Online Appendix to "Occupational Matching and Cities"

Theodore Papageorgiou

*Boston College, Department of Economics**

September 2020

1 Model with Occupation-Specific Human Capital

In this appendix I derive optimal worker behavior when I extend the baseline model to allow workers to accumulate occupation-specific human capital.

There are two experience levels in each occupation: workers are either experienced or inexperienced. A worker who enters an occupation for the first time is inexperienced. As in Kambourov and Manovskii (2009), a worker becomes experienced stochastically at rate θ. Output production increases by a known amount u for all experienced workers.

In order to solve the worker's occupational choice problem, as in the baseline model, I follow Whittle (1982) and Karatzas (1984). More specifically, I compute the retirement value at which the worker is exactly indifferent between continuing with that occupation or retiring. This retirement value serves as an index for each occupation, which corresponds to that occupation's Gittins index (see Gittins, 1979 and Bergemann and Valimaki, 2008).

Gittins Index for Experienced Worker

I start by deriving the optimal behavior of an experienced worker. The solution to the transformed problem in this case is similar to the baseline model.

The value of an experienced worker with belief p and the option of retiring and obtaining value W is given by

$$rV^{\exp}(p,W) = \alpha_G p + \alpha_B(1-p) + u + \frac{1}{2}\left(\frac{\alpha_G-\alpha_B}{\sigma}\right)^2 p^2(1-p)^2 V^{\exp}_{pp}(p,W) - \delta\left(V^{\exp}(p,W) - J\right)$$

*Maloney Hall 343, Chestnut Hill, MA, 02467, USA, theodore.papageorgiou@bc.edu

where I suppress k, the occupational index in this section, to reduce notational congestion. In other words the worker's flow value consists of his expected output, which is now increased by the amount u since he is experienced; the term that captures the value of learning; and a term that captures the possibility of an exogenous move that occurs at rate δ.

The optimal stopping rule is to retire when p reaches $\widetilde{p}^{\exp}(W)$ such that the value matching and the smooth pasting conditions hold

$$V^{\exp}\left(\widetilde{p}^{\exp}(W), W\right) = W \tag{1}$$

$$V_p^{\exp}\left(\widetilde{p}^{\exp}(W), W\right) = 0.$$

Similarly to the baseline case, the solution to the above differential equation is given by

$$\begin{aligned} V^{\exp}(p, W) &= \frac{\alpha_G p + \alpha_B (1-p) + u + \delta J}{r+\delta} \\ &+ \frac{\alpha_G - \alpha_B}{r+\delta}\left(\widetilde{p}^{\exp}(W) + \frac{1}{2}d - \frac{1}{2}\right)^{-1} \widetilde{p}^{\exp}(W)^{\frac{1}{2}+\frac{1}{2}d}\left(1-\widetilde{p}^{\exp}(W)\right)^{\frac{1}{2}-\frac{1}{2}d} \\ &\times (p)^{\frac{1}{2}-\frac{1}{2}d}(1-p)^{\frac{1}{2}+\frac{1}{2}d}, \end{aligned} \tag{2}$$

where

$$\widetilde{p}^{\exp}(W) = \frac{(d-1)\left((r+\delta)W - \alpha_B - u - \delta J\right)}{(d+1)(\alpha_G - \alpha_B) - 2\left((r+\delta)W - \alpha_B - u - \delta J\right)}, \tag{3}$$

and $d = \sqrt{\frac{8(r+\delta)}{\left(\frac{\alpha_G - \alpha_B}{\sigma}\right)^2} + 1}$.

The index of each occupation is the highest retirement value at which the worker is indifferent between working at that occupation or retiring with $W = W^{\exp}(p)$, i.e.

$$W^{\exp}(p) = V^{\exp}(p, W), \tag{4}$$

where $W^{\exp}(p) = \max\left\{\widetilde{W}\right\}$ and the set $\left\{\widetilde{W}\right\}$ includes all possible retirement values, $\widetilde{W}$, such that $\widetilde{W} = V^{\exp}\left(p, \widetilde{W}\right)$.

For eq. (4) to hold, from eq. (1), it must be the case that

$$p = \widetilde{p}^{\exp}(W). \tag{5}$$

Substituting condition (5) into the threshold condition, equation (3), obtains

$$\begin{aligned} p &= \frac{(d-1)\left((r+\delta) W^{\exp}(p)-\alpha_B-u-\delta J\right)}{(d+1)\left(\alpha_G-\alpha_B\right)-2\left((r+\delta) W^{\exp}(p)-\alpha_B-u-\delta J\right)} \Rightarrow \\ W^{\exp}(p) &= \frac{1}{r+\delta} \frac{(d+1)\left(\alpha_G-\alpha_B\right) p+(2p+d-1)\left(\alpha_B+u+\delta J\right)}{2p+d-1}, \end{aligned} \tag{6}$$

which is the Gittins index of an experienced worker with beliefs p.

