnover and therefore higher employment stability for regular workers is a reflection of the lower destruction rates faced at regular jobs. After 8 to 10 years, the number of turnovers stabilizes and increases at the same rate for all initial states, which implies that from that point onwards, employment stability is the same, independent of the initial state. Figure 2 (b) shows that, averaging across employment states, higher education is associated with higher employment stability at any stage of a worker's career, since the level and growth rate of turnovers are always lower for college graduates. That is, although the effect of the initial employment state on employment stability disappears after a few years, the effect of the education level is permanent.

We now turn to study the welfare implications of the initial employment status of the worker. Many of the papers which have studied the effects of part-time/temporary employment on the long-term employment prospects of workers focus on transition probabilities or on the length of time until a regular job is obtained. However, individuals compare the present value of the different options when making their decisions about whether or not to take a certain job. Hence, we think that it is important to study, besides the transition probabilities, the effects of taking a contingent job or remaining unemployed on the welfare of the worker. Panel A of Table 5 shows the mean values at 5, 10, and 40 year horizons of the present value at period zero of the welfare of workers who start in any of the three possible states. In other words, given that a worker starts his career at a given state, we simulate the model and calculate the present value of the utility obtained up to a certain number of years. We see that the workers with the highest mean welfare are those who start with a regular job, followed by those with an initial contingent job, and finally by the initially unemployed. This occurs for the three different time horizons, although the differences decrease as time passes, the lowest being at 40 years, which is the length of a worker's career. We also observe that this result holds when we disaggregate the data by education level. What this result indicates is that despite the fact that being unemployed is better for the worker to find a regular job in the short-run, the forgone earnings and therefore forgone utility makes the option of waiting a costly one in terms of welfare.

Finally we want to analyze the implications of our model and estimated parameters in terms of the inequality in the welfare of workers as a function of their initial jobs. In Panels B and C of Table 5 we can see, through the values of the Coefficient of Variation (CV) squared and Gini coefficient, that inequality is highest among those workers who start unemployed, followed by those who are initially in contingent jobs. The least inequality is found for workers with regular jobs initially. These results show once more that although initially holding a contingent job lowers the chances to move to a regular job compared with being initially unemployed, it reduces the risk of misfortune in terms of income and welfare. We also observe that inequality decreases as the career of the worker progresses, whereby the ordering of highest to lowest inequality is as explained above. Finally, comparing the college and high school graduates, we see the same inequality patterns as for the aggregate data, but as should be expected, high school graduates display higher inequality than college graduates.

6 Discussion

The results from the previous sections indicate that initial employment status matters for welfare due to its effect on subsequent job offers, job separations, wages, and nonpecuniary employment benefits. In particular, there are substantial differences in later careers and total earnings, not only between employed and unemployed, but also between regular and contingent workers.

Taking the estimates of the model's parameters as given and the parameters themselves as fundamental, we can perform a number of policy experiments relevant to the contingent job arrangements.

We address the following issue: what would the probability of having regular employment and welfare for each initial employment status look like if first, the rate of regular job offer to contingent workers increases; second, the non-pecuniary employment benefit for contingent workers is improved; and third, contingent employment is more stabilized than the current level?

For the first experiment, we simulate the model and generate counter factual job histories, setting the rate of regular job offer for contingent workers at the same level as that of unemployed individuals (i.e., $\lambda_c^r = \lambda_u^r$). Promotion of the conversion of contingent employment contracts into regular ones may be one possible implementation of such a policy. For the second experiment, we compute the counter factual job histories, setting the nonpecuniary benefit for contingent workers at the same level of regular workers (i.e., $b_c = b_r$). Provision to contingent workers of health insurance, pension and other fringe benefits provided to regular workers would be a policy that improves the non-pecuniary benefits of contingent employment. For the third experiment, we simulate the model setting the exogenous job separation rate for contingent workers at the same level of regular workers (i.e., $\eta_c = \eta_r$). Firing restrictions which increase the termination cost to firms is one such possible policy.

Figure 3 plots the job transition probabilities to regular employment from each initial employment status at different time horizons under the benchmark case and the three experiments for the aggregated sample. The results of the first experiment concerning the increase in the job offer rate to contingent workers indicate that first, the steady-state probability of becoming a regular employee is higher by about 20 percent points compared to the benchmark case; second, there is no difference between initially contingent workers and the unemployed in the transition probability to a regular job; and third, the effect of initial employment status lasts for less than 10 years. The probability of regular employment in the new steady state is higher than the benchmark case because workers receive more job offers for regular employment, which has a low probability of separation. The difference in the transition probability to regular jobs between initially contingent workers and the unemployed disappears because, with small rates of regular and contingent job offers to contingent workers, both contingent workers and the unemployed face an almost identical probability of obtaining regular employment. The effect of initial employment status disappears much faster than in the benchmark case because workers' job mobility across the three employment states is higher with the high job offer rate to contingent workers than in the benchmark case. These results indicate that policies such as the promotion of the conversion of contingent employment contracts into regular ones may be effective in raising the long-run share of regular employment and mitigating the adverse effect of taking contingent jobs and being unemployed initially. Moreover, they also suggest that such a policy may be effective to reduce unemployment because unemployed job searchers may want to take contingent jobs, since in that case they do not sacrifice the possibility of obtaining regular employment in the future.

The second experiment indicates that the increase in contingent employment benefits has little effect on the transition probabilities. One reason that we obtain a small effect of this policy on the transition probability may be the fact that the difference in the estimated non-pecuniary benefits between regular and contingent jobs is so small that workers do not change their job switching behaviors with higher contingent employment benefits.

The results of the third experiment with a low job separation rate of contingent workers indicate that first, the steady-state probability to regular employment is lower by about 20 percentage points than the benchmark case; and second, the effect of initial employment status lasts for about 30 years. The long-run probability of obtaining regular employment is smaller than in the benchmark case because with a low separation rate for contingent jobs, more workers end up with stable contingent jobs in the long run. Moreover, with low job mobility, the job turnover rate across employment status is low. These results imply that in the benchmark case, many initially contingent workers reach the regular employment status through unemployment. Indeed, the increase of contingent job stability raises the probability of workers being trapped at the contingent jobs.

Let us turn to the effect of these policy changes on the welfare of the workers. Table 6 contains the percentage changes of the mean, CV squared, and Gini coefficient of the 5-year, 10-year, and 40-year welfare from each initial employment status for the aggregate sample [19]. For the first experiment, the mean welfare increases for each initial employment status. Initially contingent workers benefit the most from this policy change, as we would expect. This is because the initially contingent workers face a high probability of obtaining regular jobs with high wages and non-pecuniary benefits. Inequality measures fall for all initial employment states. Note that among the three initial states, it is for the initially contingent worker that we obtain the highest equalization in a 5 year horizon. In the long run, all three states reduce the inequality in a similar way. This is due to the high long-run share of regular employment with a stable employment environment. These results indicate that policies that improve the possibility of obtaining regular employment in the short run and in the long run may be, other things equal, effective to improve welfare of workers through a reduction in welfare inequality.

For the second experiment, the mean of welfare for initially contingent workers is improved, although the magnitude is small. Note that this policy also improves the 5-year welfare of the initially unemployed, but not of the initially regular workers. This is because most of the initially regular workers stay at the first job for the rest of their job careers. The overall small effects of this policy change are predictable given the small difference in the estimated non-pecuniary benefits between regular and contingent jobs.

Let us look now at the results of the third experiment. The increase in the job stability of contingent jobs raises the welfare of initially contingent workers substantially. Moreover, the reduction of

[19] We find that the simulation results for disaggregated case by education are qualitatively similar to the aggregated case presented above. These results are available upon request.

inequality is also substantial. These results stem from the simple fact that a contingent job is better than unemployment in terms of welfare, since contingent workers earn wages and enjoy non-pecuniary benefits. These results indicate that even though policies that increase the job stability of contingent workers increase the probability of being trapped at the contingent job, contingent workers benefit from a stable employment environment, and hence, such policies may be effective in improving welfare while reducing inequality.

