14.461 Advanced Macroeconomics I:
Part 1: Search Theory

Daron Acemoglu

October 2005

1 Review of Dynamic Programming

This is a very quick review of some key aspects of dynamic programming, especially those useful in the context of search models. The notes here heavily borrow from Stokey, Lucas and Prescott (1989), but simplify the exposition a little and emphasize the results useful for search theory.

1.1 Basic Idea of Dynamic Programming

Most models in macroeconomics, and more specifically most models we will see in the macroeconomic analysis of labor markets, will be dynamic, either in discrete or in continuous time. Either formulated as a social planner's problem or formulated as an equilibrium problem, with each agent maximizing their utility or profits, these models will involve a dynamic optimization problem. Using abstract but simple notation, the canonical dynamic optimization program can be written as

$$\begin{aligned}
\textbf{Problem A1} &: \\
v^*(x_0) &= \sup_{\{x_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1}) \\
&\quad \text{subject to} \\
x_{t+1} &\in \Gamma(x_t), \qquad \text{for all } t \geq 0 \\
&\quad x_0 \text{ given.}
\end{aligned}$$

where $x_t \in X \subset \mathbb{R}^K$ for some $K \geq 1$. In many economic applications, we will have $K = 1$, so that $x_t \in \mathbb{R}$.

Here F is the payoff function, depending on x_t, which is the **state variable**, and x_{t+1}, which corresponds to the **control variable**. In this simple formulation, x_{t+1} will also directly become the state variable in the next time period.

The constraint on the problem is written as

$$x_{t+1} \in \Gamma(x_t)$$

where

$$\Gamma : X \rightrightarrows X$$

is a correspondence determining what type of x_{t+1} is allowed given the state variable x_t.

Problem A1, also referred to as the **sequence problem**, is one of choosing an infinite sequence $\{x_t\}_{t=0}^{\infty}$ from some (vector) space of infinite sequences (for example, $\{x_t\}_{t=0}^{\infty} \in \mathcal{L}^{\infty}$, where $\mathcal{L}^{\infty}$ is the vector space of infinite sequences that are bounded with the $\|\cdot\|_\infty$ norm, which I will denote throughout by the simpler notation $\|\cdot\|$). Such problems sometimes have nice features, but often are difficult to characterize both analytically and numerically.

The basic idea of dynamic programming is to turn the sequence problem into a functional equation, i.e., one of finding a function rather than a sequence. This often gives better economic insights, similar to the logic of comparing today to tomorrow. It is also often easier to characterize analytically or numerically. In this particular case, the relevant functional equation can be written as

Problem A2 :

$$v(x) = \sup_{y \in \Gamma(x)} \left[F(x,y) + \beta v(y)\right], \text{ for all } x \in X. \quad (1)$$

Part of the theory of dynamic programming is about specifying the conditions under which Problems A1 and A2 are equivalent. These are not central for us to focus upon here, but I will return to some of these issues below. This equation is commonly referred to as the **Bellman equation,** after Richard Bellman, who introduced dynamic programming to operations research and engineering applications (though identical tools and reasonings, including the contraction mapping theorem were earlier used by Lloyd Shapley in his work on stochastic games).

A couple of points are immediately worth noting. First, $v(x)$ is a function, more formally,

$$v : X \to \mathbb{R}$$

Differently from other maximization problems, here maximization itself defines the function v, as the notation makes it clear with the sup (or max) defining the function. Therefore, instead of finding a sequence $\{x_t\}_{t=0}^{\infty} \in \mathcal{L}^{\infty}$, we will try to find a function v, that satisfies (1). Second, because the function v is defined **recursively**, in the sense that it is on the right hand side of (1) as well, this is often referred to as the recursive formulation.

What makes this formulation useful is that the solution will often be a time invariant **policy function**, $g : X \to X$ determining what value of x_{t+1} to choose for a given value of the state variable x_t. [In general, there are two complications: first, a control reaching the optimal value may not exist, which was the reason why we originally used the notation sup; second, we may not have a policy function, but the policy correspondence $g : X \rightrightarrows X$, because there are more than one maximizers for a given state variable. Let me avoid these complications for now]. Moreover, as we will see, once the value function v is determined, the policy function is given straightforwardly. In particular, by definition it must be the case that

$$v(x) = [F(x, g(x)) + \beta v(g(x))], \text{ for all } x \in X,$$

which is one way of determining the policy function. This equation simply follows from the fact that $g(x)$ is the optimal policy, so reaches the maximal value $v(x)$.

The usefulness of the recursive formulation as in (1) comes from the fact that there are some powerful tools which not only establish existence of the solution, but also some of its properties. Let us briefly look at these.

1.2 Contraction Mappings

We say that (S, ρ) is a metric space, if S is a space and ρ is a metric defined over this space with the usual properties (loosely corresponding to "distance" between elements of S).

Definition 1 *Let (S, ρ) be a metric space and $T : S \to S$ be an operator mapping S into itself. T is a **contraction mapping** (with **modulus** β) if for some $\beta \in (0, 1)$,*

$$\rho(Tx, Ty) \leq \beta\rho(x, y), \text{ for all } x, y \in S.$$

In other words, a contraction mapping brings elements of the space S "closer" to each other.

For example, let us take a simple interval of the real line as our space, $S = [a, b]$, with usual metric of this space $\rho(x, y) = |x - y|$. Then $T : S \to S$ is a contraction if for some $\beta \in (0, 1)$,

$$\frac{|Tx - Ty|}{|x - y|} \leq \beta < 1, \quad \text{all } x, y \in S \text{ with } x \neq y.$$

Definition 2 *A **fixed point** of T is any element of S satisfying $Tx = x$.*

Recall also that a metric space (S, ρ) is complete if every Cauchy sequence in S converges to an element in S.

Theorem 1 ***(Contraction Mapping Theorem)*** *Let (S,ρ) be a complete metric space, and $T : S \to S$ be a contraction. Then there exists a unique $\hat{v} \in S$ such that*

$$T\hat{\nu} = \hat{\nu},$$

i.e., a unique fixed point.

Proof. Note $T^n x = T(T^{n-1}x)$ for any $n = 1, 2, ...$ Now take $\nu_0 \in S$, and a sequence $\{\nu_n\}_{n=0}^{\infty}$ with each element in S, such that $\nu_{n+1} = T\nu_n$ so that

$$\nu_n = T^n \nu_0.$$

This implies that

$$\rho(\nu_2,\nu_1) = \rho(T\nu_1, T\nu_0) \leq \beta\rho(\nu_1,\nu_0),$$

where the last inequality uses the contraction property of T. Moreover, by induction, we have

$$\rho(\nu_{n+1},\nu_n) \leq \beta^n \rho(\nu_1,\nu_0), \quad n = 1, 2, ... \tag{2}$$

Hence, for any $m > n$,

$$\begin{aligned}
\rho(\nu_m,\nu_n) &\leq \rho(\nu_m,\nu_{m-1}) + ... + \rho(\nu_{n+2},\nu_{n+1}) + \rho(\nu_{n+1},\nu_n) \\
&\leq \left[\beta^{m-1} + ... + \beta^{n+1} + \beta^n\right] \rho(\nu_1,\nu_0) \\
&= \beta^n \left[\beta^{m-n-1} + ... + \beta + 1\right] \rho(\nu_1,\nu_0) \\
&\leq \frac{\beta^n}{1-\beta}\rho(\nu_1,\nu_0),
\end{aligned}$$

where the first line uses the triangle inequality (which is true by definition for any metric), and the second line uses (2).

The last line implies that as $n, m \to \infty$, ν_m and ν_n are getting closer, so $\{\nu_n\}_{n=0}^{\infty}$ is a Cauchy sequence. Since S is complete, this establishes that

$$\nu_n \to \hat{\nu} \in S.$$

Now note that for any $\nu_0 \in S$ and any $n \in \mathbb{N}$, we have

$$\begin{aligned} \rho(T\hat{\nu}, \hat{\nu}) &\leq \rho(T\hat{\nu}, T^n \nu_0) + \rho(T^n \nu_0, \hat{\nu}) \\ &\leq \beta \rho(\hat{\nu}, T^{n-1} \nu_0) + \rho(T^n \nu_0, \hat{\nu}), \end{aligned}$$

where the first line again uses the triangle inequality, and the second line the definition of the contraction. The above argument shows that both of the terms on the right tend to zero as $n \to \infty$, which implies that $\rho(T\hat{\nu}, \hat{\nu}) = 0$, establishing that $T\hat{\nu} = \hat{\nu}$, thus a fixed point exists.

Uniqueness is proved by contradiction. Suppose that there exist $\hat{\nu}, \nu \in S$, such that $T\nu = \nu$ and $T\hat{\nu} = \hat{\nu}$ with $\hat{\nu} \neq \nu$. This implies

$$0 < a = \rho(\hat{\nu}, \nu) = \rho(T\hat{\nu}, T\nu) \leq \beta \rho(\hat{\nu}, \nu) = \beta a.$$

Since $\beta < 1$, this yields a contradiction, proving uniqueness. ■

The use of the contraction mapping theorem is that it can be applied to any metric space, so in particular to the space of functions. Applying it to equation (1) will establish the existence of a unique value function v, greatly facilitating the analysis of such dynamic models. Naturally, for this we have to prove that the recursion in (1) defines a contraction mapping. We will see below that this is often straightforward.

Before doing this, let us consider another useful result.

First, recall that if (S,ρ) is a complete metric space and S' is a closed subset of S, then (S',ρ) is also a complete metric space.

Theorem 2 *Let (S,ρ) be a complete metric space, $T : S \to S$ be a contraction mapping with $T\hat{\nu} = \hat{\nu}$. If S' is a closed subset of S, and $T(S') \subseteq S'$, then $\hat{\nu} \in S'$. Moreover, if $T(S') \subseteq S'' \subseteq S'$, then $\hat{\nu} \in S''$.*

Proof. Take an arbitrary $\nu_0 \in S'$, and consider the sequence $\{T^n \nu_0\}_{n=0}^{\infty}$. Each element of this sequence is in S' by the fact that $T(S') \subseteq S'$. $T^n\nu_0 \to \hat{\nu}$ from Theorem 1. Since S' is closed, $\hat{\nu} \in S'$, proving the first claim in the theorem. If in addition we have that $T(S') \subseteq S''$, then by virtue of the fact that $\hat{\nu} \in S'$. $T\hat{\nu} \in S''$, so $\hat{\nu} \in S''$, proving the second part. ■

The second part of this theorem is very important to prove results such as strict concavity or that a function is strictly increasing. This is because the set of strictly concave functions or the strictly increasing functions are not closed. The second part of the theorem enables us to avoid this complication.

How do we check that a mapping is a contraction? Here, the following theorem is useful, especially in the context of dynamic programming. Let us use the notation $(f+a)(x) = f(x) + a$ for some $a \in \mathbb{R}$. Then:

Theorem 3 ***(Blackwell's sufficient conditions for a contraction)*** *Let $X \subseteq \mathbb{R}^K$, and $B(X)$ be the space of bounded functions $f : X \to \mathbb{R}$. defined* *on X Suppose that $T : B(X) \to B(X)$ is an operator satisfying the following two conditions:*

1. ***(monotonicity)*** *For any $f, g \in B(X)$ and $f(x) \leq g(x)$ for all $x \in X$ implies $(Tf)(x) \leq (Tg)(x)$ for all $x \in X$.*

2. ***(discounting)*** *There exists $\beta \in (0,1)$ such that*

$$[T(f + a)](x) \leq (Tf)(x) + \beta a, \quad \text{for all } f \in B(X), a \geq 0, \ x \in X,$$

Then, T is a contraction with modulus β.

Proof. Let $f \leq g$ stand for $f(x) \leq g(x)$ for all $x \in X$. By definition

$$\text{for any } f, g \in B(X), \ \ f \leq g + \|f - g\|,$$

where again $\|\cdot\|$ is the sup norm. Now applying the operator T on both sides, we have

$$Tf \leq T(g + \|f - g\|) \leq Tg + \beta \|f - g\|,$$

where the first inequality uses monotonicity and the second discounting. Applying the same argument in reverse establishes

$$Tg \leq Tf + \beta \|f - g\|.$$

Combining these two inequalities yields

$$\|Tf - Tg\| \leq \beta \|f - g\|,$$

proving that T is a contraction. ■

1.3 Application to Dynamic Programming

Let us now apply the above tools to the problem of dynamic programming, outlined at the beginning. Consider a sequence $\{x_{t+1}\}_{t=0}^{\infty}$ which attains the supremum of Problem A1. We will now show that this sequence will satisfy the recursive equation of dynamic programming

$$v(x_t) = F(x_t, x_{t+1}) + \beta v(x_{t+1}), \text{ for all } t = 0, 1, 2, ..., \tag{3}$$

and moreover, under some boundedness conditions, any sequence that is a solution to (3) is a solution to Problem A1, in the sense that it attains its supremum. In other words, we will establish some equivalence results between the solutions to Problem A1 and Problem A2.

To prepare for these results, let us define the set of feasible sequences or **plans** starting with initial value x_0:

$$\Pi(x_0) = \{\{x_{t+1}\}_{t=0}^{\infty} : x_{t+1} \in \Gamma(x_t), \quad t = 0, 1, ...\}.$$

Let us denote a typical element of the set by $\underline{x} = (x_0, x_1, ...) \in \Pi(x_0)$, and assume:

Assumption 1 $\Gamma(x)$ is nonempty for all $x \in X$; and for all $x_0 \in X$ and $\underline{x} \in \Pi(x_0)$, $\lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1})$ exists.

Next define the supremum function $v^* : X \to \bar{\mathbb{R}}$, where $\bar{\mathbb{R}}$ is the extended real line ($\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$), as:

$$v^*(x_0) = \sup_{\underline{x}\in\Pi(x_0)} u(\underline{x}).$$

Thus $v^*(x_0)$ is the supremum in Problem A1 (i.e., the value of the program in Problem A1). Note that it follows by definition that v^* is the unique function satisfying the following three conditions for Problem A1, or the sequence problem, **SP**:

1. if $|v^*(x_0)| < \infty$, then

$$v^*(x_0) \geq u(\underline{x}), \quad \text{all } \underline{x} \in \Pi(x_0); \tag{4}$$

   and for any $\varepsilon > 0$,

$$v^*(x_0) \leq u(\underline{x}) + \varepsilon, \quad \text{some } \underline{x} \in \Pi(x_0); \tag{5}$$

2. if $v^*(x_0) = +\infty$, then there exists a sequence $\{\underline{x}^k\}$ in $\Pi(x_0)$ such that $\lim_{k\to\infty} u(\underline{x}^k) = +\infty$; and

3. if $v^*(x_0) = -\infty$, then $u(\underline{x}) = -\infty,$, for all $\underline{x} \in \Pi(x_0)$.