Gittins Index for Inexperienced Worker

I now turn to the optimal behavior of an inexperienced worker.

In the transformed problem, the value of an inexperienced worker with the option of retiring and obtaining value W is given by

$$\begin{aligned} rV^{\text{in}}(p, W) =\ & \alpha_G p+\alpha_B(1-p) \\ & +\frac{1}{2}\left(\frac{\alpha_G-\alpha_B}{\sigma}\right)^2 p^2(1-p)^2 V_{pp}^{in}(p, W) \\ & -\delta\left(V^{in}(p, W)-J\right)+\theta\left(V^{\exp}(p, W)-V^{in}(p, W)\right), \end{aligned} \tag{7}$$

i.e. the flow value of an inexperienced worker equals his expected output plus the term that captures the option value of learning; at rate δ the worker moves exogenously to a new city, whereas at rate θ the worker becomes experienced and his value becomes $V^{\exp}(p, W)$, derived above.

Substituting into (7) the expression for $V^{\exp}(p, W)$ using equations (2) and (3) derived above, leads to

$$\begin{aligned} & (r+\theta+\delta) V^{\text{in}}(p, W) \\ =\ & \frac{r+\delta+\theta}{r+\delta}\left(\alpha_G-\alpha_B\right) p+\frac{r+\delta+\theta}{r+\delta} \alpha_B \\ & +\frac{1}{2}\left(\frac{\alpha_G-\alpha_B}{\sigma}\right)^2 p^2(1-p)^2 V_{pp}^{in}(p, W) \\ & +\theta \frac{\alpha_G-\alpha_B}{r+\delta}\left(\widetilde{p}^{\exp}(W)+\frac{1}{2} d-\frac{1}{2}\right)^{-1} \widetilde{p}^{\exp}(W)^{\frac{1}{2}+\frac{1}{2} d}\left(1-\widetilde{p}^{\exp}(W)\right)^{\frac{1}{2}-\frac{1}{2} d} \\ & \times p^{\frac{1}{2}-\frac{1}{2} d}(1-p)^{\frac{1}{2}+\frac{1}{2} d} \\ & +\frac{r+\delta+\theta}{r+\delta} \delta J+\frac{\theta}{r+\delta} u \end{aligned}$$

or equivalently:

$$
\begin{aligned}
& \frac{1}{2\left(r+\theta+\delta\right)}\left(\frac{\alpha_G-\alpha_B}{\sigma}\right)^2 p^2\left(1-p\right)^2 V_{pp}^{in}\left(p,W\right) \\
= \; & V^{\text{in}}\left(p,W\right) \\
& -\frac{1}{r+\delta}\left(\alpha_G-\alpha_B\right)p \\
& -\frac{\theta}{r+\theta+\delta}\frac{\alpha_G-\alpha_B}{r+\delta}\left(\widetilde{p}^{\text{exp}}\left(W\right)+\frac{1}{2}d-\frac{1}{2}\right)^{-1}\widetilde{p}^{\text{exp}}\left(W\right)^{\frac{1}{2}+\frac{1}{2}d}\left(1-\widetilde{p}^{\text{exp}}\left(W\right)\right)^{\frac{1}{2}-\frac{1}{2}d} \\
& \times p^{\frac{1}{2}-\frac{1}{2}d}\left(1-p\right)^{\frac{1}{2}+\frac{1}{2}d} \\
& -\left(\frac{1}{r+\delta}\alpha_B+\frac{\delta}{r+\delta}J+\frac{1}{\left(r+\theta+\delta\right)}\frac{\theta}{r+\delta}u\right).
\end{aligned}
$$

As before, the optimal stopping rule is to retire when p reaches $\widetilde{p}^{\text{in}}\left(W\right)$ such that the value matching and the smooth pasting conditions hold:

$$V^{\text{in}}\left(\widetilde{p}^{\text{in}}\left(W\right),W\right)=W \tag{8}$$

$$V_p^{\text{in}}\left(\widetilde{p}^{\text{in}}\left(W\right),W\right)=0. \tag{9}$$

The above is a second-order, non-homogeneous differential equation of the form

$$k_3t^2\left(1-t\right)^2\ddot{y}\left(t\right)=y\left(t\right)+k_1t+g\left(W\right)t^{\beta}\left(1-t\right)^{1-\beta}+k_2, \tag{10}$$

where

$$y\left(t\right)=V^{in}\left(p,W\right)$$

$$k_1=-\frac{1}{r+\delta}\left(\alpha_G-\alpha_B\right)$$

$$k_2=-\left(\frac{\alpha_B+\delta J}{r+\delta}+\frac{1}{\left(r+\theta+\delta\right)}\frac{\theta}{r+\delta}u\right)$$