Finally, we analyze the effects of these policies on the overall distribution of welfare. We report in Table 6 the percentage changes of the same variables for the aggregated case in which simulated job histories are mixed according to the actual initial employment status distribution. The overall impact of these policies on the welfare distribution is close to the effect seen for initially regular workers, which is due to the fact that those types of workers are the majority in our sample. However, the effects of these policies are still substantial. In particular, policies that increase the offer rate of regular jobs to contingent workers (Experiment 1) and those that improve the job stability of contingent workers (Experiment 3) are effective in improving the mean welfare while reducing inequality, both in the short and long terms. These results suggest that policies that improve the environment surrounding contingent employment may have a substantial impact on the overall welfare distribution for young male workers.

7 Concluding Remarks

This paper addresses the important issue of the short and long-term effects of contingent employment for young workers. Part-time, dispatch, and temporary employment has increased substantially in past years in many developed countries. The increase in these contingent jobs is important especially for young workers, since such jobsbring lower wages, fewer benefits, and sometimes great stigmas, which could affect the workers later in their careers. However, taking a contingent job may be better than being unemployed if it leads to a better job.

We study the effects of the initial employment status of young workers on their careers. To do so, we present an on-and-off-the-job search model, employing Japanese data on young workers to estimate the model parameters using Maximum Likelihood, and we estimate the probability that a worker be employed at a regular job after a certain number of years by simulating career paths. We also analyze the welfare implications of the initial job choice for these workers and the effects of government policies.

The findings of our study show that for young male Japanese workers, initial employment status only matters for the probability of being employed at a regular job in the first years of a worker's career, but not in the long run. In the short term, it is more probable for an initially unemployed worker to have a regular job than for an initially contingent worker, but after 20 years this probability is the same for both types of worker. Hence, we find that contingent jobs are neither stepping stones nor dead ends. However, when analyzing the welfare implications of initial contingent employment, we find that although these types of jobs may not be dead ends, they imply a lower long-term welfare for the individual. Finally, we show that policies aimed at increasing the transition of workers from contingent to regular jobs are the most effective in terms of increasing the fraction of long-run regular employment, increasing overall welfare, and reducing inequality.

To summarize, in Japan, starting a career in a contingent job has no long term effects on the chances of a worker finding a regular job, although it does have important long-run implications in terms of welfare.

References

[1] Autor, D. H. and Houseman, S. (2008). “Do Temporary Help Jobs Improve Labor Market Outcomes for Low-Skilled Workers? Evidence from ’Work First’ ”, MIT, January 2008, *mimeo.*

[2] Booth, A. L., Francesconi, M., and Frank, J. (2002). “Temporary Jobs: Stepping Stones or Dead Ends?” *The Economic Journal*, 112 (June), pp. 189-213.

[3] Bover, O. and Gomez, R. (2004). “Another Look at Unemployment Duration: Exit to a Permanent vs. a Temporary Job”, *Investigaciones Economicas*, 28 (2), pp. 285-314.

[4] Casquel, E. and Cunyat, A. (2005). “The Dynamics of Temporary jobs: Theory and some Evidence for Spain (The Role of Skill)” *Fundacion de Estudios de Economia Aplicada*, Working Paper 2005-19.

[5] Flinn, C. J. (2002). “Labour Market Structure and Inequality: A Comparisons the U.S.” *Review of Economic Studies*, 69, pp. 611-645.

[6] Flinn, C. J. and Heckman, J. (1982). “New Methods for Analyzing Structural Models of Labor Force Dynamics” *Journal of Econometrics*, 18, pp. 115-168.

[7] Genda, Y. and Kurosawa, M. (2001). “Transition from School to Work in Japan” *Journal of the Japanese and International Economies*, 15, pp. 465-488.

[8] Genda, Y. (2008). “On the Transition from Non-Regular to Regular Employment” *The Japanese Journal of Labour Studies* (in Japanese), No. 580, pp.61-77.

[9] Gaston, N. and Timcke, D. (1999). "Do Casual Workers Find Permanent Full-Time Employment? Evidence from the Australian Youth Survey", The Economic Record, 75 (December), pp. 333-347.

[10] Houseman, S. and Osawa, M. (2003). *Nonstandard Work in Developed Economies. Causes and Consequences.* Kalamazoo, Michigan. W.E. Upjohn Institute for Employment Research. Susan Houseman and Machiko Osawa, Editors.

[11] Judd, K.L. (1998). *Numerical Methods in Economics.* Cambridge: MIT Press.

[12] Kondo, A. (2007). "Does the first job really matter? State dependency in employment status in Japan" Journal of the Japanese International Economies, 21, pp. 379–402.

[13] Kvasnicka, M. (2008). "Does Temporary Help Work Provide a Stepping Stone to Regular Employment?" *National Bureau of Economic Research*, Working Paper No. 13843.

[14] Zijl, M., van den Berg, G. J. and Heyma, A. (2004) "Stepping Stones for the Unemployed: The Effect of Temporary Jobs on the Duration until Regular Work." *Institute for Labor* Studies, Discussion Paper No. 1241.

A Total Likelihood Function

In this part of the appendix, we specify the functional form of the total likelihood, which is given with the equation (16) as the general formula for each of the thirteen labor market profiles. In what follows, we ignore the constant terms that are not related to parameters for simplicity, so that the total likelihood is proportional to the one presented below.

Case 1 The likelihood function of the profile of *RUR* is given by

$$\begin{aligned}\ell &\propto \eta_r \times \exp\{-\eta_r(t_1^r + t_2^r)\} \times \int_{w^{r*}} \exp\{-\lambda_r^r \tilde{F}^r(w_1)t_1^r - \lambda_r^c \tilde{F}^c(w^{c*}(w_1))t_1^r\} f^r(w_1)dw_1 \times \\ &\quad \int_{w^{r*}} \exp\{-\lambda_r^r \tilde{F}^r(w_2)t_2^r - \lambda_r^c \tilde{F}^c(w^{c*}(w_2))t_2^r\} f^r(w_2)(m(w_2^o/w_2)/w_2)dw_2.\end{aligned}$$

Case 2 The likelihood function of the profile of *RUC* is given by

$$\begin{aligned}\ell &\propto \eta_r \times \exp\{-\eta_r t_1^r - \eta_c t_2^c\} \times \int_{w^{r*}} \exp\{-\lambda_r^r \tilde{F}^r(w_1)t_1^r - \lambda_r^c \tilde{F}^c(w^{c*}(w_1))t_1^r\} f^r(w_1)dw_1 \times \\ &\quad \int_{w^{c*}} \exp\{-\lambda_c^r \tilde{F}^r(w^{r*}(w_2))t_2^c - \lambda_c^c \tilde{F}^c(w_2)t_2^c\} f^c(w_2)(m(w_2^o/w_2)/w_2)dw_2.\end{aligned}$$

Case 3 The likelihood function of the profile of *CUC* is given by

$$\begin{aligned}\ell &\propto \eta_c \times \exp\{-\eta_c(t_1^c + t_2^c)\} \times \int_{w^{c*}} \exp\{-\lambda_c^r \tilde{F}^r(w^{r*}(w_1))t_1^c - \lambda_c^c \tilde{F}^c(w_1)t_1^c\} f^c(w_1)dw_1 \times \\ &\quad \int_{w^{c*}} \exp\{-\lambda_c^r \tilde{F}^r(w^{r*}(w_2)t_2^c - \lambda_c^c \tilde{F}^c(w_2)t_2^c\} f^c(w_2)(m(w_2^o/w_2)/w_2)dw_2.\end{aligned}$$

Case 4 The likelihood function of the profile of *CUR* is given by

$$\begin{aligned}\ell &\propto \eta_c \times \exp\{-\eta_c t_1^c - \eta_r t_2^r\} \times \int_{w^{c*}} \exp\{-\lambda_c^r \tilde{F}^r(w^{r*}(w_1))t_1^c - \lambda_c^c \tilde{F}^c(w_1)t_1^c\} f^c(w_1)dw_1 \times \\ &\quad \int_{w^{r*}} \exp\{-\lambda_r^r \tilde{F}^r(w_2)t_2^r - \lambda_r^c \tilde{F}^c(w^{c*}(w_2))t_2^r\} f^r(w_2)(m(w_2^o/w_2)/w_2)dw_2.\end{aligned}$$