Conversely, we will say that v^* is a solution to Problem A2 (and thus satisfies the functional equation (3)), if the following three conditions for **FE** hold:

1. If $|v^*(x_0)| < \infty$, then

$$v^*(x_0) \geq F(x_0, y) + \beta v^*(y), \quad \text{all } y \in \Gamma(x_0), \tag{6}$$

   and for any $\varepsilon > 0$,

$$v^*(x_0) \leq F(x_0, y) + \beta v^*(y) + \varepsilon, \quad \text{some } y \in \Gamma(x_0); \tag{7}$$

2. if $v^*(x_0) = +\infty$, then there exists a sequence $\{y^k\}$ in $\Gamma(x_0)$ such that

$$\lim_{k\to\infty} \left[F(x_0, y^k) + \beta v^*(y^k)\right] = +\infty; \tag{8}$$

3. if $v^*(x_0) = -\infty$, then

$$F(x_0, y) + \beta v^*(y) = -\infty, \quad \text{all } y \in \Gamma(x_0). \tag{9}$$

We now have the following simple lemma:

Lemma 1 *Let X, Γ, F, and β satisfy Assumption 1. Then for any $x_0 \in X$ and any $\underline{x} = (x_0, x_1, ...) \in \Pi(x_0)$,*

$$u(\underline{x}) = F(x_0, x_1) + \beta u(\underline{x}')$$

with $\underline{x}' = (x_1, x_2, ...)$.

Proof. Under Assumption 1, for any $x_0 \in X$ and any $\underline{x} \in \Pi(x_0)$,

$$\begin{aligned} u(\underline{x}) &= \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1}) \\ &= F(x_0, x_1) + \beta \lim_{n\to\infty} \sum_{t=0}^{n} \beta^t F(x_{t+1}, x_{t+2}) \\ &= F(x_0, x_1) + \beta u(\underline{x}'). \end{aligned}$$

■

This lemma basically says that the utility from any feasible plan can be decomposed into two parts, the current return and continuation value.

Theorem 4 *Let X, Γ, F, and β satisfy Assumption 1. Then the function v^* is a solution to Problem A2.*

Proof. If $\beta = 0$, the result is trivial. Suppose that $\beta > 0$, and choose $x_0 \in X$.

Suppose $v^*(x_0)$ is finite. Then SP conditions (4) and (5) hold, and it is sufficient to show that this implies that the FE conditions (6) and (7) hold. To establish (6), let $x_1 \in \Gamma(x_0)$ and $\varepsilon > 0$ be given. Then by SP (5) there exists $\underline{x}' = (x_1, x_2, ...) \in \Pi(x_1)$ such that $u(\underline{x}') \geq v^*(x_1) - \varepsilon$. Note also that $\underline{x} = (x_0, x_1, x_2, ...) \in \Pi(x_0)$. Hence it follows from SP (4) and Lemma 1 that

$$v^*(x_0) \geq u(\underline{x}) = F(x_0, x_1) + \beta u(\underline{x}') \geq F(x_0, x_1) + \beta v^*(x_1) - \beta\varepsilon$$

for any $\varepsilon > 0$, establishing FE (6).

To establish FE (7), choose $x_0 \in X$ and $\varepsilon > 0$. From SP (5) and Lemma 1, it follows that one can choose $\underline{x} = (x_0, x_1, ...) \in \Pi(x_0)$, so that

$$v^*(x_0) \leq u(\underline{x}) + \varepsilon = F(x_0, x_1) + \beta u(\underline{x}') + \varepsilon,$$

where $\underline{x}' = (x_1, x_2, ...)$. It then follows from SP (4) that

$$v^*(x_0) \leq F(x_0, x_1) + \beta v^*(x_1) + \varepsilon.$$

Since $x_1 \in \Gamma(x_0)$, this establishes FE (7).

If $v^*(x_0) = +\infty$, then there exists a sequence $\{\underline{x}^k\}$ in $\Pi(x_0)$ such that $\lim_{k\to\infty} u(\underline{x}^k) = +\infty$. Since $x_1^k \in \Gamma(x_0)$, all k, and

$$u(\underline{x}^k) = F(x_0, x_1^k) + \beta u(\underline{x}'^k) \leq F(x_0, x_1^k) + \beta v^*(x_1^k), \quad \text{all } k,$$

it follows that FE (8) holds for the sequence $\{y^k = x_1^k\}$ in $\Gamma(x_0)$. If $v^*(x_0) = -\infty$, then

$$u(\underline{x}) = F(x_0, x_1) + \beta u(\underline{x}') = -\infty, \quad \text{all } (x_0, x_1, x_2, ...) = \underline{x} \in \Pi(x_0),$$

where $\underline{x}' = (x_1, x_2, ...)$. Since F is real-valued (thus does not take the values $-\infty$ or $+\infty$), it follows that

$$u(\underline{x}') = -\infty, \quad \text{all } x_1 \in \Gamma(x_0), \text{ all } \underline{x}' \in \Pi(x_1).$$

Hence $v^*(x_1) = -\infty$, all $x_1 \in \Gamma(x_0)$. Since F is real-valued and $\beta > 0$, (9) follows immediately. ■

Under the additional boundedness condition, we have the following converse to this theorem:

Theorem 5 *Let X, Γ, F, and β satisfy Assumption 1. If v is a solution to (FE) and satisfies*

$$\lim_{n\to\infty} \beta^n v(x_n) = 0, \quad \textit{all } (x_0, x_1, ...) \in \Pi(x_0), \ \textit{all } x_0 \in X, \tag{10}$$

then $v = v^*$.

Proof. (sketch) Condition (10) implies that v cannot take on the values $+\infty$ or $-\infty$. Hence v satisfies (6) and (7), and it is sufficient to show that this implies v satisfies (4) and (5).

Since v is the solution to Problem A2, then (6) implies that for all $x_0 \in X$ and $\underline{x} \in \Pi(x_0)$

$$\begin{aligned} v(x_0) &\geq F(x_0, x_1) + \beta v(x_1) \\ &\geq F(x_0, x_1) + \beta F(x_1, x_2) + \beta^2 v(x_2) \\ &\vdots \\ &\geq u_n(\underline{x}) + \beta^{n+1} v(x_{n+1}). \end{aligned}$$

Now taking the limit as $n \to \infty$ and using the convergence property from (10), we obtain (4) for any $\underline{x} \in \Pi(x_0)$.

Now for a given $x_0 \in X$ and $\varepsilon > 0$, choose an arbitrary sequence $\{\delta_n\}_{n=1}^{\infty}$ in $\mathbb{R}_+$ such that $\sum_{n=1}^{\infty} \beta^{n-1}\delta_n \leq \varepsilon/2$. Since (7) holds, we can choose $x_{t+1} \in \Gamma(x_t)$ so that

$$v(x_t) \leq F(x_t, x_{t+1}) + \beta v(x_{t+1}) + \delta_{t+1}.$$

Using these inequalities, we obtain that for any $\underline{x} = (x_0, x_1, x_2, ...) \in \Pi(x_0)$, we have

$$\begin{aligned} v(x_0) &\leq u_n(\underline{x}) + \beta^{n+1} v(x_{n+1}) + (\delta_1 + \beta\delta_2 + ... + \beta^n \delta_{n+1}) \\ &\leq u_n(\underline{x}) + \beta^{n+1} v(x_{n+1}) + \varepsilon/2, \quad n = 1, 2, ... \end{aligned}$$

Since (10) implies that for n sufficiently large the second term is also less than $\varepsilon/2$, it follows that as $n \to \infty$,

$$v(x_0) \leq u(\underline{x}) + \varepsilon,$$

completing the proof. ■

An important implication is that although Problem A2 may have many solutions, only one of those will satisfy the convergence condition (10). In general, we can make a lot of progress by studying solutions to Problem A2, but sometimes we need to impose (10) in order to pick the right solution (this is similar to sometimes working with necessary conditions for optimization, though of course then we need to impose the sufficiency conditions).

Naturally, our interest is mainly with optimal plans. For this we have:

Theorem 6 *Let X, Γ, F, and β satisfy Assumption 1. Let $\underline{x}^* \in \Pi(x_0)$ be a feasible plan that attains the supremum in Problem A1 starting with initial state x_0. Then*

$$v^*(x_t^*) = F(x_t^*, x_{t+1}^*) + \beta v^*(x_{t+1}^*), \quad t = 0, 1, 2, ... \tag{11}$$

Proof. Since $\underline{x}^*$ attains the supremum,

$$\begin{aligned} v^*(x_0^*) &= u(\underline{x}^*) = F(x_0, x_1^*) + \beta u(x^{*\prime}) \\ &\geq u(\underline{x}) = F(x_0, x_1) + \beta u(\underline{x}'), \quad \text{all } \underline{x} \in \Pi(x_0). \end{aligned} \tag{12}$$

Now choose $x_1 = x_1^*$, (12) still holds. Since $(x_1^*, x_2, x_3, ...) \in \Pi(x_1^*)$ implies that $(x_0, x_1^*, x_2, x_3, ...) \in \Pi(x_0)$, so that

$$u(\underline{x}^{*\prime}) \geq u(\underline{x}'), \quad \text{all } \underline{x} \in \Pi(x_1^*).$$

Therefore $u(\underline{x}^{*\prime}) = v(x_1^*)$. Substituting this into (12) yields (11) for $t = 0$. Continuing by induction establishes (11) for all t. ■

Finally, the converse to this theorem is:

Theorem 7 *Let X, Γ, F, and β satisfy Assumption 1. Let $\underline{x}^* \in \Pi(x_0)$ be a feasible plan from x_0 satisfying (11), and with*

$$\limsup_{t \to \infty} \beta^t v^*(x_t^*) \leq 0. \tag{13}$$

Then $\underline{x}^*$ attains the supremum in Problem A1 for initial state x_0.

Proof. Suppose that $\underline{x}^* \in \Pi(x_0)$ satisfies (11) and (13). Then it follows by induction on (11) that

$$v^*(x_0) = u_n(\underline{x}^*) + \beta^{n+1} v^*(x_{n+1}^*), \quad n = 1, 2, ...$$

Then using (13), we find that $v^*(x_0) \leq u(\underline{x}^*)$. Since $\underline{x}^* \in \Pi(x_0)$, the reverse inequality holds, establishing the result. ■

The above theorems are useful in showing the equivalence of Problem A1 and Problem A2. Now the usefulness of the dynamic programming formulation in Problem A2, and hence of the contraction mapping theorem, comes from the fact that its solution is often easy to characterize. So for this purpose, take the following version of the dynamic programming problem (Problem A2)

$$v(x) = \max_{y \in \Gamma(x)} \left[F(x,y) + \beta v(y) \right], \tag{14}$$

where $\beta < 1$. As before, X is the possible set of values for the state variable and $\Gamma : X \rightrightarrows X$ is the correspondence describing the constraints on the problem. We now make an additional assumption, which is not necessary, but greatly simplifies the analysis.

Assumption 2 X is a compact subset of $\mathbb{R}^K$, Γ is nonempty, compact-valued and continuous. Moreover, let $A = \{(x,y) \in X \times X : y \in \Gamma(x)\}$ and $F : A \to \mathbb{R}$ be bounded and continuous.

The importance of Assumption 2 is that it will allow us to focus on the space of bounded functions. Most importantly, since F is bounded over its effective domain, there exists some $B < \infty$, such that $|F(x,y)| < B$ for all $(x,y) \in A$. This immediately implies that $|v^*(x)| \leq B/(1-\beta)$, all $x \in X$. Consequently, we can focus our attention on value functions in the space

$C(X)$ of continuous bounded functions defined on X, with the natural norm on this space, the sup norm, $\|f\| = \sup_{x \in X} |f(x)|$.

In particular, to see the usefulness of the contraction mapping theorem, now define the operator T such that

$$(Tf)(x) = \max_{y \in \Gamma(x)} [F(x, y) + \beta f(y)]. \tag{15}$$

A fixed point of this operator, $v = Tv$, will be a solution to (14), establishing the desired results. Then we can derive the policy functions from the value function.

Theorem 8 *Let X, Γ, F, and β satisfy Assumption 2 and let $C(X)$ be the space of bounded continuous functions $f : X \to \mathbb{R}$, with the sup norm. Then the operator T maps $C(X)$ into itself, i.e., $T : C(X) \to C(X)$, and has a unique fixed point, $v \in C(X)$ satisfying (14).*

Proof. Formulated in this way, it is immediate that T is a contraction. Since the maximization problem on the right hand side of (15) is one of maximizing a bounded function over a compact set, it has a solution. Consequently, T is well defined and is easily seen to satisfy the sufficient conditions for a contraction in Theorem 3. Therefore, applying Theorem 1, a unique $v \in C(X)$ satisfying (14) exists. ■

Corollary 1 *Let $G : X \to X$ defined as*

$$G(x) = \{y \in \Gamma(x) : v(x) = F(x, y) + \beta v(y)\}, \tag{16}$$

be the policy function (correspondence). Under the assumptions of Theorem 8, G is compact valued and upper hemi-continuous.

Proof. This follows immediately from Berge's maximum theorem. ■

We can next see how Theorem 2 enables us to establish more properties of the value function and the policy correspondence. In particular, for example, let us assume

Assumption 3 For each y, $F(\cdot, y)$ is strictly increasing in each of its first K arguments, and Γ is monotone in the sense that $x \leq x'$ implies $\Gamma(x) \subseteq \Gamma(x')$.

Theorem 9 *Let X, Γ, F, and β satisfy Assumptions 2 and 3, and let v be the unique solution to (14). Then v is strictly increasing.*

Proof. Let $C'(X) \subset C(X)$ be the set of bounded, continuous, nondecreasing functions on X, and let $C''(X) \subset C'(X)$ be the set of strictly increasing functions. Since $C'(X)$ is a closed subset of the complete metric space $C(X)$, by Theorem 2, it is sufficient to show that $T[C'(X)] \subseteq C''(X)$. Assumption 3 immediately implies that for any nondecreasing f, Tf is increasing, establishing the result. ■

Furthermore, let us impose

Assumption 4 F is strictly concave, i.e.,

$$F[\theta(x,y) + (1-\theta)(x',y')] \geq \theta F(x,y) + (1-\theta)F(x',y'),$$
$$\text{all } (x,y),(x',y') \in A, \quad \text{and all } \theta \in (0,1).$$

In addition, the inequality is strict if $x \neq x'$.

Moreover, Γ is convex in the sense that for any $0 \leq \theta \leq 1$, and $x, x' \in X$,

$$\begin{aligned} y &\in \Gamma(x) \quad \text{and } y' \in \Gamma(x') \text{ implies} \\ \theta y + (1-\theta)y' &\in \Gamma[\theta x + (1-\theta)x']. \end{aligned}$$

This assumption imposes enough concavity on the problem, in particular, it rules out "increasing returns" of any form.