$$k_3=\frac{1}{2}\frac{1}{r+\theta+\delta}\left(\frac{\alpha_G-\alpha_B}{\sigma}\right)^2$$

$$
\begin{aligned}
g\left(W\right) &= \frac{-\frac{\theta}{r+\theta+\delta}\frac{\alpha_G-\alpha_B}{r+\delta}}{\frac{(d-1)((r+\delta)W-\alpha_B-u-\delta J)}{(d+1)(\alpha_G-\alpha_B)-2((r+\delta)W-\alpha_B-u-\delta J)}+\frac{1}{2}d-\frac{1}{2}}\times \\
&\left(\frac{\left(d-1\right)\left(\left(r+\delta\right)W-\alpha_B-u-\delta J\right)}{\left(d+1\right)\left(\alpha_G-\alpha_B\right)-2\left(\left(r+\delta\right)W-\alpha_B-u-\delta J\right)}\right)^{\frac{1}{2}+\frac{1}{2}d}\times \\
&\left(1-\frac{\left(d-1\right)\left(\left(r+\delta\right)W-\alpha_B-u-\delta J\right)}{\left(d+1\right)\left(\alpha_G-\alpha_B\right)-2\left(\left(r+\delta\right)W-\alpha_B-u-\delta J\right)}\right)^{\frac{1}{2}-\frac{1}{2}d}
\end{aligned}
$$

$$
d=\sqrt{\frac{8\left(r+\delta\right)}{\left(\frac{\alpha_G-\alpha_B}{\sigma}\right)^2}+1}
$$

$$
\beta=\frac{1}{2}-\frac{1}{2}\sqrt{\frac{8\left(r+\delta\right)}{\left(\frac{\alpha_G-\alpha_B}{\sigma}\right)^2}+1}
$$

subject to

$$
y\left(t_0\right)=W
$$

$$
\dot{y}\left(t_0\right)=0.
$$

In order to solve the above equation, I follow Brockett (1970), Chapter 1.

In particular, let

$$
\begin{aligned}
x_1\left(t\right) &= y\left(t\right) \\
x_2\left(t\right) &= \dot{y}\left(t\right).
\end{aligned}
$$

Then I can write

$$
\begin{aligned}
\dot{x}_1\left(t\right) &= x_2\left(t\right) \\
\dot{x}_2\left(t\right) &= \ddot{y}\left(t\right)=\frac{1}{k_3t^2\left(1-t\right)^2}x_1\left(t\right)+Z\left(t\right),
\end{aligned}
$$

where

$$
Z\left(t\right)=\frac{1}{k_3t^2\left(1-t\right)^2}\left[k_1t+g\left(W\right)t^{\beta}\left(1-t\right)^{1-\beta}+k_2\right].
$$

Moreover let

$$x\left(t\right)=\begin{bmatrix}x_{1}\left(t\right) & x_{2}\left(t\right)\end{bmatrix}^{\prime}.$$

Then

$$\dot{x}\left(t\right)=\begin{bmatrix}0 & 1\\ \frac{1}{k_{3}t^{2}\left(1-t\right)^{2}} & 0\end{bmatrix}x\left(t\right)+\begin{bmatrix}0\\ 1\end{bmatrix}Z\left(t\right).$$

I now use the "Variation of Constants formula". According to Brockett (1970), Chapter 1.6, if $\Phi\left(t,t_{0}\right)$ is the transition matrix for $\dot{x}\left(t\right)=A\left(t\right)x\left(t\right)$, then the unique solution of $\dot{x}\left(t\right)=A\left(t\right)x\left(t\right)+f\left(t\right)$; $x\left(t_{0}\right)=x_{0}$ is given by

$$x\left(t\right)=\Phi\left(t,t_{0}\right)x_{0}+\int_{t_{0}}^{t}\Phi\left(t,\tau\right)\begin{bmatrix}0\\ 1\end{bmatrix}Z\left(\tau\right)d\tau, \tag{11}$$

where $\Phi\left(t,t_{o}\right)$ is the transition matrix. The transition matrix is defined as the solution Φ of the matrix differential equation (Brockett, 1970, Chapter 1.3)

$$\begin{aligned}\frac{d}{dt}\Phi\left(t,t_{o}\right) &= A\left(t\right)\Phi\left(t,t_{o}\right)\\ \Phi\left(t_{0},t_{o}\right) &= I,\end{aligned}$$

where I is the identity matrix.