Case 5 The likelihood function of the profile of *RR* is given by

$$\begin{aligned}\ell &\propto \lambda_r^r \times \exp\{-\eta_r(t_1^r + t_2^r)\} \times \int_{w^{r*}} \int_{w_1} \exp\{-\lambda_r^r \tilde{F}^r(w_1)t_1^r - \lambda_r^c \tilde{F}^c(w^{c*}(w_1))t_1^r\} f^r(w_1) \times \\ &\quad \exp\{-\lambda_r^r \tilde{F}^r(w_2)t_2^r - \lambda_r^c \tilde{F}^c(w^{c*}(w_2))t_2^r\} f^r(w_2)(m(w_2^o/w_2)/w_2)dw_2dw_1\end{aligned}$$

Case 6 The likelihood function of the profile of *RC* is given by

$$\begin{aligned}\ell &\propto \lambda_r^c \times \exp\{-\eta_r t_1^r - \eta_c t_2^c\} \times \int_{w^{r*}} \int_{w^{c*}(w_1)} \exp\{-\lambda_r^r \tilde{F}^r(w_1)t_1^r - \lambda_r^c \tilde{F}^c(w^{c*}(w_1))t_1^r\} f^r(w_1) \times \\ &\quad \exp\{-\lambda_c^r \tilde{F}^r(w^{r*}(w_2))t_2^c - \lambda_c^c \tilde{F}^c(w_2)t_2^c\} f^c(w_2)(m(w_2^o/w_2)/w_2)dw_2dw_1.\end{aligned}$$

Case 7 The likelihood function of the profile of *CC* is given by

$$\begin{aligned}\ell &\propto \lambda_c^c \times \exp\{-\eta_c(t_1^c + t_2^c)\} \times \int_{w^{c*}} \int_{w_1} \exp\{-\lambda_c^r \tilde{F}^r(w^{r*}(w_1))t_1^c - \lambda_c^c \tilde{F}^c(w_1)t_1^c\} f^c(w_1) \times \\ &\quad \exp\{-\lambda_c^r \tilde{F}^r(w^{r*}(w_2)t_2^c - \lambda_c^c \tilde{F}^c(w_2)t_2^c\} f^c(w_2)(m(w_2^o/w_2)/w_2)dw_2dw_1.\end{aligned}$$

Case 8 The likelihood function of the profile of *CR* is given by

$$\begin{aligned}\ell &\propto \lambda_c^r \times \exp\{-\eta_c t_1^c - \eta_r t_2^r\} \times \int_{w^{c*}} \int_{w^{r*}(w_1)} \exp\{-\lambda_c^r \tilde{F}^r(w^{r*}(w_1))t_1^c - \lambda_c^c \tilde{F}^c(w_1)t_1^c\} f^c(w_1) \times \\ &\quad \exp\{-\lambda_r^r \tilde{F}^r(w_2)t_2^r - \lambda_r^c \tilde{F}^c(w^{c*}(w_2))t_2^r\} f^r(w_2)(m(w_2^o/w_2)/w_2)dw_2dw_1.\end{aligned}$$

Case 9 The likelihood function of the profile of *RU* is given by

$$\ell \propto \eta_r \times \exp\{-\eta_r t_1^r\} \times \int_{w^{r*}} \exp\{-\lambda_r^r \tilde{F}^r(w_1) t_1^r - \lambda_r^c \tilde{F}^c(w^{c*}(w_1)) t_1^r\} f^r(w_1)(m(w_1^o/w_1)/w_1) dw_1.$$

Case 10 The likelihood function of the profile of *CU* is given by

$$\ell \propto \eta_c \times \exp\{-\eta_c t_1^c\} \times \int_{w^{c*}} \exp\{-\lambda_c^r \tilde{F}^r(w^{r*}(w_1)) t_1^c - \lambda_c^c \tilde{F}^c(w_1) t_1^c\} f^c(w_1)(m(w_1^o/w_1)/w_1) dw_1.$$

Case 11 The likelihood function of the profile of *R* is given by

$$\ell \propto \exp\{-\eta_r t_1^r\} \times \int_{w^{r*}} \exp\{-\lambda_r^r \tilde{F}^r(w_1) t_1^r - \lambda_r^c \tilde{F}^c(w^{c*}(w_1)) t_1^r\} f^r(w_1)(m(w_1^o/w_1)/w_1) dw_1.$$

Case 12 The likelihood function of the profile of *C* is given by

$$\ell \propto \exp\{-\eta_c t_1^c\} \times \int_{w^{c*}} \exp\{-\lambda_c^r \tilde{F}^r(w^{r*}(w_1)) t_1^c - \lambda_c^c \tilde{F}^c(w_1) t_1^c\} f^c(w_1)(m(w_1^o/w_1)/w_1) dw_1.$$

Case 13 The likelihood function of the profile *U* is given by $\text{texp}[-h_t t_u]$. However, it is constant conditional on the partial likelihood estimates from step 1. So, this profile does not contribute the likelihood in the estimation step 2.

B Numerical Solution of the Integral Equations System

Here we present the numerical method used to solve the system of integral equations for the value functions $V^r(w)$ and $V^c(w)$ used in the paper:

$$\begin{cases} V^c(w^{c*}(w)) = V^r(w) \\ V^r(w^{r*}(w)) = V^c(w) \\ V^r(w) = S_1(V^r(w), V^c(w); w^{c*}(w)) \\ V^c(w) = S_2(V^r(w), V^c(w); w^{r*}(w)). \end{cases}$$

To solve this system, we utilize a recursive procedure. Let $(w_{(k)}^{r*}(w), w_{(k)}^{c*}(w))$ be the reservation wage functions of the kth iteration. Let $V_{(k)}^r(w)$ and $V_{(k)}^c(w)$ be the value functions that are compatible with the reservation wage functions $(w^{r*}_{(k)}(w), w^{c*}_{(k)}(w))$. Then we solve for $V^r_{(k)}(w)$ and $V^c_{(k)}(w)$ using the following integral equations:

$$V^r_{(k)}(w) = S_1(V^r_{(k)}, V^c_{(k)}, w^{c*}_{(k)}(w))$$
$$V^c_{(k)}(w) = S_2(V^r_{(k)}, V^c_{(k)}, w^{r*}_{(k)}(w)).$$

Given the solved value functions, $V^r_{(k)}(w)$ and $V^c_{(k)}(w)$, we update the reservation wage functions by the following inversion formula:

$$\begin{aligned} w^{r*}_{(k+1)}(w) &= \left(V^r_{(k)}\right)^{-1}\left[V^c_{(k)}(w)\right], \\ w^{c*}_{(k+1)}(w) &= \left(V^c_{(k)}\right)^{-1}\left[V^r_{(k)}(w)\right]. \end{aligned}$$

Using $w^{r*}_{(k+1)}(w)$ and $w^{c*}_{(k+1)}(w)$ as the $(k+1)$th iteration reservation wage functions, we can repeat the procedure to solve for $V^r_{(k+1)}(w)$ and $V^c_{(k+1)}(w)$. This procedure is repeated until a certain convergence criterion is satisfied. We use $w^{r*}_{(0)}(w) = w^{c*}_{(0)}(w) = w$ as the initial values of the iteration procedure.

In the iteration scheme presented above, the key step is to solve for the value functions, $V^r_{(k)}(w)$ and $V^c_{(k)}(w)$, given the predetermined reservation wage functions, $w^{r*}_{(k)}(w)$ and $w^{c*}_{(k)}(w)$. For simplicity, the subscript (k) is dropped in what follows. Then the system of integral equations is simply given by

$$V^r(w) = S_1(V^r(w), V^c(w); w^{c*}(w)) \tag{18}$$
$$V^c(w) = S_2(V^r(w), V^c(w); w^{r*}(w)). \tag{19}$$

Note that the reservation wage functions $w^{r*}(w)$ and $w^{c*}(w)$ are predetermined. We solve this problem numerically using a quadrature method, as shown below.