Theorem 10 *Let X, Γ, F, and β satisfy Assumptions 2, 3 and 4, and let v satisfy (14); and let G satisfy (16). Then v is strictly concave and G is a continuous, single-valued function.*

Proof. The proof again follows from Theorem 2. Let $C'(X) \subset C(X)$ be the set of bounded, continuous, (weakly) concave functions on X, and let $C''(X) \subset C'(X)$ be the set of strictly concave functions. Since $C'(X)$ is a closed subset of the complete metric space $C(X)$, by Theorem 2, $T[C'(X)] \subseteq C''(X)$ would establish the results. To see this, let $f \in C'(X)$ and let

$$x_0 \neq x_1, \quad \theta \in (0,1), \quad \text{and} \quad x_\theta = \theta x_0 + (1-\theta)x_1.$$

Let $y_i \in \Gamma(x_i)$ attain $(Tf)(x_i)$, for $i = 0, 1$. Then Assumption 4 implies that $y_\theta = \theta y_0 + (1-\theta)y_1 \in \Gamma(x_\theta)$, so that

$$\begin{aligned} (Tf)(x_\theta) &\geq F(x_\theta, y_\theta) + \beta f(y_\theta) \\ &> \theta\,[F(x_0, y_0) + \beta f(y_0)] + (1-\theta)[F(x_1, y_1) + \beta f(y_1)] \\ &= \theta(Tf)(x_0) + (1-\theta)(Tf)(x_1), \end{aligned}$$

where the first line is as simple implication of (15) and the fact that $y_\theta \in \Gamma(x_\theta)$; the second line uses the hypothesis that f is concave and the concavity restriction on F from Assumption 4. Since these relationships are true for any $f \in C'(X)$, they establish $T[C'(X)] \subseteq C''(X)$, so that the unique fixed point v is strictly concave. Since, from Assumption 4, F is also concave and for each $x \in X$, $\Gamma(x)$ is convex, it follows that the maximum in (3) is attained at a unique y value. Hence G is a single-valued function, and its continuity follows from the fact that it is upper hemi-continuous. ■

2 The Basic Sequential Search Model

Let us start with the classical McCall model of search. This model is not only elegant, but has also become a workhorse for many questions in macro, labor and industrial organization. An important feature of the model is that it is much more tractable than the original Stigler formulation of search, as one of sampling multiple offers, but we will return to this theme below. Throughout, as a preparation for what it is going to come below, I'm going to consider the problem of the worker searching for a job (or a wage), but clearly, the problem of a consumer searching for the lowest price is isomorphic.

2.1 The Partial Equilibrium Model

Imagine a partial equilibrium setup with a risk neutral individual in discrete time. At time $t = 0$, this individual has preferences given by

$$\sum_{t=0}^{\infty} \beta^t c_t$$

where c_t is his consumption. He starts life as unemployed. When unemployed, he has access to consumption equal to b (from home production, value of leisure or unemployment benefit). At each time period, he samples a job. All jobs are identical except for their wages, and wages are given by an exogenous stationary distribution of $F(w)$ with finite (bounded) support $\mathbb{W}$, i.e., F is defined only for $w \in \mathbb{W}$. Without loss of any generality, we can take the lower support of $\mathbb{W}$ to be 0, since negative wages can be ruled out. In other words, at every date, the individual samples a wage $w_t \in W$, and has to decide whether to take this or continue searching. Draws from $\mathbb{W}$ over time are independent and identically distributed.

This type of sequential search model can also be referred to as a model of **undirected search**, in the sense that the individual has no ability to seek or direct his search towards different parts of the wage distribution (or towards different types of jobs). This will contrast with models of **directed search** which we will see later.

Let us assume for now that there is no recall, so that the only thing the individual can do is to take the job offered within that date (with recall, the individual would be able to accumulate offers, so at time t, he can choose any of the offers he has received up at that point). If he accepts a job, he will be employed at that job forever, so the net present value of accepting a job of wage w_t is

$$\frac{w_t}{1-\beta}.$$

This is a simple decision problem. Let us specify the class of decision rules of the agent. In particular, let

$$a_t : \mathbb{W} \to [0,1]$$

denote the action of the agent at time t, which specifies his acceptance probability for each wage in $\mathbb{W}$ at time t. Let $a_t' \in \{0,1\}$ be the realization of the action by the individual (thus allowing for mixed strategies). Let also A_t denote the set of realized actions by the individual, and define $A^t = \prod_{s=0}^{t} A_s$. Then a strategy for the individual in this game is

$$p_t : A^{t-1} \times \mathbb{W} \to [0,1]$$

Let $\mathcal{P}$ be the set of such functions, and $\mathcal{P}^\infty$ the set of infinite sequences of such functions. The most general way of expressing the problem of the individual would be as follows. Let $\mathbb{E}$ be the expectations operator. Then the individual's problem is

$$\max_{\{p_t\}_{t=0}^{\infty} \in \mathcal{P}^\infty} \mathbb{E} \sum_{t=0}^{\infty} \beta^t c_t$$

subject to $c_t = b$ if $t < s$ and $c_t = w_s$ if $t \geq s$ where $s = \inf \{n \in \mathbb{N} : a_n' = 1\}$. Naturally, written in this way, the problem looks complicated. In fact, the point of writing it in this way is to show that in certain classes of models, while the sequence problem will be complicated, the dynamic programming formulation will be quite tractable.

To show this, let us analyze this problem by writing it recursively using dynamic programming techniques. First, let us define the value of the agent when he has sampled a job of $w \in \mathbb{W}$. This is clearly given by

$$v(w) = \max\left\{\frac{w}{1-\beta}, \beta v + b\right\}, \tag{17}$$

where

$$v = \int_{\mathbb{W}} v(\omega)\, dF(\omega) \tag{18}$$

is the continuation value of not accepting a job. Here I have made no assumptions about the structure of the set $\mathbb{W}$, which could be an interval, or might have a mass point, and the density of the distribution F may not exist. Therefore, the integral in (18) should be interpreted as a Lebesgue integral.

Equation (17) follows from the observation that the individual will either accept the job, receiving a constant consumption stream of w (valued at $w/(1-\beta)$) or will turn down this job, in which case he will enjoy the consumption level b, and receive the continuation value v. Maximization implies that the individual takes whichever of these two options gives higher net present value.

Equation (18), on the other hand, follows from the fact that from tomorrow on, the individual faces the same distribution of job offers, so v is simply the expected value of $v(w)$ over the stationary distribution of wages.

We are interested in finding both the value function $v(w)$ and the optimal policy of the individual.

Combining these two equations, we can write

$$\begin{aligned} v(w) &= \max\left\{\frac{w}{1-\beta}, b+\beta\int_{\mathbb{W}} v(\omega)\,dF(\omega)\right\}, \\ &= Tv(w), \end{aligned} \tag{19}$$

where the second line defines the mapping T. Now (19) is in a form to which we can apply the above theorems. Blackwell's sufficiency theorem (Theorem 3) applies directly and implies that T is a contraction since it is monotonic and satisfies discounting.

Next, let $v \in C[\mathbb{W}]$, i.e., the set of continuous functions real-valued (hence bounded) defined over $\mathbb{W}$, which is a complete metric space with the sup norm. Then the contraction mapping theorem, Theorem 1, immediately implies that a unique value function $v(w)$ exists in this space. Thus the dynamic programming formulation of the sequential search problem immediately leads to the existence of an optimal solution (and thus optimal strategies, which will be characterized below).

Moreover, Theorem 2 also applies by taking S' to be the space of nondecreasing continuous functions over $\mathbb{W}$, which is a closed subspace of $C[\mathbb{W}]$. Therefore, $v(w)$ is nondecreasing. In fact, using Theorem 2 we could prove that $v(w)$ is piecewise linear with first a flat portion and then an increasing portion. Let the space of such functions be S'', which is another subspace of $C[\mathbb{W}]$, but is not closed. Nevertheless, now the second part of Theorem

2 applies, since starting with any nondecreasing function $v(w)$, $Tv(w)$ will be piecewise linear with first a flat portion. Therefore, the theorem implies that the unique fixed point $v(w)$ also has this property.

In fact, in this case, this property can be immediately seen from (19), which shows $v(w)$ to be piecewise linear, with first a flat portion.

The next task is to determine the optimal policy. But the fact that $v(w)$ is non-decreasing and is piecewise linear with first a flat portion, immediately tells us that the optimal policy will take a **reservation wage** form, which is a key result of the sequential search model. More explicitly, there will exist some reservation wage R such that all wages above R will be accepted and those $w < R$ will be turned down. Moreover, this reservation wage has to be such that

$$\frac{R}{1-\beta} = b + \beta \int_{\mathbb{W}} v(\omega)\, dF(\omega), \tag{20}$$

so that the individual is just indifferent between taking $w = R$ and waiting for one more period. Next we also have that since $w < R$ are turned down, for all $w < R$

$$\begin{aligned} v(w) &= b + \beta \int_{\mathbb{W}} v(\omega)\, dF(\omega) \\ &= \frac{R}{1-\beta}, \end{aligned}$$

and for all $w \geq R$,

$$v(w) = \frac{w}{1-\beta}$$

Therefore,

$$\int_{\mathbb{W}} v(\omega)\,dF(\omega) = \frac{RF(R)}{1-\beta} + \int_{w \geq R} \frac{w}{1-\beta} dF(w).$$

Combining this with (20), we have

$$\frac{R}{1-\beta} = b + \beta\left[\frac{RF(R)}{1-\beta} + \int_{w \geq R} \frac{w}{1-\beta} dF(w)\right]$$

Manipulating this equation, we can write

$$R = \frac{1}{1-\beta F(R)}\left[b(1-\beta) + \beta \int_{R}^{+\infty} w dF(w)\right],$$

which is one way of expressing the reservation wage. More useful is to rewrite this equation as

$$\int_{w<R} \frac{R}{1-\beta} dF(w) + \int_{w \geq R} \frac{R}{1-\beta} dF(w) = b + \beta\left[\int_{w<R} \frac{R}{1-\beta} dF(w) + \int_{w \geq R} \frac{w}{1-\beta} dF(w)\right]$$

Now subtracting $\beta R \int_{w \geq R} dF(w) / (1-\beta) + \beta R \int_{w<R} dF(w) / (1-\beta)$ from both sides, we obtain

$$\begin{aligned} & \int_{w<R} \frac{R}{1-\beta} dF(w) + \int_{w \geq R} \frac{R}{1-\beta} dF(w) \\ & -\beta \int_{w \geq R} \frac{R}{1-\beta} dF(w) - \beta \int_{w<R} \frac{R}{1-\beta} dF(w) \\ = \; & b + \beta\left[\int_{w \geq R} \frac{w-R}{1-\beta} dF(w)\right] \end{aligned}$$

Collecting terms, we obtain

$$R - b = \frac{\beta}{1-\beta}\left[\int_{w \geq R} (w-R)\,dF(w)\right], \tag{21}$$

which is an important way of characterizing the reservation wage. The left-hand side is best understood as the cost of foregoing the wage of R, while the right hand side is the expected benefit of one more search. Clearly, at the reservation wage, these two are equal.

One implication of the reservation wage policy is that the assumption of no recall, made above, was of no consequence. In a stationary environment, the worker will have a constant reservation wage, and therefore has no desire to go back and take a job that he had previously rejected.

Let us define the right hand side of equation (21) as

$$g(R) \equiv \frac{\beta}{1-\beta}\left[\int_{w \geq R} (w-R)\, dF(w)\right],$$

which represents the expected benefit of one more search as a function of the reservation wage. Clearly,

$$\begin{aligned} g'(R) &= -\frac{\beta}{1-\beta}(R-R) f(R) - \frac{\beta}{1-\beta}\left[\int_{w \geq R} dF(w)\right] \\ &= -\frac{\beta}{1-\beta}[1-F(R)] < 0 \end{aligned}$$

This implies that equation (21) has a unique solution. Moreover, by the implicit function theorem,

$$\frac{dR}{db} = \frac{1}{1-g'(R)} > 0,$$

so that as expected, higher benefits when unemployed increase the reservation wage, making workers more picky.

Moreover, for future reference, also note that when the density of $F(R)$, denoted by $f(R)$, exists, the second derivative of g also exists and is

$$g''(R) = \frac{\beta}{1-\beta} f(R) \geq 0,$$

so that the right hand side of equation (21) is also convex.

The next question is to investigate how changes in the distribution of wages F affects the reservation wage. Before doing this, however, I will use this partial equilibrium McCall model to derive a very simple theory of unemployment.

2.2 Unemployment with Sequential Search

Let us now use the McCall model to construct a simple model of unemployment. In particular, let us suppose that there is now a continuum 1 of identical individuals sampling jobs from the same stationary distribution F. Moreover, once a job is created, it lasts until the worker dies, which happens with probability s. There is a mass of s workers born every period, so that population is constant, and these workers start out as unemployed. The death probability means that the effective discount factor of workers is equal to $\beta(1-s)$. Consequently, the value of having accepted a wage of w is:

$$v^a(w) = \frac{w}{1-\beta(1-s)}.$$

Moreover, with the same reasoning as before, the value of having a job offer at wage w at hand is

$$v(w) = \max\{v^a(w), b + \beta(1-s)v\}$$

with

$$v = \int_{\mathbb{W}} v(w)\, dF.$$

Therefore, the same steps lead to the reservation wage equation:

$$R - b = \frac{\beta(1-s)}{1-\beta(1-s)}\left[\int_{w\geq R}(w-R)\,dF(w)\right]. \tag{22}$$

Now what is interesting is to look at the law of motion of unemployment. Let us start time t with U_t unemployed workers. There will be s new workers born into the unemployment pool. Out of the U_t unemployed workers, those who survive and do not find a job will remain unemployed. Therefore

$$U_{t+1} = s + (1-s)\,F(R)U_t,$$

where $F(R)$ is the probability of not finding a job (i.e., a wage offer below the reservation wage), so $(1-s)\,F(R)$ is the joint probability of not finding a job and surviving, i.e., of remaining unemployed. This is a first-order linear difference at wage and determining the law of motion of unemployment. Moreover, since $(1-s)\,F(R) < 1$, it is asymptotically stable, and will converge to a unique steady-state level of unemployment.

To get more insight, subtract U_t from both sides, and rearrange to obtain

$$U_{t+1} - U_t = s\,(1-U_t) - (1-s)\,(1-F(R))\,U_t.$$

This is the simplest example of the **flow approach** to the labor market, where unemployment dynamics are determined by flows in and out of unemployment. In fact is equation has the canonical form for change in unemployment in the flow approach. The left hand-side is the change unemployment (which can be either indiscreet or continuous time), while the right hand-side consists of the job destruction rate (in this case s) multiplied by $(1 - U_t)$ minus the rate at which workers leave unemployment (in this case $(1 - s)(1 - F(R))$) multiplied with U_t.

The unique steady-state unemployment rate where $U_{t+1} = U_t$ is given by

$$U = \frac{s}{s + (1 - s)(1 - F(R))}.$$

This is again the canonical formula of the flow approach. The steady-state unemployment rate is equal to the job destruction rate (here the rate at which workers die, s) divided by the job destruction rate plus the job creation rate (here in fact the rate at which workers leave unemployment, which is different from the job creation rate). Clearly, an increases in s will raise steady-state unemployment. Moreover, an increase in R, that is, a higher reservation wage, will also depress job creation and increase unemployment.

2.3 Aside on Riskiness and Mean Preserving Spreads

To investigate the effect of changes in the distribution of wages on the reservation wage, let us introduce the concept of **mean preserving spreads.** Loosely speaking, a mean preserving spread is a change in distribution that increases risk. Let a family of distributions over some set $X \subset \mathbb{R}$ with generic element x be denoted by $F(x, r)$, where r is a shift variable, which changes the distribution function. An example will be $F(x, r)$ to stand for mean zero normal variables, with r parameterizing the variance of the distribution. In fact, the normal distribution is special in the sense that, the mean and the variance completely describe the distribution, so the notion of risk can be captured by the variance. This is generally not true. The notion of "riskier" is a more stringent notion than having a greater variance. In fact, we will see that "riskier than" is a partial order (while, clearly, comparing variances is a complete order).

Here is a natural definition of one distribution being riskier than another, first introduced by Blackwell, and then by Rothschild and Stiglitz.

Definition 3 *$F(x,r)$ is less risky than $F(x,r')$, written as $F(x,r) \succeq_R F(x,r')$, if for all concave and increasing $u : \mathbb{R} \to \mathbb{R}$, we have*

$$\int_X u(x)\, dF(x,r) \geq \int_X u(x)\, dF(x,r').$$

A related definition is that of second-order stochastic dominance.