For the problem at hand, I have

$$\dot{x}\left(t\right)=A\left(t\right)x\left(t\right),$$

or

$$\dot{x}\left(t\right)=\begin{bmatrix}0 & 1\\ \frac{1}{k_{3}t^{2}\left(1-t\right)^{2}} & 0\end{bmatrix}x\left(t\right).$$

Moreover let

$$\begin{aligned}x_{0} &= \begin{bmatrix}x_{10} & x_{20}\end{bmatrix}^{\prime}\\ x_{10} &= y\left(t_{0}\right)\\ x_{20} &= \dot{y}\left(t_{0}\right).\end{aligned}$$

I know that the solution to the homogeneous differential equation

$$\ddot{y}(t) = \frac{1}{k_3 t^2 (1-t)^2} y(t),$$

is given by

$$y(t) = c_1 t^q (1-t)^{1-q} + c_2 t^{1-q} (1-t)^q,$$

where $q = \frac{1}{2}\left(1 - \sqrt{\frac{4+k_3}{k_3}}\right)$ and c_1 and c_2 are undetermined coefficients.

Given the solution to the homogeneous differential equation, I can write

$$\Phi(t, t_0) x_0 = \begin{bmatrix} t^q (1-t)^{1-q}, & t^{1-q} (1-t)^q \\ \frac{d}{dt}\left(t^q (1-t)^{1-q}\right), & \frac{d}{dt}\left(t^{1-q} (1-t)^q\right) \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

or

$$\Phi(t, t_0) x_0 = \begin{bmatrix} t^q (1-t)^{1-q}, & t^{1-q} (1-t)^q \\ t^{q-1} (1-t)^{-q} (q-t), & t^{-q} (1-t)^{q-1} (1-q-t) \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \tag{12}$$

where c_1 and c_2 are undetermined coefficients.

I next solve for c_1 and c_2 as a function of x_0.

At $t = t_0$, since $\Phi(t_0, t_0) = I$, I have

$$\begin{aligned} \begin{bmatrix} x_{10} \\ x_{20} \end{bmatrix} &= \begin{bmatrix} t_0^q (1-t_0)^{1-q}, & t_0^{1-q} (1-t_0)^q \\ t_0^{q-1} (1-t_0)^{-q} (q-t_0), & t_0^{-q} (1-t_0)^{q-1} (1-q-t_0) \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \Leftrightarrow \\ \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \frac{1}{1-2q} \begin{bmatrix} t_0^{-q} (1-t_0)^{q-1} (1-q-t_0), & -t_0^{1-q} (1-t_0)^q \\ -t_0^{q-1} (1-t_0)^{-q} (q-t_0), & t_0^q (1-t_0)^{1-q} \end{bmatrix} \begin{bmatrix} x_{10} \\ x_{20} \end{bmatrix}. \end{aligned}$$

Therefore, substituting in for $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, in (12) leads to

$$\begin{aligned} \Phi(t, t_0) x_0 &= \begin{bmatrix} t^q (1-t)^{1-q}, & t^{1-q} (1-t)^q \\ t^{q-1} (1-t)^{-q} (q-t), & t^{-q} (1-t)^{q-1} (1-q-t) \end{bmatrix} \\ &\quad \frac{1}{1-2q} \begin{bmatrix} t_0^{-q} (1-t_0)^{q-1} (1-q-t_0), & -t_0^{1-q} (1-t_0)^q \\ -t_0^{q-1} (1-t_0)^{-q} (q-t_0), & t_0^q (1-t_0)^{1-q} \end{bmatrix} \begin{bmatrix} x_{10} \\ x_{20} \end{bmatrix} \end{aligned}$$

or

$$\Phi(t,t_0) = \frac{1}{1-2q}\begin{bmatrix} t^q(1-t)^{1-q}, & t^{1-q}(1-t)^q \\ t^{q-1}(1-t)^{-q}(q-t), & t^{-q}(1-t)^{q-1}(1-q-t) \end{bmatrix} \begin{bmatrix} t_0^{-q}(1-t_0)^{q-1}(1-q-t_0), & -t_0^{1-q}(1-t_0)^q \\ -t_0^{q-1}(1-t_0)^{-q}(q-t_0), & t_0^q(1-t_0)^{1-q} \end{bmatrix}, \tag{13}$$

since $x_0 = \begin{bmatrix} x_{10} & x_{20} \end{bmatrix}'$ cancels on both sides of the equation.

In other words, the solution to the homogeneous is given by

$$x(t) = \Phi(t,t_0)\, x_0, \tag{14}$$

where $\Phi(t,t_0)$ is defined in (13) above.