Following the standard classification scheme for integral equations, the equations (18) and (19) are considered as, so called, Volterra Type 2. Rewrite equation (18) as

$$\begin{aligned} &V^r(w) = D_{(1)}(w) \\ &+ \int_{w^{r*}} G_{(11)}(w)V^r(x)dF^r(x) + \int_w G_{(12)}(w)V^r(x)dF^r(x) \\ &+ \int_{w^{c*}} H_{(11)}(w)V^c(x)dF^c(x) + \int_{w^{c*}(w)} H_{(12)}(w)V^c(x)dF^c(x), \end{aligned} \tag{20}$$

where $D_{(1)}$, $G_{(11)}$, $G_{(12)}$, $H_{(11)}$ and $H_{(12)}$ are defined by

$$\begin{aligned}
D_{(1)}(w) &= \frac{(\rho+\lambda_r^r+\lambda_r^c)(b_r+w)+\eta_r(b_r+w^{r*})}{(\rho+\eta_r+\lambda_r^r\tilde{F}^r(w)+\lambda_r^c\tilde{F}^c(w^{c*}(w)))(\rho+\lambda_r^r+\lambda_r^c)} \\
G_{(11)}(w) &= \frac{\eta_r\lambda_r^r}{(\rho+\eta_r+\lambda_r^r\tilde{F}^r(w)+\lambda_r^c\tilde{F}^c(w^{c*}(w)))(\rho+\lambda_r^r+\lambda_r^c)} \\
G_{(12)}(w) &= \frac{\lambda_r^r}{(\rho+\eta_r+\lambda_r^r\tilde{F}^r(w)+\lambda_r^c\tilde{F}^c(w^{c*}(w)))} \\
H_{(11)}(w) &= \frac{\eta_r\lambda_r^c}{(\rho+\eta_r+\lambda_r^r\tilde{F}^r(w)+\lambda_r^c\tilde{F}^c(w^{c*}(w)))(\rho+\lambda_r^r+\lambda_r^c)} \\
H_{(12)}(w) &= \frac{\lambda_r^c}{(\rho+\eta_r+\lambda_r^r\tilde{F}^r(w)+\lambda_r^c\tilde{F}^c(w^{c*}(w)))}.
\end{aligned}$$

Similarly, we can rewrite equation (19) as

$$\begin{aligned}
&V^c(w) = D_{(2)}(w) \\
&+\int_{w^{r*}} G_{(21)}(w)V^r(x)dF^r(x)+\int_{w^{r*}(w)} G_{(22)}(w)V^r(x)dF^r(x) \\
&+\int_{w^{c*}} H_{(21)}(w)V^c(x)dF^c(x)+\int_{w} H_{(22)}(w)V^c(x)dF^c(x),
\end{aligned} \tag{21}$$

where $D_{(2)}$, $G_{(21)}$, $G_{(22)}$, $H_{(21)}$ and $H_{(22)}$ are defined by

$$\begin{aligned}
D_{(2)}(w) &= \frac{(\rho+\lambda_c^r+\lambda_c^c)(b_c+w)+\eta_c(b_c+w^{r*})}{(\rho+\eta_c+\lambda_c^r\tilde{F}^r(w^{r*}(w))+\lambda_c^c\tilde{F}^c(w))(\rho+\lambda_c^r+\lambda_c^c)} \\
G_{(21)}(w) &= \frac{\eta_c\lambda_c^r}{(\rho+\eta_c+\lambda_c^r\tilde{F}^r(w^{r*}(w))+\lambda_c^c\tilde{F}^c(w))(\rho+\lambda_c^r+\lambda_c^c)} \\
G_{(22)}(w) &= \frac{\lambda_c^r}{(\rho+\eta_c+\lambda_c^r\tilde{F}^r(w^{r*}(w))+\lambda_c^c\tilde{F}^c(w))} \\
H_{(21)}(w) &= \frac{\eta_c\lambda_c^c}{(\rho+\eta_c+\lambda_c^r\tilde{F}^r(w^{r*}(w))+\lambda_c^c\tilde{F}^c(w))(\rho+\lambda_c^r+\lambda_c^c)} \\
H_{(22)}(w) &= \frac{\lambda_c^c}{(\rho+\eta_c+\lambda_c^r\tilde{F}^r(w^{r*}(w))+\lambda_c^c\tilde{F}^c(w))}.
\end{aligned}$$

To solve the Volterra Type 2 integral equations (20) and (21) numerically, we replace the integrals with the quadrature sums. Let q_1^l denote the quadrature weight at w^l for the integral $\int V^r(x)dF^r(x)$. We assume that there are L quadrature points so that $l = 1, \cdots, L$. Define the following kernel functions:

$$g_{(11)}(w) = \begin{cases} 0 \text{ if } w < w^{r*} \\ 1 \text{ otherwise,} \end{cases} \qquad g_{(12)}(x, w) = \begin{cases} 0 \text{ if } x < w \\ 1 \text{ otherwise.} \end{cases}$$

Given the previous kernel functions, the integrals concerning V^r are approximated by

$$\begin{aligned} \int_{w^{r*}} G_{(21)}(w^m)V^r(x)dF^r(x) &\simeq \sum_{l=1}^{L} G_{(21)}(w^m)g_{(11)}(w^l)V^r(w^l)f^r(w^l)q_1^l \\ \int_{w} G_{(22)}(w^m)V^r(x)dF^r(x) &\simeq \sum_{l=1}^{L} G_{(22)}(w^m)g_{(12)}(w^l, w^m)V^r(w^l)f^r(w^l)q_1^l. \end{aligned}$$

Similarly, let q_2^l denote the quadrature weight for the integral $\int V^c(x)dF^c(x)$. Then the integral equations concerning V^c are approximated by

$$\begin{aligned} \int_{w^{c*}} H_{(11)}(w^m)V^c(x)dF^c(x) &\simeq \sum_{l=1}^{L} H_{(11)}(w)h_{(11)}(w^l)V^c(w^l)f^c(w^l)q_2^l, \\ \int_{w^{c*}(w)} H_{(12)}(w^m)V^c(x)dF^c(x) &\simeq \sum_{l=1}^{L} H_{(12)}(w)h_{(12)}(w^l, w^m)V^c(w^l)f^c(w^l)q_2^l, \end{aligned}$$

where the kernel functions are defined by

$$h_{(11)}(w) = \begin{cases} 0 \text{ if } w < w^{c*} \\ 1 \text{ otherwise,} \end{cases} \qquad h_{(12)}(x, w) = \begin{cases} 0 \text{ if } x < w^{c*}(w) \\ 1 \text{ otherwise.} \end{cases}$$

Using the approximation formulas presented above, we obtain the following approximation for (20): for $m = 1, \cdots, L$

$$\begin{aligned} V^r(w^m) &= D_{(1)}(w^m) \\ &\quad + \sum_{l=1}^{L} \left[G_{(11)}(w^m)g_{(11)}(w^l) + G_{(12)}(w^m)g_{(12)}(w^l, w^m) \right] f^r(w^l)q_1^l V^r(w^l) \\ &\quad + \sum_{l=1}^{L} \left[H_{(11)}(w^m)h_{(11)}(w^l) + H_{(12)}(w^m)h_{(12)}(w^l, w^m) \right] f^c(w^l)q_2^l V^c(w^l), \end{aligned} \tag{22}$$

A similar approximation technique is applied to (21), and we obtain the following approximation: for