Definition 4 *$F(x,r)$ second order stochastically dominates $F(x,r')$, written as $F(x,r) \succeq_{SD} F(x,r')$, if*

$$\int_{-\infty}^{c} F(x,r)\, dx \leq \int_{-\infty}^{c} F(x,r')\, dx, \text{ for all } c \in X.$$

In other words, this definition requires the distribution function of $F(x,r)$ to start lower and always keep a lower integral than that of $F(x,r')$. One easy case where this will be satisfied is when both distribution functions have the same mean and they intersect only once: "single crossing") with $F(x,r)$ cutting $F(x,r')$ from below.

The definitions above use weak inequalities. Alternatively, they can be strengthened to strict inequalities. In particular, the first definition would require a strict inequality for functions that are strictly concave over some range, while the second definition will require strict inequality for some c.

Theorem 11 ***(Blackwell, Rothschild and Stiglitz)*** $F(x,r) \succeq_R F(x,r')$ *if and only if* $F(x,r) \succeq_{SD} F(x,r')$.

Therefore, there is an intimate link between second-order stochastic dominance and the notion of riskiness. This also shows that variance is not a good measure of riskiness, since second order stochastic dominance is a partial order.

Now **mean preserving spreads** are essentially equivalent to second-order stochastic dominance with the additional restriction that both distributions have the same mean. As the term suggests, a mean preserving spread is equivalent to taking a given distribution and shifting some of the weight from around the mean to the tails. Alternative representations also include one distribution being obtained from the other by adding "white noise" to the other.

Second-order stochastic dominance plays a very important role in the theory of learning, and also more generally in the theory of decision-making under uncertainty. Here it will be useful for comparative statics.

2.4 Back to the McCall Search Model

Let us return to the McCall search model. To investigate the effect of changes in the riskiness (or dispersion) of the wage distribution on reservation wages, and thus on search and unemployment behavior, let us express the reservation wage somewhat differently. Start with equation (21) above, which is reproduced here for convenience,

$$R - b = \frac{\beta}{1-\beta}\left[\int_{w \geq R}(w-R)\,dF(w)\right].$$

Rewrite this as

$$\begin{aligned} R - b &= \frac{\beta}{1-\beta}\left[\int_{w \geq R}(w-R)\,dF(w)\right] + \frac{\beta}{1-\beta}\left[\int_{w \leq R}(w-R)\,dF(w)\right] \\ &\quad -\frac{\beta}{1-\beta}\left[\int_{w \leq R}(w-R)\,dF(w)\right], \\ &= \frac{\beta}{1-\beta}(Ew - R) - \frac{\beta}{1-\beta}\left[\int_{w \leq R}(w-R)\,dF(w)\right], \end{aligned}$$

where Ew is the mean of the wage distribution, i.e.,

$$Ew = \int_{\mathbb{W}} w dF(w).$$

Now rearranging this last equation, we have

$$R - b = \beta(Ew - b) - \beta\int_{w \leq R}(w-R)\,dF(w).$$

Applying integration by parts to the integral on the right hand side, in particular, noting that

$$\begin{aligned} \int_{w \leq R} w dF(w) &= \int_0^R w dF(w) \\ &= wF(w)|_0^R - \int_0^R F(w)\,dw \\ &= RF(R) - \int_0^R F(w)\,dw, \end{aligned}$$

this equation can be rewritten as

$$R - b = \beta (Ew - b) + \beta \int_0^R F(w)\, dw. \tag{23}$$

Now consider a shift from F to $\tilde{F}$ corresponding to a mean preserving spread. This implies that Ew is unchanged, but by definition of a mean preserving spread (second-order stochastic dominance), the last integral increases. Therefore, the mean preserving spread induces a shift in the reservation wage from R to $\tilde{R} > R$. This reflects the greater option value of waiting when faced with a more dispersed wage distribution; lower wages are already turned down, while higher wages are now more likely.

A different way of viewing this result is that the analysis above established that the value function $v(w)$ is convex. While Theorem 11 shows that concave utility functions like less risky distributions, convex functions like more risky distributions.

3 Paradoxes of Search

The search framework is attractive especially when we want to think of a world without a Walrasian auctioneer, or alternatively a world with "frictions". How do prices get determined? How do potential buyers and sellers get together? Can we think of Walrasian equilibrium as an approximation to such a world under some conditions?

Search theory holds the promise of potentially answering these questions, and providing us with a framework for analysis.

3.1 The Rothschild Critique

The McCall model is an attractive starting point. It captures the intuition that individuals may be searching for the right types of job (e.g., jobs offering higher wages), trading off the prospects of future benefits (high wages) for the costs of foregoing current wages.

But everything hinges on the distribution of wages, $F(w)$. Where does this come from? Presumably somebody is offering every wage in the support of this distribution.

The basis of the Rothschild critique is that it is difficult to rationalize the distribution function $F(w)$ as resulting from profit-maximizing choices of firms.

Imagine that the economy consists of a mass 1 of identical workers similar to our searching agent. On the other side, there are N firms that can productively employ workers. Imagine that firm j has access to a technology such that it can employ l_j workers to produce

$$y_j = x_j l_j$$

units of output (with its price normalized to one as the numeraire, so that w is the real wage). Suppose that each firm can only attract workers by posting a single vacancy. Moreover, to simplify life, suppose that firms post a vacancy at the beginning of the game at $t = 0$, and then do not change the wage from then on. This will both simplify the strategies, and imply that the wage distribution will be stationary, since all the same wages will remain active throughout time. [Can you see why this simplifies the discussion? Imagine, for contrast, the case in which each firm only hires one worker; then think of the wage distribution at time t, $F_t(w)$, starting with some arbitrary $F_0(w)$. Will it remain constant?]

Suppose that the distribution of x in the population of firms is given by $G(x)$ with support $X \subset \mathbb{R}_+$. Also assume that there is some cost $\gamma > 0$ of posting a vacancy at the beginning, and finally, that $N >> 1$ (i.e., $N = \int_{-\infty}^{\infty} dG(x) >> 1$) and each worker samples one firm from the distribution of posting firms.

As before, we will assume that once a worker accepts a job, this is permanent, and he will be employed at this job forever. Moreover let us set $b = 0$, so that there is no unemployment benefits. Finally, to keep the environment entirely stationary, assume that once a worker accepts a job, a new worker is born, and starts search.

Will these firms offer a non-degenerate wage distribution $F(w)$?

The answer is no.

First, note that an endogenous wage distribution equilibrium would correspond to a function

$$p : X \to \{0, 1\},$$

denoting whether the firm is posting a vacancy or not, and if it is, i.e., $p = 1$,

$$h : X \to \mathbb{R}_+,$$

specifying the wage it is offering.

It is intuitive that $h(x)$ should be non-decreasing (higher wages are more attractive to high productivity firms). Let us suppose that this is so, and denote its set-valued inverse mapping by h^{-1}. Then, the along-the-equilibrium path wage distribution is

$$F(w) = \frac{\int_{-\infty}^{h^{-1}(w)} p(x)\, dG(x)}{\int_{-\infty}^{\infty} p(x)\, dG(x)}.$$

Why?

In addition, the strategies of workers can be represented by a function

$$a : \mathbb{R}_+ \to [0,1]$$

denoting the probability that the worker will accept any wage in the "potential support" of the wage distribution, with 1 standing for acceptance. This is general enough to nest non-symmetric or mixed strategies.

The natural equilibrium concept is subgame perfect Nash equilibrium, whereby the strategies of firms (p, h) and those of workers, a, are best responses to each other in all subgames.

The same arguments as above imply that all workers will use a reservation wage, so

$$\begin{aligned} a(w) &= 1 \text{ if } w \geq R \\ &= 0 \text{ otherwise} \end{aligned}$$

Since all workers are identical and the equation above determining the reservation wage, (21), has a unique solution, all workers will all be using the same reservation rule, accepting all wages $w \geq R$ and turning down those $w < R$. Workers' strategies are therefore again characterized by a reservation wage R.

Now take a firm with productivity x offering a wage $w' > R$. Its net present value of profits from this period's matches is

$$\pi\left(p = 1, w' > R, x\right) = -\gamma + \frac{1}{n}\frac{\left(x - w'\right)}{1 - \beta}$$

where

$$n = \int_{-\infty}^{\infty} p\left(x\right) dG\left(x\right)$$

is the measure of active firms, $1/n$ is the probability of a match within each period (since the population of active firms and searching workers are constant), and $x - w'$ is the profit from the worker discounted at the discount factor β.

Notice two (implicit) assumptions here: (1) wage posting: each job comes with a commitment to a certain wage; (2) undirected search: the worker makes a random draw from the distribution F, and the only way he can seek higher wages is by turning down lower wages that he samples.

This firm can deviate and cut its wage to some value in the interval $[R, w')$. All workers will still accept this job since its wage is above the reservation wage, and the firm will increase its profits to

$$\pi\left(p = 1, w \in [R, w'), x\right) = -\gamma + \frac{1}{n}\frac{x - w}{1 - \beta} > \pi\left(p = 1, w', x\right)$$

So there should not be any wages strictly above R.

Next consider a firm offering a wage $\tilde{w} < R$. This wage will be rejected by all workers, and the firm would lose the cost of posting a vacancy, i.e.,

$$\pi (p = 1, w < R, x) = -\gamma,$$

and this firm can deviate to $p = 0$ and make zero profits. Therefore, in equilibrium when workers use the reservation wage rule of accepting only wages greater than R, all firms will offer the same wage R, and there is no distribution and no search.

This establishes

Theorem 12 *When all workers are homogeneous and engage in undirected search, all equilibrium distributions will have a mass point at their reservation wage R.*

In fact, the paradox is even deeper.

3.2 The Diamond Paradox

The following result is one form of the Diamond paradox:

Theorem 13 ***(Diamond Paradox)*** *For all $\beta < 1$, the unique equilibrium in the above economy is $R = 0$.*

Given the Theorem 12, this result is easy to understand. Theorem 12 implies that all firms will offer the same wage, R.

Suppose $R > 0$, and $\beta < 1$. What is the optimal acceptance function, a, for a worker?

If the answer is

$$\begin{aligned} a(w) &= 1 \text{ if } w \geq R \\ &= 0 \text{ otherwise} \end{aligned}$$

then we can support all firms offering $w = R$ as an equilibrium (notice that the acceptance function needs to be defined for wages "off-the-equilibrium path"). Why is this important?

However, we can prove:

Lemma 2 *There exists $\varepsilon > 0$ such that when "almost all" firms are offering $w = R$, it is optimal for each worker to use the following acceptance strategy:*

$$\begin{aligned} a(w) &= 1 \textit{ if } w \geq R - \varepsilon \\ &= 0 \textit{ otherwise} \end{aligned}$$

Note: think about what "almost all" means here and why it is necessary.

Proof. If the worker accepts the wage of $R - \varepsilon$ today his payoff is

$$u^{accept} = \frac{R - \varepsilon}{1 - \beta}$$

If he rejects and waits until next period, then since "almost all" firms are offering R, he will receive the wage of R, so

$$u^{reject} = \frac{\beta R}{1 - \beta}$$

where the additional β comes in because of the waiting period. For all $\beta < 1$, there exists $\varepsilon > 0$ such that

$$u^{accept} > u^{reject},$$

proving the claim. ■

What is the intuition for this lemma?

But this implies that, starting from an allocation where all firms offer R, any firm can deviate and offer a wage of $R-\varepsilon$ and increase its profits. This proves that no wage $R>0$ can be the equilibrium, proving the proposition.

Notice that subgame perfection is important here. Here is a result that is for you to prove in the homework:

Theorem for Homework *Let $\bar{R}$ such that $\int_{\bar{R}+(1-\beta)\gamma}^{\infty} dG(x) = 1$, then all active firms offering $w = R$ for all $R \in \left[0, \bar{R}\right]$ can be supported as a Nash equilibrium.*

Of course, we know that these are non-subgame perfect Nash equilibria, and this highlights the importance of using the right equilibrium concept in the context of dynamic economies.

So now we are in a conundrum. Not only does there fail to be a wage distribution, but irrespective of the distribution of productivities or the degree of discounting, all firms offer the lowest possible wage, i.e., they are full monopsonists.

How do we resolve this paradox?

1. By assumption: assume that $F(w)$ is not the distribution of wages, but the distribution of "fruits" exogenously offered by "trees". This is clearly unsatisfactory, both from the modeling point of view, and from the point of view of asking policy questions from the model (e.g., how

does unemployment insurance affect the equilibrium? The answer will depend also on how the equilibrium wage distribution changes).

2. Introduce other dimensions of heterogeneity: to be done later.

3. Modify the wage determination assumptions: to be done in a little bit.

4 Diamond's Coconut Model: Search Without Prices

One line of attack that leads to a lot of insights is to maintain the flavor of search, but get rid of prices. The classic paper here is Diamond (1982), which builds on the methodological advances by Mortensen during the 1970s, in particular the use of Nash bargaining solution in markets with frictions.

4.1 Basic Environment

There is a continuum of agents in an island normalized to 1. Each agent enjoys the consumption of coconuts and has a linear utility function with the utility from one unit of consumption equal to y and with discount rate r in continuous time. Thus utility is

$$V(t) = \int_t^\infty e^{-rs} [u(s) - c(s)] \, ds$$

where $u(s) = y$ if the agent consumes at time s, and $u(s) = 0$ otherwise, and $c(s)$ is the cost of collecting coconuts, to be specified below.

The problem is that agents cannot consume coconuts that they themselves have collected, so there is need for "trade" in this economy. In addition, what makes this model non-trivial is that there are frictions, and agents have to "search" for other people willing to trade coconuts.

By assumption, trade takes an extremely simple form "without prices": two agents who meet exchange coconuts.

In addition, no agent can carry more than one coconut, and when the total measure of agents with coconuts looking for partners is e, they run into each other at the flow rate $b(e)$, in other words, the process for running into another agent with a coconut is Poisson at the rate $b(e)$, so during a short interval of length Δt, each potential trader will meet another trader with probability $b(e)\Delta t$ (and formally, the probability that he will meet two or more traders is $o(\Delta t)$, i.e., it goes to zero faster than Δt as $\Delta t \to 0$).

The key assumption is:

Assumption $b'(e) > 0$: **Thick Market Externalities.**

Think about why I am referring to this as "externalities"?

Finally, after consumption, agents become "unemployed" and they start running into coconut trees at the rate a, and conditional on running into a coconut tree, they decide whether to collect the coconut. Different trees have coconuts at different altitudes, so that the cost of collecting them varies. Assume they are iid draws from a distribution $G(c)$ with support $[\underline{c}, \bar{c}]$, where $\underline{c} > 0$.

With this structure, life on this Coconut Island is a simple. Once agents have a coconut, they simply wait until they meet another agent with a coconut, and trade. Without a coconut, they wait until they run into the coconut tree, and then they decide whether to collect the coconut given the cost.