From the value matching and smooth pasting conditions, (8) and (9) above, I know that for $t_0 = \widetilde{p}^{\text{in}}+$

$$y(t_0) = W$$

$$\dot{y}(t_0) = 0.$$

Changing the notation we have

$$x_0 = \begin{bmatrix} x_{10} & x_{20} \end{bmatrix}' \tag{15}$$

$$x_{10} = y(t_0) = W \tag{16}$$

$$x_{20} = \dot{y}(t_0) = 0. \tag{17}$$

Then the solution to the homogeneous, (14), becomes

$$x(t) = \Phi(t,t_0)\begin{bmatrix} W \\ 0 \end{bmatrix}.$$

Using the variation of constants formula, equation (11), I can now solve the original differential equation, by substituting in for $\Phi(t,t_0)$ and $\Phi(t,\tau)$ from (13) and for x_0 from (15) through (17). We have:

$$
\begin{aligned}
\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} &= \frac{1}{1-2q} \begin{bmatrix} t^q (1-t)^{1-q}, & t^{1-q}(1-t)^q \\ t^{q-1}(1-t)^{-q}(q-t), & t^{-q}(1-t)^{q-1}(1-q-t) \end{bmatrix} \\
& \begin{bmatrix} t_0^{-q}(1-t_0)^{q-1}(1-q-t_0), & -t_0^{1-q}(1-t_0)^q \\ -t_0^{q-1}(1-t_0)^{-q}(q-t_0), & t_0^q(1-t_0)^{1-q} \end{bmatrix} \begin{bmatrix} W \\ 0 \end{bmatrix} \\
& + \int_{t_0}^{t} \frac{1}{1-2q} \begin{bmatrix} t^q (1-t)^{1-q}, & t^{1-q}(1-t)^q \\ t^{q-1}(1-t)^{-q}(q-t), & t^{-q}(1-t)^{q-1}(1-q-t) \end{bmatrix} . \\
& \begin{bmatrix} \tau^{-q}(1-\tau)^{q-1}(1-q-\tau), & -\tau^{1-q}(1-\tau)^q \\ -\tau^{q-1}(1-\tau)^{-q}(q-\tau), & \tau^q(1-\tau)^{1-q} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} Z(\tau)\, d\tau,
\end{aligned}
$$

or equivalently

$$
\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \frac{1}{1-2q}. \tag{18}
$$

$$
\begin{bmatrix} t^q(1-t)^{1-q} t_0^{-q}(1-t_0)^{q-1}(1-q-t_0)W - t^{1-q}(1-t)^q t_0^{q-1}(1-t_0)^{-q}(q-t_0)W \\ t^{q-1}(1-t)^{-q}(q-t) t_0^{-q}(1-t_0)^{q-1}(1-q-t_0)W - t^{-q}(1-t)^{q-1}(1-q-t) t_0^{q-1}(1-t_0)^{-q}(q-t_0)W \end{bmatrix}
$$

$$
+\frac{1}{1-2q} \begin{bmatrix} t^q (1-t)^{1-q}, & t^{1-q}(1-t)^q \\ t^{q-1}(1-t)^{-q}(q-t), & t^{-q}(1-t)^{q-1}(1-q-t) \end{bmatrix} .
$$

$$
\begin{bmatrix} \int_{t_0}^{t} \frac{1}{k_3} \left[k_1\tau + g(W)\tau^\beta (1-\tau)^{1-\beta} + k_2\right] \left(-\tau^{-1-q}(1-\tau)^{q-2}\right) d\tau \\ \int_{t_0}^{t} \frac{1}{k_3} \left[k_1\tau + g(W)\tau^\beta (1-\tau)^{1-\beta} + k_2\right] \left(\tau^{q-2}(1-\tau)^{-1-q}\right) d\tau \end{bmatrix}.
$$

In other words, (18) implies that the solution, $x_1(t) = y(t)$, to the differential equation (10), is of the following form (with no undermined coefficients)

$$
\begin{aligned}
y(t) &= \frac{1}{1-2q} \left[t^q(1-t)^{1-q} t_0^{-q}(1-t_0)^{q-1}(1-q-t_0)W - t^{1-q}(1-t)^q t_0^{q-1}(1-t_0)^{-q}(q-t_0)W \right] \\
&+ \frac{1}{k_3(1-2q)} t^q (1-t)^{1-q} \left[\int_{t_0}^{t} k_1\tau + g(W)\tau^\beta(1-\tau)^{1-\beta} + k_2 \right] \left(-\tau^{-1-q}(1-\tau)^{q-2}\right) d\tau \\
&+ \frac{1}{k_3(1-2q)} t^{1-q}(1-t)^q \int_{t_0}^{t} \left[k_1\tau + g(W)\tau^\beta(1-\tau)^{1-\beta} + k_2\right] \left(\tau^{q-2}(1-\tau)^{-1-q}\right) d\tau,
\end{aligned}
$$