$m = 1, \cdots, L$

$$
\begin{aligned}
V^c(w^m) \quad = \quad & D_{(2)}(w^m) \qquad\qquad (23) \\
& + \sum_{l=1}^{L} \left[G_{(21)}(w^m) g_{(21)}(w^l) + G_{(22)}(w^m) g_{(22)}(w^l, w^m) \right] f^r(w^l) q_1^l V^r(w^l) \\
& + \sum_{l=1}^{L} \left[H_{(21)}(w^m) h_{(11)}(w^l) + H_{(22)}(w^m) g_{(12)}(w^l, w^m) \right] f^c(w^l) q_2^l V^c(w^l),
\end{aligned}
$$

where kernel functions are defined by

$$
g_{(21)}(w) = \begin{cases} 0 \text{ if } w < w^{r*} \\ 1 \text{ otherwise,} \end{cases} \qquad g_{(22)}(x, w) = \begin{cases} 0 \text{ if } x < w^{r*}(w) \\ 1 \text{ otherwise,} \end{cases}
$$

$$
h_{(21)}(w) = \begin{cases} 0 \text{ if } w < w^{r*}(w) \\ 1 \text{ otherwise,} \end{cases} \qquad h_{(22)}(x, w) = \begin{cases} 0 \text{ if } x < w \\ 1 \text{ otherwise.} \end{cases}
$$

Using the quadrature equations (22) and (23), the value functions, $V^r(w)$ and $V^c(w)$, are solved at all the quadrature nodes, $(w^1, \cdots, w^m, \cdots, w^L)$. For notational simplicity, define the following $L \times 1$ vectors:

$$
\begin{aligned}
\mathbf{V}^r &= \left[V^r(w^1), \cdots, V^r(w^L) \right]', \\
\mathbf{V}^c &= \left[V^c(w^1), \cdots, V^c(w^L) \right]', \\
\mathbf{D}_1 &= \left[D_{(1)}(w^1), \cdots, D_{(1)}(w^L) \right]'.
\end{aligned}
$$

Then equation (22) is expressed by

$$
\mathbf{D}_1 = [\mathbf{F}_1 \quad \mathbf{P}_1] \begin{bmatrix} \mathbf{V}^r \\ \mathbf{V}^c \end{bmatrix}, \qquad (24)
$$

where the previous matrices are given by

$$\mathbf{F}_1 = \begin{pmatrix} 1-G_{(1)}(w^1,w^1) & \cdots & -G_{(1)}(w^1,w^l) & \cdots & -G_{(1)}(w^1,w^L) \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ -G_{(1)}(w^m,w^1) & \cdots & 1-G_{(1)}(w^m,w^l) & \cdots & -G_{(1)}(w^m,w^L) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -G_{(1)}(w^L,w^1) & \cdots & -G_{(1)}(w^m,w^l) & \cdots & 1-G_{(1)}(w^L,w^L) \end{pmatrix}.$$

$$\mathbf{P}_1 = \begin{pmatrix} -H_{(1)}(w^1,w^1) & \cdots & -H_{(1)}(w^1,w^l) & \cdots & -H_{(1)}(w^1,w^L) \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ -H_{(1)}(w^m,w^1) & \cdots & -H_{(1)}(w^m,w^l) & \cdots & -H_{(1)}(w^m,w^L) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -H_{(1)}(w^L,w^1) & \cdots & -H_{(1)}(w^m,w^l) & \cdots & -H_{(1)}(w^L,w^L) \end{pmatrix}.$$

The typical elements of $\mathbf{F}_1$ and $\mathbf{P}_1$ are given by

$$G_{(1)}(w^m,w^l) = \left[G_{(11)}(w^m)g_{(11)}(w^l) + G_{(12)}(w^m)g_{(12)}(w^l,w^m)\right] f^r(w^l)q_1^l,$$
$$H_{(1)}(w^m,w^l) = \left[H_{(11)}(w^m)h_{(11)}(w^l) + H_{(12)}(w^m)h_{(12)}(w^l,w^m)\right] f^c(w^l)q_2^l.$$

Similarly, using the following $L \times 1$vectors

$$\begin{aligned} \mathbf{D}_2 &= [D_{(2)}(w^1),\cdots,D_{(2)}(w^L)]' \\ \mathbf{V}^r &= \left(V^r(w^1),\cdots,V^r(w^L)\right)' \\ \mathbf{V}^c &= \left(V^c(w^1),\cdots,V^c(w^L)\right)'. \end{aligned}$$

Then equation (22) is expressed by

$$\mathbf{D}_2 = [\mathbf{F}_2\ \ \mathbf{P}_2] \begin{bmatrix} \mathbf{V}^r \\ \mathbf{V}^c \end{bmatrix}, \tag{25}$$

where we define

$$\mathbf{F}_2 = \begin{pmatrix} -G_{(2)}(w^1, w^1) & \cdots & -G_{(2)}(w^1, w^l) & \cdots & -G_{(2)}(w^1, w^L) \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ -G_{(2)}(w^m, w^1) & \cdots & -G_{(2)}(w^m, w^l) & \cdots & -G_{(2)}(w^m, w^L) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -G_{(2)}(w^L, w^1) & \cdots & -G_{(2)}(w^m, w^l) & \cdots & -G_{(2)}(w^L, w^L) \end{pmatrix},$$

$$\mathbf{P}_2 = \begin{pmatrix} 1 - H_{(2)}(w^1, w^1) & \cdots & -H_{(2)}(w^1, w^l) & \cdots & -H_{(2)}(w^1, w^L) \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ -H_{(2)}(w^m, w^1) & \cdots & 1 - H_{(2)}(w^m, w^l) & \cdots & -H_{(2)}(w^m, w^L) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -H_{(2)}(w^L, w^1) & \cdots & -H_{(2)}(w^m, w^l) & \cdots & 1 - H_{(2)}(w^L, w^L) \end{pmatrix},$$

The typical elements of $\mathbf{F}_2$ and $\mathbf{P}_2$ are given by

$$\begin{aligned} G_{(2)}(w^m, w^l) &= \left[G_{(21)}(w^m) g_{(11)}(w^l) + G_{(22)}(w^m) g_{(22)}(w^l, w^m) \right] f^r(w^l) q_1^l, \\ H_{(2)}(w^m, w^l) &= \left[H_{(21)}(w^m) h_{(11)}(w^l) + H_{(22)}(w^m) g_{(12)}(w^l, w^m) \right] f^c(w^l) q_2^l. \end{aligned}$$

Combining these matrix equations (24) and (25), we have

$$\begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{P}_1 \\ \mathbf{F}_2 & \mathbf{P}_2 \end{bmatrix} \begin{bmatrix} \mathbf{V}^r \\ \mathbf{V}^c \end{bmatrix}.$$

Therefore, the value functions, $V^r(w)$ and $V^c(w)$, evaluated at the quadrature nodes, $(w^1, \cdots, w^m, \cdots, w^L)$, are obtained by

$$\begin{bmatrix} \mathbf{V}^r \\ \mathbf{V}^c \end{bmatrix} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{P}_1 \\ \mathbf{F}_2 & \mathbf{P}_2 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \end{bmatrix}.$$

We finally consider a quadrature method used to approximate the integrals. Although other quadrature methods might be applicable, we find that the *Gauss-Legendre* quadrature is appropriate for the system of the integral equations presented above. The Gauss-Legendre quadrature formula is defined by [20]

$$\int_{-1}^{1} g(z) dz = \sum_{l=1}^{L} g(z_l) p_l + \frac{2^{2L+1}(L!)^4}{(2L+1)!(2L)!} \frac{g^{(2L)}(\xi)}{(2L)!}$$

for some $\xi \in [-1, 1]$, where the quadrature nodes, z_l, are given by the roots of the Legendre polynomi-

[20] See Judd (1988) Chapter 7 for details.

als $P_L(z)$. The quadrature weights, p_l, are given by

$$p_l = \frac{A_{l+1}\gamma_l}{A_l P_l'(z_l)P_{l+1}(z_l)},$$

where $A_l = \frac{(2l)!}{2^l(l!)^2}$ is the coefficient of x^l in $P_l(z)$ and $\gamma_l = \int_{-1}^{1}[P_l(z)]^2 dz = \frac{2}{2l+1}$.