In its most general form, and a strategy would be a mapping

$$\tilde{p} : [\underline{c}, \bar{c}] \times [0,1] \times \mathbb{R}_+ \to [0,1] \, ,$$

which determines a probability of collecting a coconut given its cost $c \in [\underline{c}, \bar{c}]$ as also a function of the current measure of traders looking for partners, $e \in [0,1]$, and time $t \in \mathbb{R}_+$. There is no obvious reason for why the strategy should be time-dependent, but there are natural reasons for it to depend on aggregate state variables, in particular e (the measure of agents looking for a coconut tree is simply $1-e$). It is natural to restrict attention to stationary policies, so

$$p : [\underline{c}, \bar{c}] \times [0,1] \to [0,1] \, .$$

An equilibrium is therefore a set of policies that are best response to each other in all subgames, or a symmetric equilibrium is a policy p that is the best response to itself, in the sense that when all other agents use p, it is optimal for each of them to do so as well.

4.2 Digression: Continuous Time Dynamic Programming

To characterize the equilibrium (at least the steady-state equilibrium) is very straightforward. To do this, let us distinguish two states in which the individual can be:

State E; holding a coconut and wants to trade, value V_E

State U; no coconut and wants to find a tree, value V_U.

Both of these values are given by continuous time dynamic programming equations. In particular,

$$rV_E(t) - \dot{V}_E(t) = b\left(e(t)\right)\left[y + V_U(t) - V_E(t)\right] \tag{24}$$

Intuition: holding a coconut is equivalent to holding an asset. This asset pays a dividend of y at the flow rate $b\left(e\left(t\right)\right)$, and at this point the individual changes "status" meaning he loses his asset and receives another asset worth $V_U\left(t\right)$—the value of an agent searching for a tree. In addition, while he's holding it, this asset might depreciate or appreciate as captured by the term $\dot{V}_E(t)$. The rate of return on his coconut asset, $rV_E(t)$, should equal the value of dividend plus appreciation, i.e., $\dot{V}_E(t) + b\left(e(t)\right)\left[y + V_U(t) - V_E(t)\right]$.

Mathematically, we can derive (24) as follows. Let E_t be the expectations at time t. Then:

$$\begin{aligned}
V_E\left(t\right) &= E_t \int_t^{\infty} e^{-r(s-t)}\left[u\left(s\right) - c\left(s\right)\right] ds \\
&= \left[\Delta t b\left(e(t)\right) + o\left(\Delta t\right)\right]\left[y + V_U(t)\right] \\
&\quad + \left[1 - \Delta t b\left(e(t)\right) - o\left(\Delta t\right)\right] e^{-r\Delta t} E_t \int_{t+\Delta t}^{\infty} e^{-r(s-t-\Delta t)}\left[u\left(s\right) - c\left(s\right)\right] ds \\
&= \left[\Delta t b\left(e(t)\right) + o\left(\Delta t\right)\right]\left[y + V_U(t)\right] \\
&\quad + \left[1 - \Delta t b\left(e(t)\right) - o\left(\Delta t\right)\right] e^{-r\Delta t} V_E\left(t + \Delta t\right)
\end{aligned}$$

The second line follows because a trading opportunity comes at the Poisson rate $b(e(t))$, and is always taken, and from then on the individual has value $V_U(t)$. By definition of the Poisson process during an interval of length Δt, the probability that there is one arrival is $\Delta t b(e(t))$, with probability of more than one arrival being of the order $o(\Delta t)$, meaning that this probability goes to zero faster than Δt as $\Delta t \to 0$. With probability $[1 - \Delta t b(e(t)) - o(\Delta t)]$, there is no arrival of a trading opportunity, so from then on the individual faces the same problem. Finally, E_t is used instead of $E_{t+\Delta t}$ because of the law of iterated expectations (i.e., $E_t[E_{t+\Delta t}(x)] = E_t x$).

Now rewriting this equation as

$$\begin{aligned} V_E(t) - V_E(t+\Delta t) &= [\Delta t b(e(t)) + o(\Delta t)]\left[y + V_U(t) - e^{-r\Delta t} V_E(t+\Delta t)\right] \\ &\quad + e^{-r\Delta t} V_E(t+\Delta t) - V_E(t+\Delta t) \end{aligned}$$

$$\begin{aligned} V_E(t) - V_E(t+\Delta t) &= [\Delta t b(e(t)) + o(\Delta t)]\left[y + V_U(t) - e^{-r\Delta t} V_E(t+\Delta t)\right] \\ &\quad + \left(e^{-r\Delta t} - e^{-r\times 0}\right) V_E(t+\Delta t) \end{aligned}$$

Now dividing all terms by Δt and taking limits

$$\begin{aligned} \lim_{\Delta t\to 0} \frac{V_E(t) - V_E(t+\Delta t)}{\Delta t} &= \lim_{\Delta t\to 0}\left[b(e(t)) + \frac{o(\Delta t)}{\Delta t}\right] \lim_{\Delta t\to 0}\left[y + V_U(t) - e^{-r\Delta t} V_E(t+\Delta t)\right] \\ &\quad + \lim_{\Delta t\to 0} \frac{e^{-r\Delta t} - e^{-r\times 0}}{\Delta t} V_E(t+\Delta t), \end{aligned}$$

which, assuming that $V_E(t)$ is continuously differentiable, implies

$$-\dot{V}_E(t) = b(e(t))\,[y + V_U(t) - V_E(t)] - rV_E(t),$$

which is identical to (24).

4.3 Back to the Model

With a similar reasoning, we can write

$$\begin{aligned} rV_U(t) - \dot{V}_U(t) &= \max_{p(e(t),c)} \left\{ a \int_{\underline{c}}^{\bar{c}} p\left(e\left(t\right), c\right) \left[V_E(t) - V_U(t) - c\right] dG(c) \right\} \\ &= a \int_{\underline{c}}^{\bar{c}} \max\left\{ \left[V_E(t) - V_U(t) - c\right]; 0 \right\} dG(c) \end{aligned}$$

where $p\left(e\left(t\right), c\right)$ denotes whether when the aggregate state is $e\left(t\right)$, the individual will collect a coconut of cost c. Given the monotonicity of this equation, standard arguments again imply that there will be a reservation cost policy, i.e., $p\left(e\left(t\right), c\right) = 1$ for all $c \leq c^*(t)$ and $= 0$ otherwise (where I write $c^*(t)$ instead of $c^*(e\left(t\right))$ to simplify notation). Moreover, at $c = c^*(t)$, $V_E(t) - V_U(t) - c = 0$, so

$$c^*(t) = V_E(t) - V_U(t)$$

Therefore, we can write

$$rV_U(t) - \dot{V}_U(t) = a \int_{\underline{c}}^{c^*(t)} \left[V_E(t) - V_U(t) - c\right] dG(c) \tag{25}$$

with $c^*(t) = V_E(t) - V_U(t)$.

Given this reservation cost policy, we also have the behavior of the aggregate state of the economy as

$$\dot{e}(t) = a\left(1 - e(t)\right) G\left(c^*(t)\right) - b\left(e(t)\right) e(t) \tag{26}$$

Intuition: the number (measure) of agents looking for a trading opportunity, $e\left(t\right)$ declines because the existing $e(t)$ agents meet each other at the rate

$b(e(t))$, hence the term $-b(e(t))\,e(t)$, and it increases because of the $1-e(t)$ agents looking for a coconut tree, a fraction a of them find a tree, and a fraction of those decide to collect the coconut.

Now let us impose steady state, so that $\dot{V}_U(t) = \dot{V}_E(t) = 0$, $e(t) = e^*$, and $c^*(t) = c^*$, then solving (24) and (25), we obtain:

$$c^* = \frac{b(e^*)y + a\int_{\underline{c}}^{c^*} cdG(c)}{r + b(e^*) + aG(c^*)}$$

and

$$e^* = \frac{aG(c^*)}{aG(c^*) + b(e)}$$

as the two steady state equilibrium relationships. Equilibria are given by the intersection of these two schedules.

Note that both are upward sloping :

$$\left.\frac{de^*}{dc^*}\right|_{\dot{e}=0} = \frac{a(1-e^*)G'(c^*)}{b(e^*) + e^*b'(e^*) + aG(c^*)} > 0$$

and

$$\left.\frac{dc^*}{de^*}\right|_{c^*=V_E-V_U} = \frac{(y-c^*)b'(e^*)}{r + b(e^*) + aG(c^*)} > 0$$

Moreover,

$$\left.\frac{d^2c^*}{d(e^*)^2}\right|_{c^*=V_E-V_U} = \frac{(y-c^*)b''(e^*) - 2b'(e^*)\,(dc^*/de^*) - aG'(c)\,(dc^*/de^*)^2}{r + b(e^*) + aG(c^*)} < 0$$

as long as b is concave (and of course we have $y - c^* > 0$, why?)

In addition, note that if $c^* \leq \underline{c}$, then $e^* = 0$ (why?)

Since both curves are upward sloping, a multiplicity of equilibria is possible. In fact, multiple steady-state equilibria is quite common here, since $e^* = 0$ is always a steady state equilibrium (why?).

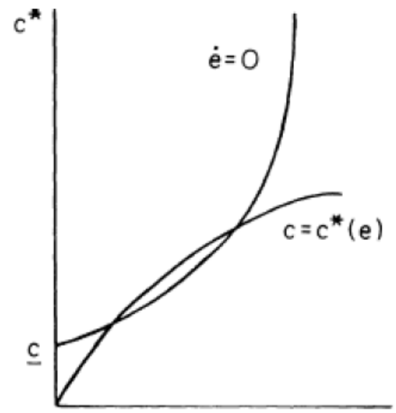

Figure 1:

Also, it is straightforward to prove that in all equilibria, the number of agents looking for a trade, i.e., the level of e, is suboptimally low. To see this, imagine that a planner forces all agents to accept using a policy $\bar{p}$ such that all $c \leq c^* + \delta$ are accepted, where $\delta > 0$ and small. Let us do this informally and compare steady states (why is this not the right thing? Think about the cost of moving from one steady state to another...)

First, assuming that $g(c^*)$ exist and letting $e^* = e^*[\delta]$, we have

$$\left.\frac{de^*}{d\delta}\right|_{\delta=0} = \frac{ag(c^*)\,b(e^*)}{[aG(c^*) + b(e^*)]^2}.$$

The value of agents searching for trees in steady state is now:

$$rV_U[\delta] = a\int_{\underline{c}}^{c^*+\delta} [V_E[\delta] - V_U[\delta] - c]\, dG(c).$$

Totally differentiating this with respect to δ, we obtain

$$V_U' [\delta = 0] = \frac{aG(c^*) V_E' [\delta = 0]}{r + aG(c^*)} \tag{27}$$

Moreover, in steady state:

$$V_E [\delta] = \frac{b(e^*(\delta))(y + V_U[\delta])}{r + b(e^*(\delta))},$$

so using (27), we have

$$\begin{aligned} V_E' [\delta = 0] &= \frac{b'(e^*)[y + V_U]}{[r + b(e^*)]^2} + V_U' [\delta = 0] \\ &= \frac{b'(e^*)[y + V_U]}{[r + b(e^*)]^2} + \frac{aG(c^*) V_E' [\delta = 0]}{r + aG(c^*)} \\ &= \frac{b'(e^*)(r + aG(c^*))[y + V_U]}{r[r + b(e^*)]^2} > 0 \end{aligned}$$

which is strictly positive given the thick market externalities, i.e., $b'(e) > 0$. Moreover, this implies from (27) that $V_U' [\delta = 0] > 0$. Consequently, a small increase in c^* (i.e., in the reservation cost of agents looking for coconuts) will result in an increase in the values of both types of agents. Therefore, none of the equilibria are constrained Pareto efficient. Some type of subsidy to "search" would improve the welfare of all agents. This is a consequence of the thick market externalities.

Moreover, the three steady-state equilibria are Pareto Ranked. The one that has higher e Pareto dominates the others, the proof of which is left for you as homework exercise.

All of these follow from the thick market externalities, $b'(e) > 0$.

Therefore, two special features of this model:

1. No prices

2. Thick market externalities.

Both of these features look natural in this context because there is only one type of agent (everything is symmetric). But this is deceptive.

Imagine that there are coconut sellers and coconut buyers. Then the matching probability of a buyer should be of the form

$$b^B(e_B, e_S)$$

where e_B is the measure of buyers and e_S is the measure of sellers. We naturally expect this function to be decreasing in its first argument and increasing in its second argument. These are the two sister externalities, congestion and thick market that we will be playing with throughout.

Then a symmetric situation with $e = e_B = e_S$ is simply one where

$$b(e) = b^B(e, e),$$

and $b'(e) > 0$ is equivalent to the function b^B exhibiting increasing returns to scale. Is this reasonable? It depends on the model of search and what we think the search process approximates in reality (more on this below).

Moreover, when we have two sides to the trades, thinking of prices becomes more compelling; if there are too few buyers and too many sellers, the prices that buyers pay should fall etc.

5 Search With Bargained Prices

We now investigate how these two assumptions matter for the results, and in the process develop a baseline search and matching model for macroeconomic analysis. The model is basically that of Diamond (RES, 1982), Mortensen (AER, 1982), and Pissarides (AER, 1985). Pissarides (2004) gives a very accessible exposition. The results I present here are largely based on Hosios (RES, 1989), but the exposition is very different (much more standard here than in the original article). The importance of this model also stems from the fact that it is very closely related to the Mortensen-Pissarides model we will use to analyze unemployment fluctuations later in the class.

5.1 Environment and Preliminaries

The model is again continuous time, infinite horizon, and agents are risk neutral with discount rate r, i.e., maximizing

$$V(t) = \int_t^{\infty} e^{-rs} y(s)\, ds$$

where $y(t)$ is their net income at time t.

Let's assume for now that there is an exogenously given stock of buyers (firms) of measure N, who can employ workers productively. The productivity of each firm is determined as a draw from the distribution

$$x \sim F(x),$$

with support X, after the match between the firm and the worker (i.e., only ex post heterogeneity). This productivity remains constant throughout the life of the match.

All workers and firms are ex ante identical.

The population of workers, i.e., stock of sellers, is L. I also denote the number (measure) of unemployed workers by U, and the number of firms is N and the number of vacant firms looking for workers by V.

Frictions in the labor market are modeled by way of a matching function

$$M(U,V)$$

which determines the flow rate of matches per instant when the stock of unemployed workers is U and the stock of vacancies is V. We make a number of assumptions on this function:

1. Increasing: more vacancies and more unemployed workers result in more matches $M_U, M_V > 0$ (this could be modified to ≥ 0)

2. Externalities:
$$\frac{\partial \frac{M(U,V)}{U}}{\partial V} > 0, \frac{\partial \frac{M(U,V)}{U}}{\partial U} < 0, \frac{\partial \frac{M(U,V)}{V}}{\partial V} < 0, \text{ and } \frac{\partial \frac{M(U,V)}{V}}{\partial U} > 0,$$
meaning that when there are more vacancies, the matching probability of a given unemployed worker increases, and the matching probability of a given unfilled vacancy decreases, and similarly in response to an increase in unemployment.

One important point here is that although there are frictions, there is a sense in which the market is "regular" meaning that when there are more vacancies demanding matches, it's easier for unemployed workers to find matches and vice versa.

The second assumption clarifies that there are both negative and positive externalities in this world.

Digression on pecuniary and non-pecuniary externalities: why do I refer to those in (2) as externalities? Is it an externality if I demand one more apple? What is the difference between me demanding one more apple and a firm demanding one more worker by posting a vacancy?

A couple of other points are useful to note:

- If $M(U, V) \to \infty$, matching frictions are disappearing, so it will be interesting to investigate whether we approach a competitive model (some care needs to be taken here; frictionless matching corresponds to the case where at an instant $\min\{U, V\}$ jobs are created; as a flow rate this corresponds to $M(U, V) \to \infty$).