or equivalently

$$
\begin{aligned}
y(t) \quad = \quad & \frac{1}{1-2q} t^{q} (1-t)^{1-q} [t_0^{-q} (1-t_0)^{q-1} (1-q-t_0) W \\
& + \frac{1}{k_3} (-k_1 \int_{t_0}^{t} \tau^{-q} (1-\tau)^{q-2} \, d\tau \\
& - g(W) \int_{t_0}^{t} \tau^{\beta-1-q} (1-\tau)^{-\beta+q-1} \, d\tau - k_2 \int_{t_0}^{t} \tau^{-1-q} (1-\tau)^{q-2} \, d\tau)] \\
& + \frac{1}{1-2q} t^{1-q} (1-t)^{q} [t_0^{q-1} (1-t_0)^{-q} (t_0 - q) W \\
& + \frac{1}{k_3} (k_1 \int_{t_0}^{t} \tau^{q-1} (1-\tau)^{-1-q} \, d\tau \\
& + g(W) \int_{t_0}^{t} \tau^{q+\beta-2} (1-\tau)^{-\beta-q} \, d\tau + k_2 \int_{t_0}^{t} \tau^{q-2} (1-\tau)^{-1-q} \, d\tau)],
\end{aligned} \tag{19}
$$

where $t_0 \equiv \widetilde{p}^{\text{in}}$ and $q = \frac{1}{2}\left(1 - \sqrt{\frac{4+k_3}{k_3}}\right)$.

But using integration by substitution (changing the variable in each of the integrals to $\frac{1}{\tau}$), the six integrals above are equal to

$$
\int_{t_0}^{t} \tau^{-q} (1-\tau)^{q-2} \, d\tau = \frac{1}{q-1} \left(\left(\frac{1}{t_0} - 1 \right)^{q-1} - \left(\frac{1}{t} - 1 \right)^{q-1} \right)
$$

$$
\int_{t_0}^{t} \tau^{\beta-1-q} (1-\tau)^{-\beta+q-1} \, d\tau = \frac{1}{q-\beta} \left(\left(\frac{1}{t_0} - 1 \right)^{q-\beta} - \left(\frac{1}{t} - 1 \right)^{q-\beta} \right)
$$

$$
\int_{t_0}^{t} \tau^{-1-q} (1-\tau)^{q-2} \, d\tau = \frac{1}{q(q-1)} \left(\left(\frac{1}{t_0} - 1 \right)^{q-1} \left(q\frac{1}{t_0} - \frac{1}{t_0} + 1 \right) - \left(\frac{1}{t} - 1 \right)^{q-1} \left(q\frac{1}{t} - \frac{1}{t} + 1 \right) \right)
$$

$$
\int_{t_0}^{t} \tau^{q-1} (1-\tau)^{-1-q} \, d\tau = \frac{1}{q} \left(\left(\frac{1}{t} - 1 \right)^{-q} - \left(\frac{1}{t_0} - 1 \right)^{-q} \right)
$$

$$
\int_{t_0}^{t} \tau^{q+\beta-2} (1-\tau)^{-\beta-q} \, d\tau = \frac{1}{\beta+q-1} \left(\left(\frac{1}{t} - 1 \right)^{1-q-\beta} - \left(\frac{1}{t_0} - 1 \right)^{1-q-\beta} \right)
$$

$$
\int_{t_0}^{t} \tau^{q-2} (1-\tau)^{-1-q} \, d\tau = \frac{1}{q(q-1)} \left(\left(\frac{1}{t} - 1 \right)^{-q} \left(q\frac{1}{t} - 1 \right) - \left(\frac{1}{t_0} - 1 \right)^{-q} \left(q\frac{1}{t_0} - 1 \right) \right).
$$

So the solution (19) above becomes

$$
\begin{aligned}
y\left(t\right) \quad = \quad & \frac{1}{1-2q}t^{q}\left(1-t\right)^{1-q}t_{0}^{-q}\left(1-t_{0}\right)^{q-1}\left(1-q-t_{0}\right)W \\
& +\frac{1}{1-2q}\frac{1}{k_{3}}t^{q}\left(1-t\right)^{1-q}[-k_{1}\frac{1}{q-1}\left(\left(\frac{1}{t_{0}}-1\right)^{q-1}-\left(\frac{1}{t}-1\right)^{q-1}\right) \\
& -g\left(W\right)\frac{1}{q-\beta}\left(\left(\frac{1}{t_{0}}-1\right)^{q-\beta}-\left(\frac{1}{t}-1\right)^{q-\beta}\right) \\
& -k_{2}\frac{1}{q\left(q-1\right)}\left(\left(\frac{1}{t_{0}}-1\right)^{q-1}\left(q\frac{1}{t_{0}}-\frac{1}{t_{0}}+1\right)-\left(\frac{1}{t}-1\right)^{q-1}\left(q\frac{1}{t}-\frac{1}{t}+1\right)\right)] \\
& +\frac{1}{1-2q}t^{1-q}\left(1-t\right)^{q}t_{0}^{q-1}\left(1-t_{0}\right)^{-q}\left(t_{0}-q\right)W \\
& +\frac{1}{1-2q}\frac{1}{k_{3}}t^{1-q}\left(1-t\right)^{q}[k_{1}\frac{1}{q}\left(\left(\frac{1}{t}-1\right)^{-q}-\left(\frac{1}{t_{0}}-1\right)^{-q}\right) \\
& +g\left(W\right)\frac{1}{\beta+q-1}\left(\left(\frac{1}{t}-1\right)^{1-q-\beta}-\left(\frac{1}{t_{0}}-1\right)^{1-q-\beta}\right) \\
& +k_{2}\frac{1}{q\left(q-1\right)}\left(\left(\frac{1}{t}-1\right)^{-q}\left(q\frac{1}{t}-1\right)-\left(\frac{1}{t_{0}}-1\right)^{-q}\left(q\frac{1}{t_{0}}-1\right)\right)].
\end{aligned}
$$

Finally I still need to pin down the value of t_0. I consider the case where $p \to 1$.