It should be noted that, in order to apply the Gauss-Legendre quadrature formula, we need to change the variable x (wage) with domain $(0, \infty)$ so that the domain of the new variable (call it z) becomes $[-1, 1]$. The change of the variable $w = \phi(z) = \tau\frac{1+z}{1-z}$ with $\tau > 0$ with Jacobian $\phi'(z) = \frac{2\tau}{(1-z)^2}$ accomplishes this task. Then we have, for a generic function $M(w)$, and any type $k \in \{r, c\}$,

$$\begin{aligned}
&\int_0^\infty K(w)V^k(w)dF^k(w) \\
&= \int_{-1}^{1} M\left(\tau\frac{1+z}{1-z}\right) V^k\left(\tau\frac{1+z}{1-z}\right) f^k\left(\tau\frac{1+z}{1-z}\right) \frac{2\tau}{(1-z)^2}dz \\
&\simeq \sum_{l=1}^{L}\left[M(w_l)V^k(w_l)f^l(w_l)\right] q_l,
\end{aligned}$$

where $w_l = \tau\frac{1+z_l}{1-z_l}$ and $q_l = p_l\frac{2\tau}{(1-z_l)^2}$ for $l = 1, \cdots, L$.

C Simulation Details

In this paper, we conduct simulation exercises for at least two purposes. The first is to obtain transition probabilities from initial employment state to other employment states after a certain number of periods. The second is to calculate the welfare of the first 5, 10 and 40 years in the life of a worker. In what follows, we provide details on how we calculate the welfare for the first 40 years of the life of a worker. The welfare for other time periods can be calculated by the same method.

We simulate the job careers of young workers for each initial employment status (unemployed, contingent employed, and regular employed). Time is continuous, but one "period" in the model and the simulation corresponds to one month. The generic spell is indexed by i, and it begins at time τ_i, where $\tau_1 = 0$. The total duration of spell is denoted by t_i.

If spell i is an unemployment spell, we first generate a draw t_i from an exponential distribution with parameter $\lambda_u^r\tilde{F}^r(w^{r*}) + \lambda_u^c\tilde{F}^c(w^{c*}) = \lambda_u^r + \lambda_u^c$. Workers may exit from the unemployment pool with either a regular or a contingent job offer. We draw x from a uniform distribution on the unit interval. If $x < \lambda_u^c/(\lambda_u^r + \lambda_u^c)$, we then draw a wage w_{i+1} for the contingent job from the accepted wage offer distribution $F^c(w|w \geq w^{c*})$. If $x \geq \lambda_u^c/(\lambda_u^r + \lambda_u^c)$, we then draw a wage w_{i+1} for the regular job from the accepted wage offer distribution $F^r(w|w \geq w^{r*})$. The contribution of this unemployment spell to the welfare is given by

$$\begin{aligned} V_i &= \exp(-\rho\tau_i)\int_0^{t_i} b_u \exp(-\rho u)du \\ &= \rho^{-1}\exp(-\rho\tau_i)\left[1-\exp(-\rho t_i)\right] b_u. \end{aligned}$$

Since we normalize the unemployment benefit to zero, the contribution of the unemployment spell to the welfare is also zero. The next spell starts from $\tau_{i+1} = \tau_i + t_i$.

If spell i is an employment spell with a contingent job with wage w_i, we first generate a draw t_i from an exponential distribution with parameter $D_c(w_i) = \eta_c + \lambda_c^r \tilde{F}^r(w^{r*}(w_i)) + \lambda_c^c \tilde{F}^c(w_i)$. There are three ways to end this spell, either through an exogenous separation, a quit into a higher-paying contingent job, or a quit into a regular job with wage above the corresponding reservation wage. We generate x from a uniform distribution on the unit interval.

If

$$x < \frac{\eta_c}{\eta_c + \lambda_c^r \tilde{F}^r(w^{r*}(w_i)) + \lambda_c^c \tilde{F}^c(w_i)},$$

the spell is considered to have ended in a dismissal so that spell $i+1$ is an unemployment spell.

If

$$\frac{\eta_c}{\eta_c + \lambda_c^r \tilde{F}^r(w^{r*}(w_i)) + \lambda_c^c \tilde{F}^c(w_i)} \leq x < \frac{\eta_c + \lambda_c^c \tilde{F}^c(w_i)}{\eta_c + \lambda_c^r \tilde{F}^r(w^{r*}(w_i)) + \lambda_c^c \tilde{F}^c(w_i)},$$

then the next spell is an employment spell with a higher-paying contingent job. In this case, we draw a wage w_{i+1} for the next contingent job from the wage offer distribution $F^c(w|w \geq w_i)$.

If

$$\frac{\eta_c + \lambda_c^c \tilde{F}^c(w_i)}{\eta_c + \lambda_c^r \tilde{F}^r(w^{r*}(w_i)) + \lambda_c^c \tilde{F}^c(w_i)} \leq x,$$

then the next spell is an employment spell with a regular job. In this case, we draw a wage w_{i+1} for the next regular job from the wage offer distribution $F^r(w|w \geq w^{r*}(w_i))$. The contribution of this employment spell to the welfare is given by

$$V_i = \rho^{-1}\exp(-\rho\tau_i)\left[1-\exp(-\rho t_i)\right]\left[b_c + w_i\right].$$

If spell i is an employment spell with a regular job with w_i, we proceed in a similar fashion, but with different distributions of wage and job spell. We first generate a draw t_i from an exponential distribution with parameter $D_r(w_i) = \eta_r + \lambda_r^r \tilde{F}^r(w_i) + \lambda_r^c \tilde{F}^c(w^{c*}(w_i))$. Then we generate x from a uniform distribution on the unit interval.

If

$$x < \frac{\eta_r}{\eta_r + \lambda_r^r \tilde{F}^r(w_i) + \lambda_r^c \tilde{F}^c(w^{c*}(w_i))},$$

the spell is considered to have ended in a dismissal so that spell $i+1$ is an unemployment spell.

If

$$\frac{\eta_r}{\eta_r + \lambda_r^r \tilde{F}^r(w_i) + \lambda_r^c \tilde{F}^c(w^{c*}(w_i))} \leq x < \frac{\eta_r + \lambda_r^c \tilde{F}^c(w^{c*}(w_i))}{\eta_r + \lambda_r^r \tilde{F}^r(w_i) + \lambda_r^c \tilde{F}^c(w^{c*}(w_i))},$$

then the next spell is an employment spell with a higher-paying contingent job. In this case, we draw a wage w_{i+1} for the next contingent job from the wage offer distribution $F^c(w|w \geq w^{c*}(w_i))$.

If

$$\frac{\eta_r + \lambda_r^c \tilde{F}^c(w^{c*}(w_i))}{\eta_r + \lambda_r^r \tilde{F}^r(w_i) + \lambda_r^c \tilde{F}^c(w^{c*}(w_i))} \leq x,$$

then the next spell is an employment spell with a regular job. In this case, we draw a wage w_{i+1} for the next regular job from the wage offer distribution $F^r(w|w \geq w_i)$. The contribution of this employment spell to the welfare is given by

$$V_i = \rho^{-1} \exp(-\rho\tau_i) \left[1 - \exp(-\rho t_i)\right] \left[b_r + w_i\right].$$

Let N_j denote the number of spells which commenced prior to the 480th month (i.e., 40 years) for worker j. The last spell is truncated at $\tau_{N_j} = 480$. Then worker j's 40 years' labor market career of the worker j has value $\sum_{i=1}^{N_j} V_i$.