- A natural benchmark, which seems to be consistent with the data both in the US and the UK is that $M(U, V)$ exhibits constant returns to scale, so when there is a doubling in the number of unemployed workers and vacancies, the total number of matches within a given period also doubles (and matching probabilities remain unchanged). I do not impose constant returns to scale yet, but this will play an important role below.

Finally, let us assume that there are exogenous separations once a worker and the firm come together at the flow rate s.

The cost of a vacancy is γ, and when unemployed, workers receive income or benefit from leisure equal to b.

Let ask focus on steady-state equilibria here (non-steady-state equilibria are for the homework).

A steady state equilibrium in this economy will specify a wage rate function

$$w : X \to \mathbb{R}_+$$

assigning a wage for every level of productivity, and a pair of acceptance functions for the worker and the firm

$$\begin{aligned} a_w &: X \to [0,1] \\ a_f &: X \to [0,1] \end{aligned}$$

assigning a probability of accepting a job of productivity x after this productivity is realized following the match. We will see that the decision of the worker and the firm can be collapsed into a single function without loss of any generality below, so we can work with the simpler setup where there is only one function

$$a : X \to [0,1]$$

to be determined.

I still need to specify the wage determination process, which will be done below, but for now note that wages will be "bargained" between the worker and the firm.

Let us define

$$q = \frac{M(U,V)}{V}$$

as the matching probability for an unfilled vacancy, and

$$p = \frac{M(U,V)}{U},$$

as the matching probability for an unemployed worker.

Imposing steady states, the asset value of a filled job of productivity x can be written as

$$rJ^F(x) = x - w(x) + s\left(J^V - J^F(x)\right), \tag{28}$$

where J^V is the value of an unfilled vacancy, and x is the output of the match. Now an immediate application of Theorem 2 implies that $J^F(x)$ is strictly monotonically increasing in x as long as $x - w(x)$ is increasing in x (i.e., when differentiable, $w'(x) < 1$ everywhere)—can you see why this will be true? Think of wage bargaining or see below. This implies that for the firm a reservation productivity rule will be optimal. One complication is that some matches may be acceptable to the firm, but not to the worker. We will see that this will not be the case, but this may need to be borne in mind.

Now the value of an unfilled vacancy is

$$rJ^V = -\gamma + q\int_X \max\left\{\left(J^F(x) - J^V\right); 0\right\} dF(x), \tag{29}$$

which imposes the decision that the firm will only accept a match if this yields a higher value $J^F(x)$ for it. This expression ignores the decision of the worker. More generally we should write

$$rJ^V = -\gamma + q\int_X \max\left\{a_w\left(x\right)\left(J^F(x) - J^V\right);0\right\}dF(x),$$

which means that the firm can only choose to create the match if $a_w\left(x\right) = 1$, i.e., if the worker also want to create the match. Let us ignore this for now and work with (29)—we will see below, why it is okay to ignore this, and also why even if this were not the case, the analysis would not be much more complicated.

From the monotonicity of $J^F(x)$, (29) can be simplified to

$$rJ^V = -\gamma + q\int_{x^*}^{\infty}\left(J^F(x) - J^V\right)dF(x) \tag{30}$$

where x^* is the reservation productivity of the firm.

The value functions for the workers are similar

$$rJ^E(x) = w(x) + s\left(J^U - J^E(x)\right),$$

Now presuming that $w\left(x\right)$ is strictly increasing in x (again see below), Theorem 2 immediately implies that this is strictly increasing in x, so a reservation productivity rule will also be optimal for the worker in deciding whether to accept a job or not.

Imposing that this is the same threshold for the worker as for the firm, we can then write

$$rJ^U = b + p\int_{x^*}^{\infty}(J^E(x) - J^U)dF(x). \tag{31}$$

Naturally, both equations (30) and (31) will be valid even if the worker and the firm use different reservation productivity rules, with x^* corresponding to the maximum of these two thresholds. (Can you see why? Think of the case in which workers and firms use two different cutoff levels x_w^* and x_f^*; what would happen then?)

Now given this, we can write the law motion of the number of vacancies in unemployment as

$$\begin{aligned}\dot{U} &= s(L-U) - p(1-F(x^*))U \\ \dot{V} &= s(N-V) - q(1-F(x^*))V\end{aligned}$$

How are wages determined? Nash Bargaining.

Why do we need bargaining? Because of bilateral monopoly, or much more specifically: match-specific surplus (or as is sometimes called *quasi-rents*).

Think of a competitive labor market, at the margin the firm is indifferent between employing the marginal worker or not, and the worker is indifferent between supplying the marginal hour or not (or working for this firm or another firm). We can make both parties indifferent at the same time—no match-specific surplus.

In a frictional labor market, if we choose the wage such that $J^E(x) = 0$, we will typically have $J^F(x) > 0$ and vice versa. There is some surplus to be shared.

Nash solution to bargaining is a natural benchmark.

5.2 Digression: Nash's Solution to Bargaining

Nash's bargaining theorem considers the bargaining problem of choosing a point x from a set $X \subset \mathbb{R}^N$ for some $N \geq 1$ by two parties with utility functions $u_1(x)$ and $u_2(x)$, such that if they cannot agree, they will obtain respectively d_1 and d_2. The theorem is that if we impose the following four conditions: (1) $u_1(x)$ and $u_2(x)$ are Von Neumann-Morgenstern utility functions, i.e., concave, increasing and unique up to positive linear transformations; (2) Pareto optimality, the agreement point will be along the frontier; (3) Independence of Irrelevant Alternatives: suppose $X' \subset X$ and the choice when bargaining over the set X is $x' \in X'$, then x' is also the solution when bargaining over X'; (4) Symmetry: identities of the players do not matter, only their utility functions; then there is a unique bargaining solution which is

$$x^{NS} = \arg\max_{x \in X} \left(u_1(x) - d_1\right)\left(u_2(x) - d_2\right)$$

If we relax the symmetry assumption, so that the identities of the players can matter (e.g., worker versus firm have different "bargaining powers"), then we obtain:

$$x^{NS} = \arg\max_{x \in X} \left(u_1(x) - d_1\right)^{\beta}\left(u_2(x) - d_2\right)^{1-\beta} \tag{32}$$

where $\beta \in [0, 1]$ is the bargaining power of player 1.

Next note that if both utility functions are linear and defined over their share of some pie, and the set $X \subset \mathbb{R}^2$ is given by $x_1 + x_2 \leq 1$, then the solution to (32) is given by

$$x_2 = (1 - \beta)(1 - d_1 - d_2) + d_2$$

or

$$(1 - \beta)(1 - x_2 - d_1) = \beta (x_2 - d_2)$$

and $x_1 = 1 - x_2$.

With Nash bargaining in place, we can already answer the question of why both firms and workers are using the same threshold x^*. Recall that $w(x)$ resulting from Nash bargaining satisfies "Pareto optimality". Suppose that there exists some x' such that $w(x')$ is the equilibrium wage, and $J^E(x') < J^U$ and $J^F(x') > J^V$. In that case, if $J^E(x') + J^F(x') > J^U + J^V$, we can find some $\tilde{w}(x') > w(x')$ such that both $J^E(x')$ and $J^F(x')$ are positive. Thus the two parties can agree to produce together, improving their welfare relative to disagreement, conflicting the presumption that $w(x')$ was part of an equilibrium (it would not have been Pareto optimal). Conversely, if $J^E(x') + J^F(x') < J^U + J^V$, then there should in fact be a separation at x'.

5.3 Back to the Model

Put differently, bargaining is going to ensure that separations are mutually beneficial, and thus x^* will be such that $J^E(x^*) + J^F(x^*) = J^U + J^V$.

Applied to our setting, let's assume that the worker has bargaining power β. Then, the Nash bargaining solution implies that the wage function $w(x)$

will be a solution to the equation:

$$\beta\left(J^F(x)-J^V\right)=(1-\beta)\left(J^E(x)-J^U\right) \tag{33}$$

Now using the value functions, we obtain

$$\begin{aligned} J^E(x)-J^U &= \frac{w(x)-rJ^U}{r+s} \\ J^F(x)-J^V &= \frac{x-w(x)-rJ^V}{r+s} \end{aligned} \tag{34}$$

then substituting into (33), we have

$$\begin{aligned} w(x) &= \beta\left(x-rJ^U-rJ^V\right)+rJ^U \\ &= \beta x+(1-\beta)rJ^U-\beta rJ^V \end{aligned} \tag{35}$$

This wage equation is very intuitive. The worker receives a fraction β of total surplus of the flow value of match, $x-rJ^U-rJ^V$, plus his outside option (more appropriately disagreement point), rJ^U.

Digression: when is the Nash solution the equilibrium of a well-specified bargaining game?

In addition, we have

$$rJ^U=b+p\int_{x^*}^{\infty}\left[\frac{w(x)-rJ^U}{r+s}\right]dF(x)$$

Now substituting for $w(x)$ from (35), gives

$$rJ^U=b+\frac{p\beta\bar{x}}{r+s}-\frac{p\beta\phi^*}{r+s}\left(rJ^U+rJ^V\right)$$

where

$$\bar{x}\equiv\int_{x^*}^{\infty}xdF(x)$$

Note that this is *not* the expectation of x conditional on $x \geq x^*$. That conditional expectation would be $E[x \mid x \geq x^*] = \int_{x^*}^{\infty} x dF(x) / [1 - F(x^*)]$, so $\bar{x}$ is the conditional expectation times the probability that x is indeed greater than x^*.

Moreover, let that probability be denoted by

$$\phi^* \equiv 1 - F(x^*).$$

Similarly

$$rJ^V = -\gamma + \frac{q(1-\beta)\bar{x}}{r+s} - \frac{q(1-\beta)\phi^*(rJ^U + rJ^V)}{r+s}$$

This implies that the sum of the disagreement points for a firm and a worker is:

$$\Rightarrow rJ^U + rJ^V = \frac{(r+s)(b-\gamma) + q(1-\beta)\bar{x} + p\beta\bar{x}}{r+s+q(1-\beta)\phi^* + p\beta\phi^*} \tag{36}$$

The preceding argument already establishes that in equilibrium

$$x^* = rJ^U + rJ^V$$

(to derive this equation, can use (34) together with the fact that $J^U + J^V = J^E(x^*) + J^F(x^*)$; alternatively, use $w(x^*) = rJ^U$ and $x^* - w(x^*) = rJ^V$)

Combining this with (36), we have

$$x^* = \frac{q(1-\beta)\bar{x} + p\beta\bar{x} + (r+s)(b-\gamma)}{r+s+q(1-\beta)\phi^* + p\beta\phi^*} \tag{37}$$

This equation defines x^* implicitly. (Recall that $\bar{x} \equiv \int_{x^*}^{\infty} x dF(x)$, $\phi^* \equiv 1 - F(x^*)$).

In addition, the equation (37) may be rewritten as follows:

$$\begin{aligned}(r+s)\left(x^*-(b-\gamma)\right) &= \left(q(1-\beta)+p\beta\right)\left(\bar{x}-\phi^* x^*\right)\\ &= \left(q(1-\beta)+p\beta\right)\int_{x^*}^{\infty}(x-x^*)dF(x)\end{aligned}$$

From the point of view of an individual worker or firm, who takes q and p as given, this equation characterises a unique value for x^*. The left hand side is increasing in x^*. Holding q and p constant

$$\frac{dRHS}{dx^*}=-\left(q(1-\beta)+p\beta\right)\left(1-F(x^*)\right)<0$$

which establishes the result. The result means that all workers and firms in the economy deduce the same threshold x^* from observing U and V. For the economy as a whole, changing the value of x^* will cause changes in q and p. Characterizing uniqueness in this context is more demanding. In particular, as we will see below, steady-state unemployment is

$$U=\frac{sL}{s+p\left(1-F\left(x^*\right)\right)}$$

and steady-state vacancy level is

$$V=\frac{sN}{s+q\left(1-F\left(x^*\right)\right)}$$

Inverting these equations, we have

$$p=\frac{s\left(L-U\right)}{U\left(1-F\left(x^*\right)\right)}=\frac{M\left(U,V\right)}{U}\text{ and }q=\frac{s\left(N-V\right)}{V\left(1-F\left(x^*\right)\right)}=\frac{M\left(U,V\right)}{V},$$

which jointly solve for U and V as functions of x^*, and thus pin down p and q as functions of x^*. If we make further assumptions on $M\left(U,V\right)$, we can derive conditions under which x^* will be uniquely determined in general equilibrium. But I will not pursue this here.

Now continuing with the analysis, we also have

$$\begin{aligned} rJ^U &= b + \frac{\beta p \bar{x} - (b-\gamma)\beta p \phi^*}{r+s+q(1-\beta)\phi^* + p\beta\phi^*} \\ rJ^V &= -\gamma + \frac{(1-\beta) q \bar{x} - (b-\gamma)(1-\beta) q \phi^*}{r+s+q(1-\beta)\phi^* + p\beta\phi^*} \end{aligned} \tag{38}$$

which completes the description of the equilibrium (recall that $\bar{x} \equiv \int_{x^*}^{\infty} x dF(x)$).

We can also calculate the number of unemployed workers (or the unemployment rate in this economy) now. Notice that we could characterize x^* without worrying about unemployment. This is a common feature of many search models, sometimes referred to as "block recursiveness".

The unemployment evolution is given by

$$\dot{U} = s\left(L-U\right) - p\left(1-F\left(x^*\right)\right)U \tag{39}$$

where the first term is separations from existing jobs, of which there are $L-U$, and the second term is job creation, which happens at the flow rate $p\left(1-F\left(x^*\right)\right)$. Thus in steady-state

$$U = \frac{sL}{s+p\left(1-F\left(x^*\right)\right)},$$

or defining the unemployment rate as $u = U/L$,

$$u = \frac{s}{s+p\left(1-F\left(x^*\right)\right)}.$$

It is natural to look towards comparative statics now.

First, consider an increase in b, the level of utility or benefits in unemployment. For given q and p, a higher b would increase x^* from (37), reducing the probability of job creation conditional on a match, and increasing unemployment. However, the effect of b on overall unemployment also needs to take into account the changes in q and p. Can you derive this effect? Can you derive the effect of γ on unemployment?

Also note that an increase in β, the bargaining power of the workers, and an increase in r, the discount rate, have ambiguous effects. Why?

Now consider an extended model where new agents can enter at the per period cost c_W for workers and c_F for firms, and the initial stocks of workers and firms is small enough (why is this caveat necessary?). What is the equilibrium of this extended model? It is straightforward to see that as long as $c_W < rJ^U$, there will be entry. Similarly for firms. Then in equilibrium we also have

$$c_W = rJ^U; \quad c_F = rJ^V.$$

In fact, the standard Mortensen-Pissarides search model, which will be analyzed in greater detail later in the class, is one where L is constant, $c_W = \infty$, $N = 0$ and $c_F = 0$. Now in this case, we have another equilibrium condition, given by

$$rJ^V = -\gamma + \frac{(1-\beta)q\bar{x} - (b-\gamma)(1-\beta)q\phi^*}{r + s + q(1-\beta)\phi^* + p\beta\phi^*} = 0.$$

Calculating the comparative statics in this model is not as easy as it seems. This is because when the matching function has decreasing or increasing returns to scale, there can be difficulties in establishing comparative statics. We will do much more of these comparative statics when we look at the Mortensen-Pissarides model in the context of understanding unemployment fluctuations later, but these models will impose constant returns to scale matching. Here intuitively, we expect

$$\frac{\partial U}{\partial b} > 0,\ \frac{\partial U}{\partial \gamma} > 0,\ \frac{\partial U}{\partial \beta} > 0,\ \text{and}\ \frac{\partial U}{\partial r} > 0,$$

but for now, you will be asked to derive these results only in the case with constant returns to scale matching and free entry in the homework (you can think about the intuition more generally, however, if you want).