For an experienced worker I know from equation (2) that

$$V^{\exp}\left(1,W\right)=\frac{\alpha_{G}+u+\delta J}{r+\delta}. \tag{20}$$

Similarly for an inexperienced worker, from equation (7) I have

$$\left(r+\delta+\theta\right)V^{\text{in}}\left(1,W\right)=\alpha_{G}+\delta J+\theta V^{\exp}\left(1,W\right). \tag{21}$$

Substituting in for the value of an experienced worker from (20), equation (21) becomes

$$V^{\text{in}}\left(1,W\right)=\frac{\alpha_{G}+\delta J}{r+\delta+\theta}+\frac{\theta}{r+\delta+\theta}\frac{\alpha_{G}+u+\delta J}{r+\delta}.$$

Put differently, using the notation above, I have

$$\lim_{t\to 1}y\left(t\right)=\frac{\alpha_{G}+\delta J}{r+\delta+\theta}+\frac{\theta}{r+\delta+\theta}\frac{\alpha_{G}+u+\delta J}{r+\delta}. \tag{22}$$

The above equation pins down t_0 (i.e. $\widetilde{p}^{\text{in}}$) for a a given W.

As before the Gittins index of an occupation is the highest retirement value at which the worker is indifferent between working at that occupation or retiring with $W = W^{\text{in}}(p)$, i.e.

$$W^{\text{in}}(p) = V^{\text{in}}(p, W), \tag{23}$$

where $W^{\text{in}}(p) = \max\left\{\widetilde{W}\right\}$ and the set $\left\{\widetilde{W}\right\}$ includes all possible retirement values, $\widetilde{W}$, such that $\widetilde{W} = V^{\text{in}}\left(p, \widetilde{W}\right)$.

For eq. (23) to hold, it must be the case that

$$p = \widetilde{p}^{\text{in}}(W). \tag{24}$$

In practice, for a given value of beliefs, p, and a given value of moving, J, I use equation (22) to solve out for the associated Gittins index of the inexperienced worker, $W^{\text{in}}(p)$. In particular, substitute p for t_0 and solve (22) for W, while letting t got to 1.

Gittins Index for Moving

I also allow workers the option of moving to another city, which provides known value to the worker, J. As in the model of the main text, J is the retirement value associated with the option of moving and therefore corresponds to its Gittins index. A worker therefore moves, if and only if the retirement value (Gittins index) of all other arms is lower than J. In practice, the value of moving, J, can be computed numerically by simulating the worker's problem many times and computing the mean present discounted value of a worker who has just moved.

2 Additional Results on City Size and Number of Occupations

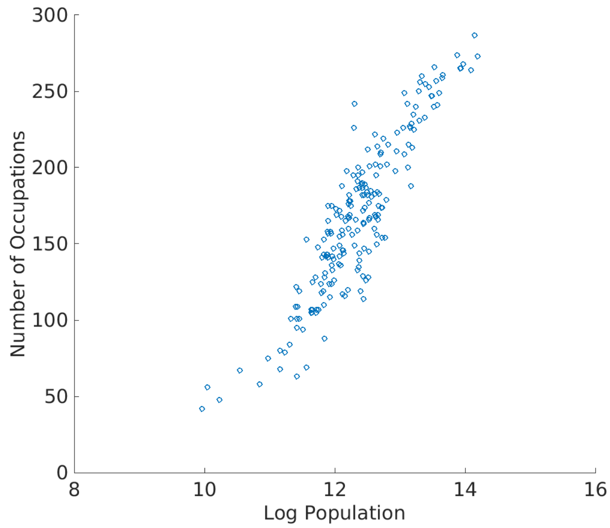

Figure 1: Number of Occupations vs. Log County Population - JobCentrePlus Vacancy Data, UK, June through September 2012