Table 1: The Distribution of Labor Market Profiles

Profiles	Distribution (%)
R-U-R	2.27
R-U-C	1.45
C-U-R	0.86
C-U-C	0.84
R-R	0.69
R-C	0.53
R-U	3.03
C-R	0.46
C-C	0.3
C-U	1.96
R	76.33
C	6.64
U	4.63

Table 2: Summary Statistics

	All samples $n = 6257$	High School graduates $n = 2613$	College graduates $n = 2552$
Current Labor Status Distribution (%)			
Regular	79.61	76.54	86.09
Contingent	11.06	13.24	6.43
Unemployed	9.33	10.22	7.48
Mean Duration (month)			
Employment Spell	18.37	18.24	18.75
	(12.59)	(12.64)	(12.40)
Unemployment Spell	7.43	7.03	7.58
	(7.16)	(6.89)	(6.48)
Mean Hourly Wage (yen)	1666.61	1477.01	1903.43
	(857.76)	(903.41)	(789.58)
Mean Age	21.87	19.87	24.17
	(2.42)	(1.20)	(1.18)

Standard errors are in parentheses

Table 3: Partial Likelihood Estimation Result

Parameter	All Samples	High School Grad	College Grad
$\hat{\mu}_r$	7.2538	7.1664	7.4345
	(0.0404)	(0.5887)	(0.0369)
$\hat{\mu}_c$	7.1036	7.0334	7.2149
	(0.0445)	(0.0614)	(0.1953)
$\hat{\sigma}_r$	0.3485	0.3318	0.2886
	(0.0383)	(0.2781)	(0.0337)
$\hat{\sigma}_c$	0.4856	0.4779	0.5913
	(0.0291)	(0.2540)	(0.1140)
$\hat{\sigma}_\epsilon$	0.0481	0.0535	0.0436
	(0.0675)	(0.3341)	(0.0588)
$\hat{\lambda}_u^r$	0.0310	0.0301	0.0324
	(0.0023)	(0.0033)	(0.0040)
$\hat{\lambda}_u^c$	0.0227	0.0319	0.0115
	(0.0019)	(0.0034)	(0.0024)
$\hat{w}^{r*}$	432.10	432.10	816.67
$\hat{w}^{c*}$	496.45	496.45	729.17
Log Likelihood	2525.53	1280.14	678.54
N	632	333	164

Standard errors are in parentheses

Table 4: Total Likelihood Estimation Result

Parameter	All Samples	High School Grad	College Grad
$\hat{\eta}_r$	0.0041	0.0052	0.0029
	(0.0002)	(0.0004)	(0.0003)
$\hat{\eta}_c$	0.0228	0.0262	0.0156
	(0.0016)	(0.0024)	(0.0026)
$\hat{\lambda}_r^r$	0.0011	0.0016	0.0005
	(0.0002)	(0.0003)	(0.0002)
$\hat{\lambda}_r^c$	0.0063	0.0057	0.0011
	(0.0011)	(0.0013)	(0.0006)
$\hat{\lambda}_c^c$	0.0038	0.0043	0.0017
	(0.0009)	(0.0013)	(0.0011)
$\hat{\lambda}_c^r$	0.0031	0.0017	0.0060
	(0.0006)	(0.0006)	(0.0017)
$\hat{b}^r$	3072.66	2455.40	3578.41
$\hat{b}^c$	2874.99	2417.92	2960.17
Log Likelihood	47650.99	19623.61	19364.54
N	6040	2492	2498

Standard errors are in parentheses

Table 5: Welfare Implication of the Initial Employment Status

	All Samples			College Graduates			High school Graduates		
	5 years	10 years	40 years	5 years	10 years	40 years	5 years	10 years	40 years
A. Mean									
Initially regular	2.234E+005	3.839E+005	8.126E+005	2.705E+005	4.731E+005	1.022E+006	1.861E+005	3.205E+005	6.767E+005
Initially contingent	1.856E+005	3.343E+005	7.602E+005	2.209E+005	4.031E+005	9.439E+005	1.579E+005	2.827E+005	6.339E+005
Initially unemployed	1.433E+005	2.971E+005	7.241E+005	1.634E+005	3.557E+005	9.014E+005	1.247E+005	2.531E+005	6.056E+005
Aggregated case	2.130E+005	3.720E+005	8.000E+005	2.610E+005	4.620E+005	1.010E+006	1.780E+005	3.110E+005	6.660E+005
B. CV squared									
Initially regular	0.042	0.039	0.019	0.030	0.028	0.016	0.045	0.039	0.024
Initially contingent	0.099	0.063	0.023	0.103	0.067	0.023	0.097	0.061	0.031
Initially unemployed	0.234	0.100	0.030	0.283	0.109	0.029	0.202	0.090	0.036
Aggregated case	0.066	0.049	0.021	0.050	0.037	0.018	0.067	0.049	0.026
C. Gini coefficient									
Initially regular	0.108	0.107	0.077	0.086	0.087	0.069	0.112	0.108	0.081
Initially contingent	0.173	0.139	0.085	0.173	0.141	0.084	0.172	0.136	0.091
Initially unemployed	0.275	0.177	0.096	0.302	0.184	0.096	0.254	0.168	0.100
Aggregated case	0.135	0.121	0.081	0.111	0.101	0.073	0.137	0.121	0.085

Table 6: Welfare Implication of the Policy Experiments

	Experiment 1 ($\lambda_c^r = \lambda_u^r$)			Experiment 2 ($b_c = b_r$)			Experiment 3 ($\eta_c = \eta_r$)		
	5 years	10 years	40 years	5 years	10 years	40 years	5 years	10 years	40 years
A. Mean									
Initially regular	0.92	2.31	4.43	0.54	0.79	1.10	2.35	4.51	7.49
Initially contingent	8.01	8.96	7.81	3.13	2.52	1.93	19.65	17.83	14.21
Initially unemployed	2.90	4.64	5.70	1.30	1.39	1.39	8.09	9.63	10.22
Aggregated case	1.80	3.18	4.91	0.96	1.10	1.27	4.40	6.16	8.37
B. CV squared									
Initially regular	-5.66	-12.95	-17.80	-1.89	-3.37	-4.71	-20.52	-28.50	-27.23
Initially contingent	-17.61	-19.17	-19.74	-0.81	-1.12	-3.43	-58.50	-46.81	-30.04
Initially unemployed	-2.86	-10.14	-16.27	0.13	-1.71	-3.39	-14.05	-24.20	-22.71
Aggregated case	-8.78	-14.87	-18.78	-3.48	-4.48	-5.16	-26.63	-31.77	-28.17
C. Gini coefficient									
Initially regular	-3.99	-7.86	-9.94	-1.30	-1.87	-2.48	-12.35	-17.12	-16.21
Initially contingent	-11.17	-11.42	-10.93	-0.29	-0.58	-1.76	-39.76	-30.49	-19.51
Initially unemployed	-1.69	-5.71	-8.84	0.04	-0.79	-1.77	-8.59	-14.54	-14.03
Aggregated case	-6.14	-9.18	-10.37	-2.29	-2.56	-2.72	-17.00	-19.44	-17.04