Now we can investigate what happens to equilibrium as the amount of frictions are diminished, i.e., $M(U,V) \to \infty$. For this case, seem that N and W are constant (i.e., no free entry). This implies that $p, q \to \infty$, and $p/q \to\ P$ finite. In this case, also assumed that there is bounded support on x, in particular, let the upper bound be $x^{\sup}$. Taking limits, we have from (37)

$$x^* \to \frac{(1-\beta) + P\beta}{(1-\beta) + P\beta} \frac{\int_{x^*}^{x^{\sup}} x dF(x)}{1 - F(x^*)} = \frac{\int_{x^*}^{x^{\sup}} x dF(x)}{1 - F(x^*)} = E\left[x \mid x \geq x^{\sup}\right]$$

which is only possible if $x^* = x^{\sup}$ (why?). Thus exactly as in a competitive equilibrium, only the most productive jobs are active in equilibrium (homework question: what happens to wages in the limit?).

How does this equilibrium compare to the "second-best", that is the solution to the planner's problem where the constraints are the same as those imposed on the decentralized economy. Therefore, the planner's problem is to maximize output subject to search constraints.

The following current value Hamiltonian describes the problem of the planner. To simplify, I have already imposed the cutoff rule that all jobs above some $\tilde{x}^*$ will be active (otherwise, the program has to follow each x and choose $a(x)$ again as the probability of a match conditional on productivity x).

To write this Hamiltonian, reason as follows. The planner creates $M(U,V)$ matches at every instant when there are U unemployed workers and V vacancies. Only a fraction $1-F(\tilde{x}^*)$ of the matches are turned into jobs. Each job is worth on average $E[x \mid x \geq \tilde{x}^*]$ conditional on being created. Finally, a job of productivity x has a discounted net present value equal to $x/(r+s)$, because of discounting and potential future separations. Thus the net return to the planner during an instant can be written as

$$\frac{E[x \mid x \geq \tilde{x}^*]}{r+s}(1-F(\tilde{x}^*))M(U,V)+bU-\gamma V$$

where the last two terms are the net income flows from unemployed workers and vacant firms. This is simply equal to

$$\frac{\int_{\tilde{x}^*}^{\infty} x dF(x)}{r+s}M(U,V)+bU-\gamma V$$

Thus the objective function of the planner is to maximize

$$\int_0^{\infty} e^{-rt}\left[\frac{\int_{\tilde{x}^*}^{\infty} x dF(x)}{r+s}M(U,V)+bU-\gamma V\right]$$

where I suppressed time dependence to save on notation.

Now adding the constraints with corresponding multipliers, the Hamiltonian is

$$\begin{aligned}
H &= \left[\frac{\int_{\tilde{x}^*}^{\infty} xdF(x)}{r+s} M(U,V) + bU - \gamma V\right] \\
&\quad + \lambda \left[s(L-U) - (1-F(\tilde{x}^*))M(U,V)\right] \qquad \text{search constraint} \\
&\quad + \mu \left[N - V - L + U\right] \qquad \text{adding up constraint.}
\end{aligned}$$

Here the control variables are $\tilde{x}^*, V$, and the stock variable is U (recall the constraint (39)).

In addition, the multipliers are:

λ: social value of one more match.

μ: social value of one more vacancy. (Why? Why not the value of one more worker?)

This is a standard optimal control problem, with necessary conditions

$$\begin{aligned}
\frac{\partial H}{\partial \tilde{x}^*} &= 0 \\
\frac{\partial H}{\partial U} &= r\lambda - \dot{\lambda} \\
\frac{\partial H}{\partial V} &= 0
\end{aligned}$$

As in the equilibrium, let us focus on steady states: $\dot{\lambda} = 0$.

The first-order conditions are:

With respect to $\tilde{x}^*$

$$\left(-\frac{\tilde{x}^* f(\tilde{x}^*)}{r+s}\right) M(U,V) + \lambda f(\tilde{x}^*) M(U,V) = 0.$$

Or rearranging:

$$\tilde{x}^* = (r+s)\lambda \tag{40}$$

Thus the cutoff threshold has to be proportional to the shadow value of one more unemployed worker appropriately discounted. What is the intuition?

With respect to U :

$$\frac{\int_{\tilde{x}^*}^{\infty} xdF(x)}{r+s} M_U + b - \lambda(1 - F(\tilde{x}^*))M_U - \lambda(r+s) + \mu = 0 \tag{41}$$

Finally, with respect to V :

$$\frac{\int_{\tilde{x}^*}^{\infty} xdF(x)}{r+s} M_V - \gamma - \lambda(1 - F(\tilde{x}^*))M_V - \mu = 0 \tag{42}$$

Now adding (41) and (42) to eliminate μ, we obtain:

$$\frac{\int_{\tilde{x}^*}^{\infty} xdF(x)}{r+s} (M_U + M_V) + b - \gamma - \lambda(1 - F(\tilde{x}^*)) (M_U + M_V) - \lambda(r+s) = 0$$

or rearranging to solve for λ,

$$\lambda = \frac{\left(\frac{\int_{\tilde{x}^*}^{\infty} xdF(x)}{r+s}\right)(M_U + M_V) + b - \gamma}{r + s + \tilde{\phi}^* M_U + \tilde{\phi}^* M_V} \tag{43}$$

where I have defined

$$\tilde{\phi}^* \equiv 1 - F(\tilde{x}^*)$$

For future reference, we also have

$$\mu = -\gamma + \frac{\left(\int_{\tilde{x}^*}^{\infty} xdF(x)\right) M_V - (b - \gamma)\,\tilde{\phi}^* M_V}{r + s + \tilde{\phi}^* M_U + \tilde{\phi}^* M_V} \tag{44}$$

Now combining this with (40), we have

$$\tilde{x}^* = \frac{\left(\int_{\tilde{x}^*}^{\infty} xdF(x)\right)(M_U + M_V) + (r+s)(b-\gamma)}{r + s + \tilde{\phi}^* M_U + \tilde{\phi}^* M_V} \tag{45}$$

When will the decentralized allocation be efficient?

In the model without entry, we only need $x^* = \tilde{x}^*$ where these two thresholds are given by (37) and (45). (Why?)

In the model with entry, we need two more conditions to ensure optimal entry. To see what these are, note that the planner will add unemployed workers in vacancies up to the point where

$$\frac{dH}{dL} = c_W \text{ and } \frac{dH}{dN} = c_F$$

Thus for full constrained efficiency, we need the following three conditions:

(a)

$$x^* = \tilde{x}^*$$

(b)

$$rJ^U = \frac{dH}{dL} = s\lambda - \mu$$

(c)

$$rJ^V = \frac{dH}{dN} = \mu$$

Now comparing dH/dL and dH/dN as implied from (43) and (44) with (38) and recalling that $p \equiv M/U, \quad q \equiv M/V$, we obtain the following simple conditions for the equilibrium to coincide with the constrained efficient allocation.

(a)

$$M_U + M_V = \beta\frac{M}{U} + (1-\beta)\frac{M}{V}$$

(b)

$$\frac{M_U\left(\int_{\tilde{x}^*}^{\infty} x dF(x)\right)}{r+s+\tilde{\phi}^* M_U+\tilde{\phi}^* M_V}=\frac{\beta p \bar{x}}{r+s+\phi^*(1-\beta) q+\phi^* \beta p}$$

(c)

$$\frac{M_V\left(\int_{\tilde{x}^*}^{\infty} x dF(x)\right)}{r+s+\tilde{\phi}^* M_U+\tilde{\phi}^* M_V}=\frac{(1-\beta) q \bar{x}}{r+s+\phi^*(1-\beta) q+\phi \beta p}$$

First, suppose $M(U,V)$ exhibits increasing returns to scale or decreasing returns to scale. Then (b) + (c) are jointly impossible. Why? Part of the homework exercise...

Next, suppose that $M(U,V)$ exhibits constant returns to scale. Then (a), (b), (c) all hold true if and only if

$$\beta=\frac{M_U \cdot U}{M} \qquad \left(\text{or} \quad 1-\beta=\frac{M_V \cdot V}{M}\right)$$

(This is not obvious, you need to play with the equations to convince yourself).

This is the famous *Hosios condition*. It requires the bargaining power of a factor to be equal to the elasticity of the matching function with respect to the corresponding factor.

What is the intuition?

It is not easy to give an intuition for this result, but here is an attempt: as a planner you would like to increase the number of vacancies to the point where the marginal benefit in terms of additional matches is equal to the cost. In equilibrium, vacancies enter until the marginal benefits in terms of their bargained returns is equal to the cost. So if β is too high, they are getting too small a fraction of the return, and they will not enter enough. If β is too high, then they are getting too much of the surplus, so there will be excess entry. The right value of β turns out to be the one that is equal to the elasticity of the matching function with respect to unemployment (thus $1-\beta$ is equal to the elasticity of the matching function with respect to vacancies, by constant returns to scale).

The acceptance externalities are then easy to understand, since turning down a job is just like entering this economy by paying some cost.

[Important observation: Job Acceptance externalities ((a)) are easier to internalize than entry externalities ((b) + (c)). Why?]

Does the Hosios result imply that the decentralized equilibrium is going to be efficient? Possible, but unlikely unless the planner chooses β.

Other important observations:

- No Scale Effects, unless the matching technology is Increasing Returns to Scale.
- Inefficiencies look more like distorted prices (very neo-classical).

6 Frictions and Investment

In the above model, there are investment-like activities; workers and firms decide to enter before the matching stage. Nevertheless, these are limited to the extensive margin, and somewhat miraculously, the Hosios condition internalizes these externalities. So a naïve intuition might be that it will also internalize investment externalities at the intensive margin.

We now see that this is not the case.

Consider a continuous-time economy with measure 1 infinitely lived risk-neutral workers, and larger continuum of risk-neutral firms, all agents discounting the future at the rate r. This model is based on Acemoglu-Shimer (IER, 1999). For future reference, I will refer to this model as the **search and investment model**.

Production still requires 1 firm - 1 worker, but now there is the intensive margin of capital per worker. In particular, this pair produces $f(k)$, where k is capital per worker. We assume

$$f' > 0, \quad f'' < 0$$

The most important feature is that k is to be chosen ex ante and is irreversible. The important economic implications of this are two:

1. If there is bargaining, at this stage of bargaining, the capital is already sunk and the capital to labor ratio is irreversibly determined.

2. While looking for a worker, the firm incurs an opportunity cost equal to be user cost of capital times the amount of capital that has, i.e., $u_k \times k$, where u_k is the user cost which will be determined below.

Trading frictions will be modeled in a way similar to before, but since my interest here is with "inefficiency", which is easily possible with increasing or decreasing returns to scale in the matching technology, I will assume constant returns to scale from the beginning. I will also develop the notation that will be useful when we look at wage posting and directed search.

First note that if $M = M(U, V)$ exhibits constant returns to scale, then exploiting the standard linear homogeneity properties, we can write

$$\begin{aligned} q &= \frac{M}{V} = M\left(\frac{U}{V}, 1\right) \\ &= q(\theta) \end{aligned}$$

where $\theta \equiv V/U$ is the tightness of the labor market (the vacancy to unemployment ratio), and the function $q(\theta)$ is decreasing in θ given our assumptions above. This means that vacancies have a harder time finding matches in a tighter labor market.

This is the standard notation in the Diamond-Mortensen-Pissarides macro search models.

Moreover,

$$\begin{aligned} p &= \frac{M}{U} = \frac{V}{U} M\left(\frac{U}{V}, 1\right) \\ &= \theta q(\theta) \end{aligned}$$

where $\theta q(\theta)$ is increasing in θ. This means that unemployed workers have an easier time finding matches in a tighter labor market.

Now let us develop a slightly different notation. Assume that if there are Q workers searching for 1 job (think of the analogy to *queues*), Q is equivalent to $1/\theta$ in the above notation.

Then with constant returns to scale, we have

$\mu(Q)$: flow rate of match for workers, assumed it is continuously differentiable and $\mu' < 0$

$\eta(Q) \equiv Q\mu(Q)$: flow rate of match for vacancy, with $\eta' > 0$

The fact that μ, η are simply functions of Q is equivalent to assuming **Constant Returns to Scale.**

As before let r be the rate of time preference, and s be the separation rate due to destruction of capital

Here let us change the order a little, and start with the efficient allocation, which is again a solution to the planner's problem subject to the search constraints.

The objective function of the planner can be written as:

$$\int_0^\infty e^{-rt} \left[\underbrace{\left(\mu(Q_t) \frac{f(k_t) - (r+s)k_t}{r+s} \right) u_t}_{\text{net output of a matched worker}} - \underbrace{(r+s)k_t \frac{u_t}{Q_t}}_{\text{cost of unfilled vacancies}} \right] dt$$

where u_t is the measure of unemployed workers, or alternatively the unemployment rate, at time t.

Here it is easy to see that $(r+s)k$ is the flow cost of investment, or user cost of capital, k. (k paid up front and rk opportunity cost, sk cost of destruction). The planner incurs this cost for $V_t = u_t/Q_t$ vacancies

Less obvious at first, but equally intuitive is that the value of an unemployed worker is that with probability $\mu(Q_t)$ he will find a job, in which case he will produce a net output of $f(k_t) - (r+s)k_t$, until the job is destroyed, which has discounted value $\frac{f(k_t)-(r+s)k_t}{r+s}$, thus the value of an unemployed worker is

$$\mu(Q_t)\frac{f(k_t)-(r+s)k_t}{r+s}.$$

This expression already imposes that all firms will choose the same capital level, and no segmentation in the market (Homework exercise: setup and solve this problem when the planner allows firms to choose different levels of capital).

The constraint that the planner faces is very similar to the flow constraints we saw above:

$$\dot{u}_t = s(1-u_t)-\mu(Q_t)u_t$$

This equation says that the evolution of unemployment is given by the flows into unemployment, $s(1-u_t)$, and exits from unemployment, i.e., job creation, $\mu(Q_t)u_t$.

Now we can write the Current Value Hamiltonian as

$$H(k,Q,u,\lambda) = u\left[\mu(Q)\left(\frac{f(k)}{r+s}-k\right)-\frac{(r+s)k}{Q}\right]+\lambda\left[s(1-u)-\mu(Q)u\right]$$

The necessary conditions are

$$\begin{aligned}
H_k &= u\left(\mu(Q)\left(\frac{f'(k)}{r+s}-1\right)-\frac{(r+s)}{Q}\right)=0\\
H_Q &= u\left(\mu'(Q)\left(\frac{f(k)}{r+s}-k-\lambda\right)+\frac{(r+s)}{Q^2}k\right)=0\\
H_u &= \mu(Q)\left(\frac{f(k)}{r+s}-k\right)-\frac{(r+s)}{Q}k-\lambda(s+\mu(Q))=r\lambda-\dot{\lambda}
\end{aligned}$$

Again, focusing on steady state, we impose

$$\dot{\lambda}=0$$

$$H_u = r\lambda \Longrightarrow \lambda = \frac{\mu(Q)\left(\frac{f(k)}{r+s} - k\right) - \frac{(r+s)}{Q}k}{r+s+\mu(Q)}$$

which is the shadow value of an unemployed worker. This equation has a very intuitive interpretation. The shadow value of a worker is given by the probability (flow rate) that he will create a job, which is $\mu(Q)$, and the value of the job is $\left(\frac{f(k)}{r+s} - k\right)$. While unemployed, the worker induces the planner to have more vacancies open (so as to keep Q constant), hence the term $-\frac{(r+s)}{Q}k$. Finally, once the job is destroyed, which happens at the rate s, a new cycle begins, at the rate $\mu(Q)$, which gives the denominator for discounting.