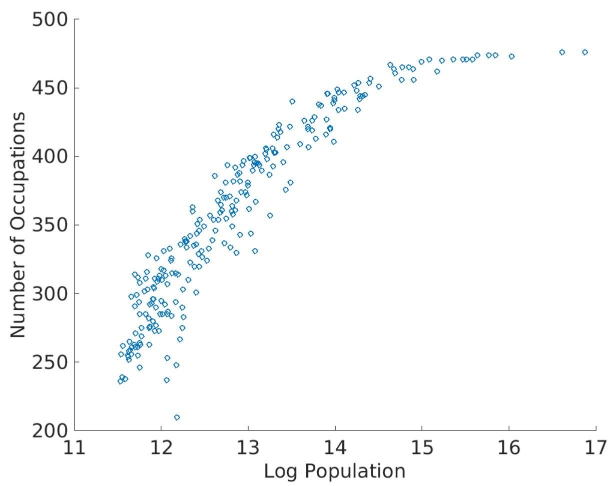

Figure 2: Number of Occupations vs. Log MSA Population - 2000 Census Data (5%)

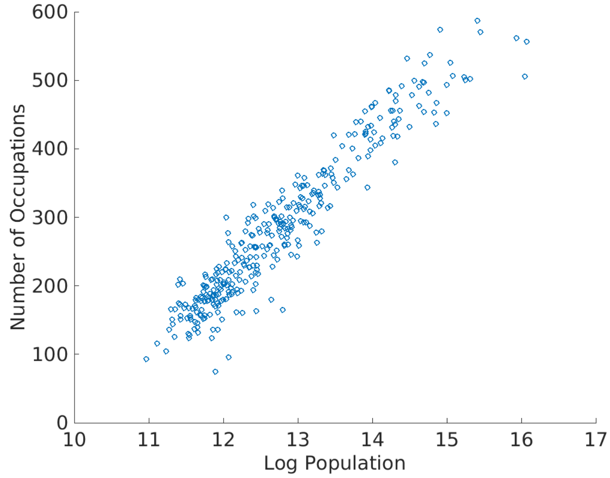

Figure 3: Number of Occupations vs. Log MSA Population - 2000 Occupational Employment Statistics

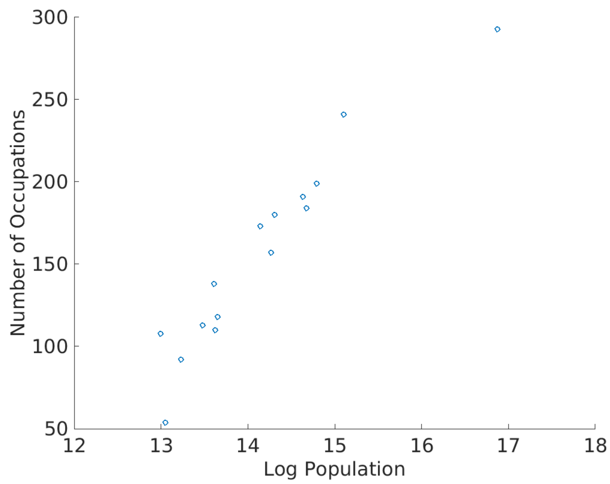

Figure 4: Number of Occupations vs. Log City Population - Brazilian RAIS data. Males with a College Degree in the State of São Paulo, 1995-2005

3 Full Set of Controls of Mincer Regression

In Table 1, I present the full set of the Mincerian controls for the wage premium regression. Their signs and magnitudes are consistent with the prior literature.

	Full Sample
	ln(wage)
ln(current city pop)	0.041
	(0.001)
Age	-0.033
	(0.008)
Age^2	0.002
	(0.0003)
Age^3	-0.00004
	(0.0000054)
Age^4	0.000000218
	(0.0000000316)
Female dummy	-0.15
	(0.005)
High School dummy	0.132
	(0.005)
College dummy	0.248
	(0.01)
Married dummy	0.073
	(0.005)
Non-White dummy	-0.058
	(0.005)
Firm>100 employees	0.131
	(0.005)
Firm 25 - 99 employees	0.03
	(0.006)
constant	1.916
	(0.127)
1-Digit Industry & Occup fixed effects	Yes
Number of Obs	169536

Table 1: Wage Premium. Source: 1996 Panel of Survey of Income and Program Participation. Population data from 2000 Census. Controls include 11 industry dummies and 13 occupation dummies. Standard errors in parentheses are clustered by individual.

References

BERGEMANN, D. AND J. VALIMAKI (2008): "Bandit Problems," in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf and L. E. Blume, Macmillan.

BROCKETT, R. W. (1970): *Finite Dimensional Linear Systems*, New York: Wiley.

GITTINS, J. C. (1979): "Bandit Processes and Dynamic Allocation Indices," *Journal of the Royal Statistical Society. Series B (Methodological)*, 41, 148–177.

Kambourov, G. and I. Manovskii (2009): "Occupational Mobility and Wage Inequality," *Review of Economic Studies*, 76, 731–759.

Karatzas, I. (1984): "Gittins Indices in the Dynamic Allocation Problem for Diffusion Processes," *The Annals of Probability*, 12, 173–192.

Whittle, P. (1982): *Optimization over Time: Programming and Stochastic Control*, New York: Wiley.