All numbers are the percentage change from the benchmark case

Figure 1: Transition Probabilities to Regular Employment

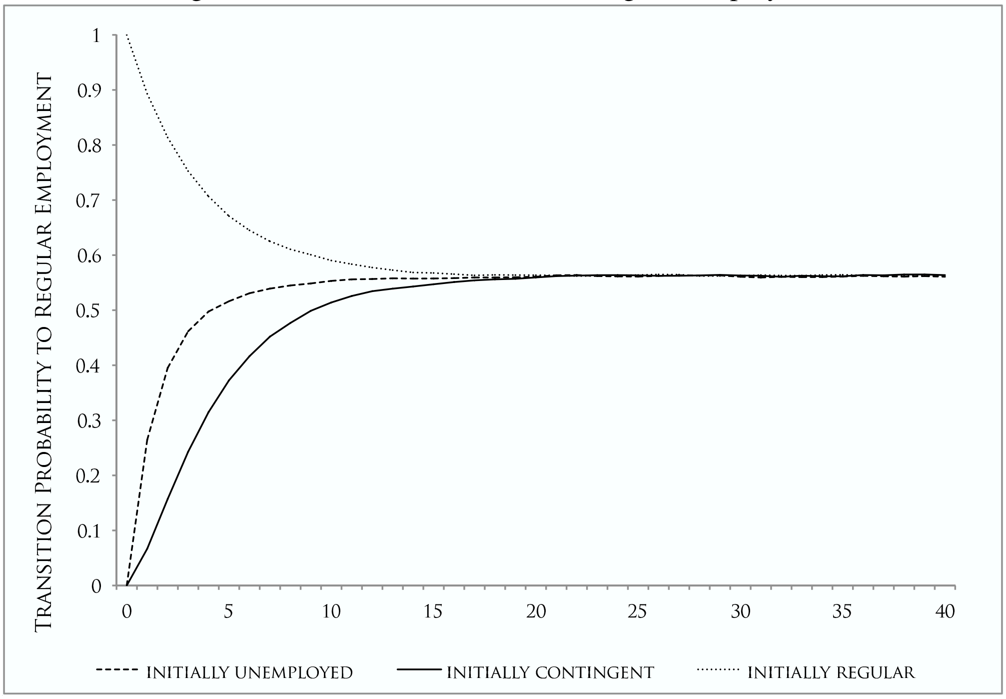
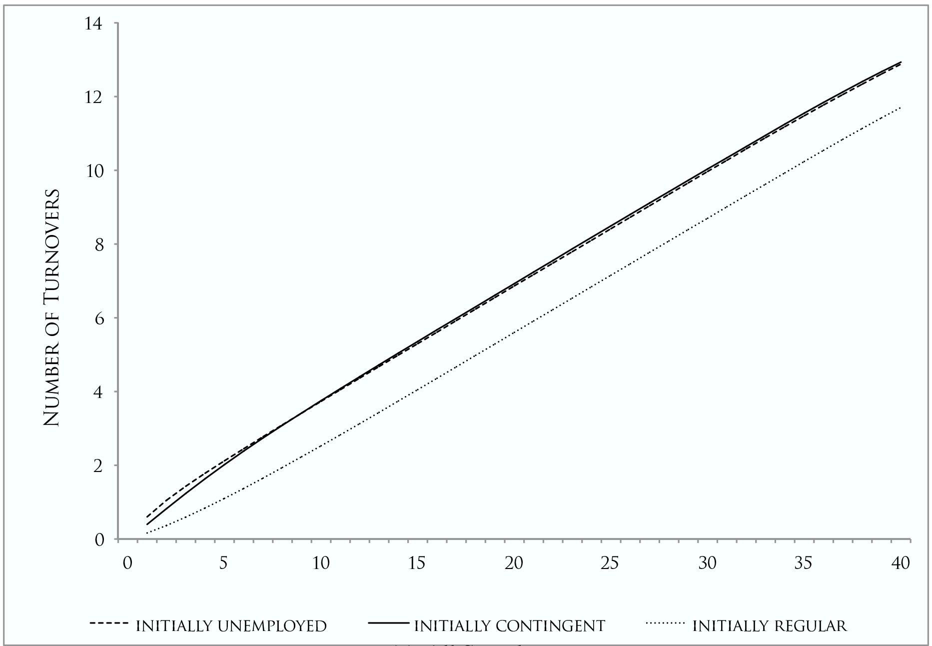
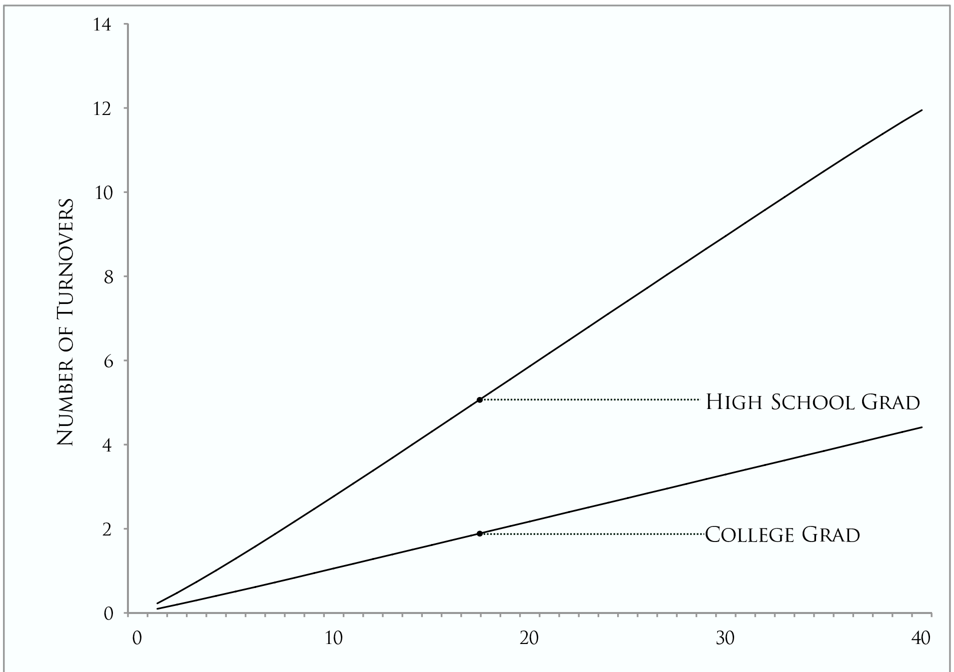


Figure 2: The Average Number of Turnovers

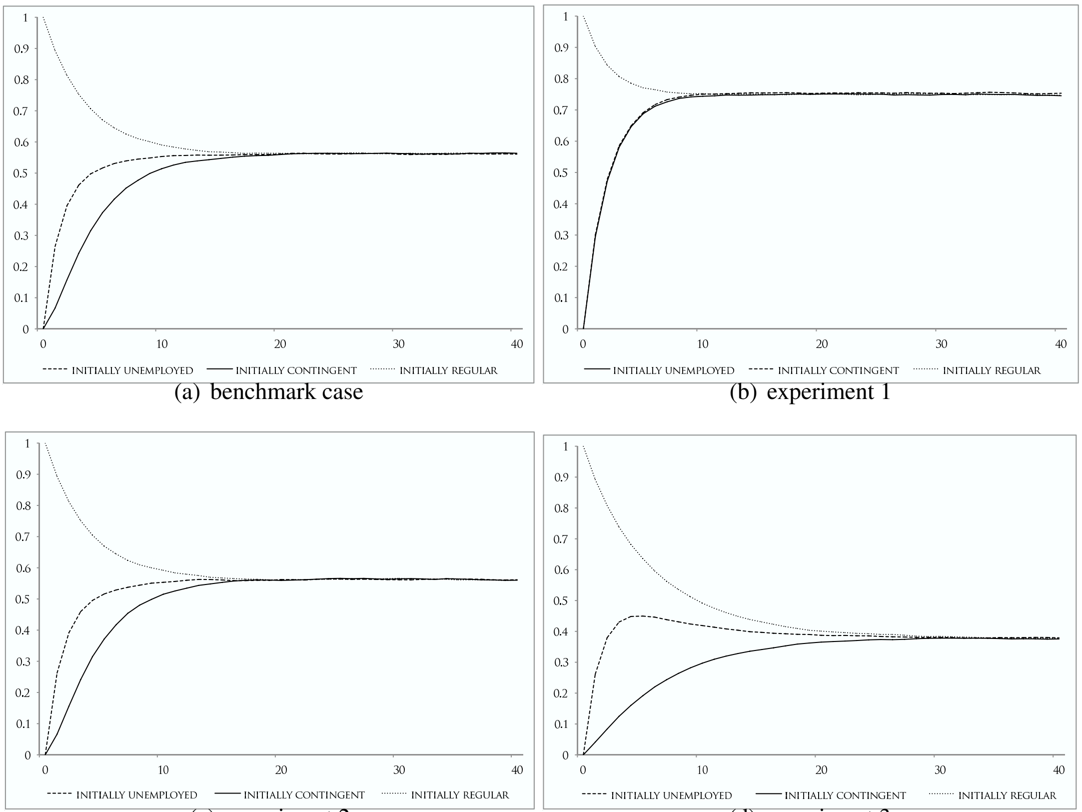


(a) All Samples



(b) Dissaggregate Samples By Education

Figure 3: Policy Simulation: Transition Probabilities to Regular Employment



(a) benchmark case

(b) experiment 1

(c) experiment 2

(d) experiment 3