The condition that $H_k = 0$ gives

$$\Longrightarrow \frac{Q^S\mu(Q^S)f'(k^s)}{(r+s)(r+s+Q^S\mu(Q^S))} = 1 \tag{46}$$

Now combining this and the value of λ obtained about with $H_u = 0 \Longrightarrow$

$$f(Q^S)\frac{\mu'(Q^S)}{r+s} + \frac{r+s+\mu(Q^S)+Q^S\mu'(Q^S)-(Q^S)^2\mu'(Q^S)}{(Q^S)^2}k = 0 \tag{47}$$

Conditions (46) and (47) characterize the constrained efficient allocation.

Next, consider the equilibrium allocation. With bargaining this corresponds to:

$$\begin{aligned} rJ^F(k) &= f(k) - w(k) - sJ^F(k) \\ rJ^V(k) &= \eta(Q)(J^F(k) - J^V(k)) - sJ^V(k) \end{aligned}$$

Recall that there is random matching, so Q workers for each vacancy. Then

I can write

$$\begin{aligned} rJ^E(k) &= w(k) + s(J^U - J^E(k)) \\ rJ^U &= \mu(Q) \int a(k)(J^E(k) - J^U) dF(k) \end{aligned}$$

where $a(k)$ is the decision rule of the worker on whether to match with a firm with capital k, and $F(k)$ is the endogenous distribution of capital (please do not confuse this with f which is the production function).

Nash Bargaining again implies:

$$(1-\beta)(J^E(k) - J^U) = \beta(J^F(k) - J^V(k))$$

Now we will impose free entry as in the basic Mortensen-Pissarides models, so

$$J^V(k) - k = 0$$

That is, opening a job costs k (the sunk investment), and has a return of $J^V(k)$.

$$\Longrightarrow w(k) = \beta\left(f(k) - (r+s)k\right) + (1-\beta)rJ^U$$

Now use this wage rule with J^V and J^F

$$J^V(k) = \frac{\eta(Q)\left((1-\beta)f(k) + \beta\left(r+s\right)k - (1-\beta)rJ^U\right)}{(r+s)(r+s+\eta(Q))} \tag{48}$$

Also recall that $\eta(Q) = Q\mu\left(Q\right)$.

How is the capital-labor ratio chosen? Firms will clearly choose it to maximize profits: that is,

$$k \text{ maximizes } J^V(k) - k.$$

Since this is a strictly concave problem, this implies that all firms will choose the same level of capital, k^B

$\Longrightarrow$

$$F(k) \text{ is a degenerate distribution with all of its mass at } k^B$$

where

$$\frac{\eta(Q^B)(1-\beta)f'(k^B)}{(r+s)(r+s+(1-\beta)\eta(Q^B))} = 1 \tag{49}$$

with Q^B as the equilibrium queue length in the economy.

Now use (48) with J^V and J^E to obtain an equation determining Q^B.

$$\frac{\eta(Q^B)(1-\beta)f(k^B)}{r+s} = \left(r+s+(1-\beta)\eta(Q^B)+\beta\mu(Q^B)\right)k^B \tag{50}$$

The equations (49) and (50) characterize the equilibrium, and can be directly compared to the conditions (46) and (47) for the efficient allocation.

First, compare k^S to k^B: we can see that for all $\beta > 0$, $k^B < k^S$. In other words, there will be underinvestment as long as workers have ex post bargaining power. This is a form of **holdup**, in the sense that the firm makes an investment and the returns from the investments are shared between the worker and the firm. Because the investment is made before there is a match, there is no feasible way of contracting between the worker and the firm in order to avoid this holdup problem.

Thus the only way of obtaining efficiency is to set $\beta = 0$.

What about Q^S versus Q^B?

To compare Q^S versus Q^B, let $f(k^B) = f(k^S)$, then we obtain

$$\beta = \beta^*(Q) \equiv \frac{\eta'(Q)Q}{\eta(Q)} \equiv 1 + \frac{\mu'(Q)Q}{\mu(Q)},$$

is necessary and sufficient for $Q^S = Q^B$.

In other words, with $f(k^B) = f(k^S)$, we are back to the model without investment, so all we need is the Hosios condition for efficiency.

$$M = \mu \cdot U \Longrightarrow M_U = \mu' Q + \mu,$$

$$\Longrightarrow \frac{M_U U}{M} = 1 + \frac{\mu' Q}{\mu},$$

which can be verified as the Hosios condition in this case.

Thus when $f(k^B) = f(k^S)$, the Hosios condition is necessary and sufficient for efficiency.

This is not surprising, since with $f(k^B) = f(k^S)$, the economy is identical to the one with fixed capital.

The key question is whether it is possible to ensure both $f(k^B) = f(k^S)$ and $Q^S = Q^B$ simultaneously.

Of course, from the analysis the answer is no.

If $\beta > 0$, hold-up problem and $k^S > k^B$

If $\beta = 0$, the excessive entry of firms $Q^B < Q^S$.

Theorem 14 *Constrained efficiency is impossible with ex ante investments and ex post bargaining.*

The intuition is quite straightforward: as long as $\beta > 0$, there is rent sharing on the marginal increase in productivity, thus hold-up. But $\beta = 0$ is inconsistent with optimal entry.

7 Two-sided Investments

The problem becomes even worse when reconsider two-sided investments.

Consider the following two-period model based on Acemoglu (QJE, 1996).

$t = 0$ Invest

$t = 1$ Match and Produce

Two types of risk-neutral agents $\nearrow$ Firms ____ j ; $\searrow$ Workers ___ i

"Leontieff" technology

Firms and workers undertake investments, and then search for partners.

As in the above model of ex ante investment and search, theThe crucial assumption is one of anonymity. At the time investments are undertaken, there is anonymity, and this leads to an "incompleteness of contracts": at the search stage, workers cannot write contracts with their future partners (it is not known who their future partners will be!)

Worker i and firm j produce output

$$y_{ij} = Ah_i^{\alpha}k_j^{1-\alpha}$$

where h_i is the human capital of the worker, and k_i is the physical capital of the firm.

Notice that there is also a Leontieff assumption built in this production technology, since each firm can at most employ 1 worker.

Workers maximize their utility given by

$$V_i\left(c_i, h_i\right) = c_i - \frac{1}{\delta_i} \frac{h_i^{1+\Gamma}}{1+\Gamma}$$

Notice that the cost of human capital investment varies across workers as captured by the parameter δ_i.

Cost of capital to firms is denoted by μ, and firms are assumed to maximize profits.

I assume that each firm always meets a worker, and each worker always meets a firm. The frictional nature of the labor market will be modeled by assuming that this meeting is random, and firms and workers cannot change partners after this meeting (alternatively, they can only change partners at a cost, for example foregone output during the search process).

This is a short-cut one-period way of modeling search frictions.

An allocation in this economy is:

1. Investment functions for workers and firms.

2. Allocation of firms to workers.

3. Wage and rental return functions, $w\left(h\right)$ and $r\left(k\right)$, which determine the (expected) payments to workers and firms.

7.1 The Walrasian Allocation

First, consider the Walrasian equilibrium. Here the price functions are $w\left(h\right)$ and $r\left(k\right)$ (why are these not conditioned on who the worker or the firm produces with?)

These functions will be given by marginal products (conditional on equilibrium matching), and workers and firms will be allocated "assortatively," in the sense that high human capital firms will be matched with high physical capital workers.

[If we want to be more formal, we could define efficient allocation as follows

$$\begin{aligned} \Omega_w(i) &= \int_{s\in S_w : h_s > h_i} \\ \Omega_F(j) &= \int_{s\in S_F : k_s > k_i} \end{aligned}$$

where S_w is the set of workers looking for a match, and $\Omega_w(i)$ and $\Omega_F(j)$ give the rat coworkers and firms.

Define the relation P such that $(i, j) \in P$ (matched together) if and only if

(i) $\Omega_w(i) = \Omega_F(j)$

or

(ii) $\Omega_w(i) > \Omega_F(j)$, but $\forall i^*$s.t.

$\Omega_w(i^*) < \Omega_w(i)$; $(i^*, j^*) \in P$

$\Longrightarrow \Omega_F(j^*) \leq \Omega_F(j)$

(iii) $\Omega_w(i) < \Omega_F(j)$, but $\forall j^*$ s.t.

$\Omega_F(j^*) < \Omega_F(j)$; $(i^*, j^*) \in P$

$\Longrightarrow \Omega_w(i^*) \leq \Omega_w(i)$

This formalizes what we mean by efficient matching; basically equivalents to "assortative" matching...]

Taking prices as given, firms maximize:

$$Ah^{\alpha}k^{1-\alpha} - w(h)\,h - \mu k.$$

The first-order condition for the firm is $(1-\alpha)\,Ah^{\alpha}k^{-\alpha} = \mu$. Notice that the firm is taking $w(h)$ as a price schedule, independent of its own investment decision, which is a feature of a competitive/Walrasian economy.

The first-order condition for profit maximization then implies:

$$\left(\frac{h_i}{k_j}\right) = \left(\frac{\mu}{(1-\alpha)\,A}\right)^{1/\alpha}$$

In addition, worker choices are given by

$$h_i = \left[\alpha\,(1-\alpha)^{\frac{1-\alpha}{\alpha}}\,A^{\frac{1}{\alpha}}\mu^{-\frac{1-\alpha}{\alpha}}\delta_i\right]^{\gamma} \tag{51}$$

where $\gamma \equiv 1/\Gamma$.

The Walrasian equilibrium is unique and efficient. Moreover, the factor distribution of income is purely determined by technology—the share of capital income in national product is simply $1-\alpha$.

Finally, the rate of return on human capital is independent of the distribution of human capital investment costs, $\{\delta_i\}$. In fact, more generally, the return on human capital will be decreasing in $\{\delta_i\}$ (in the sense that as δ's increase and workers invest more in human capital, the rate of return on human capital will decline).

This can be captured here by assuming that the cost of capital faced by firms is a function of the aggregate demand for physical capital investments, i.e., $\mu(K)$ where $K = \int k_j dj$.

Now as δ_i's increase, the demand for physical capital by firms increases (at the same cost of capital, they would like to hire more capital to keep the physical to human capital ratio constant). As K increases, the cost of capital rises and the desired physical to human capital ratio falls.

7.2 The Frictional Allocation

Next, consider the equilibrium without the Walrasian auctioneer. Here workers and firms first make investments, and then match. There are two major assumptions:

1. Matching is random. This approximates a situation in which there are frictions in finding the right production partner. Random matching means that two workers with different human capital levels have the same probability of meeting with each firm (the alternative would be a situation in which high human capital workers are somehow more likely to be matched with high physical capital firms).

2. Changing partners is costly. In particular, in the two period model here, this corresponds to no switching of partners.

These assumptions imply that once a worker and a firm meet, there are quasi-rents in the relationship that need to be shared. I simply assume that the worker receives a fraction β of the output, i.e.,

$$W_i = \beta y_{ij} \tag{52}$$

(this can be derived as the equilibrium bargaining rule in the presence of a la Rubinstein outside option bargaining, see Acemoglu, QJE 1996).

Then, workers and firms will choose their investments to maximize their expected returns given by

$$\begin{aligned} EW(h_i, \{k_j\}) &= \beta A h_i^{\alpha} \left(\int k_j^{1-\alpha} dj \right) \\ ER(k_j, \{h_i\}) &= (1-\beta) A k_j^{1-\alpha} \left(\int h_i^{\alpha} di \right) \end{aligned}$$

First, suppose that $\delta_i = \delta_1$ for all i. Then, since $EW(h_i, \{k_j\})$ is strictly concave, all workers will choose the same level of human capital investments, and the first-order conditions for the firms and the workers give

$$\begin{aligned} (1-\beta)(1-\alpha) A k^{-\alpha} h^{\alpha} &= \mu \\ \beta \alpha A k^{1-\alpha} h^{-(1-\alpha)} &= \frac{h^{\Gamma}}{\delta_1} \end{aligned} \tag{53}$$

Solving these equations together, we obtain the equilibrium human capital investment as

$$h_i = h_R \equiv \left[\alpha\beta \left((1-\alpha)(1-\beta) \right)^{\frac{1-\alpha}{\alpha}} A^{\frac{1}{\alpha}} \mu^{-\frac{1-\alpha}{\alpha}} \delta_1 \right]^{\gamma} \tag{54}$$

Comparison of (54) with (51) shows that human capital investments are always inefficient in the random matching economy. For example, when

$$\begin{aligned} \beta &= 0 \text{ we have } h_i = 0 \\ \beta &= 1 \text{ we have } k_j = 0\text{, and therefore } h_i = 0 \end{aligned}$$

Intuitively, in contrast to the competitive economy where wages are taken parametrically by firms, here firms anticipate that by increasing their investments they will also increase the wages they pay to workers. Similarly, workers understand that they are not the full residual claimant of the increase in output due to their human capital investments.

This feature arises because of the incompleteness of contracts induced by matching. If a worker knew which firm he would be matched with, he could write a binding contract with that firm, specifying a particular division of the surplus, encouraging greater investment by both parties. The problem is that frictional allocation means that it is not known who will be matched with whom. Such contracts are therefore impossible.

Using (53) and (54), output in this economy is obtained as

$$Y = A\left(\frac{(1-\beta)}{\mu}\right)^{1/\alpha}\left[\alpha\beta\left((1-\alpha)(1-\beta)\right)^{\frac{1-\alpha}{\alpha}} A^{\frac{1}{\alpha}}\mu^{-\frac{1-\alpha}{\alpha}}\delta_1\right]^{\gamma}$$

This output level is maximized when

$$\beta = \frac{\gamma\alpha}{1-\alpha+\gamma},$$

but this maximized value of output is less than the Walrasian output level.

It is important to note that the market structure (institutional structure) as captured by β now matters for physical and human capital investments. This differs from the Walrasian equilibrium.

It is also interesting that factor shares in the data now reveal the institutional distribution parameter β, not the production function parameter α. This is also relevant for efforts to deduce the elasticity of output with respect to investment (or distortions) using factor shares.

This framework not only highlights a number of potential determinants of human capital investments, but also leads to "social increasing returns" in human capital. To see this suppose that there are two types of workers, a fraction ρ of type δ_2 and a fraction $1-\rho$ of type $\delta_1 < \delta_2$. Denote the investments of these two types of workers by h_1 and h_2, and recall that here the δ_2 workers face lower costs of human capital investments, so they can be thought as the "high type" workers.

Now, the expected return function of firms becomes

$$ER\left(k_j, \{h_i\}\right) = (1-\beta)\, k_j^{1-\alpha}\left((1-\rho)\, h_1^{\alpha} + \rho h_2^{\alpha}\right)$$

Again this return function is strictly concave, so all firms will choose a unique physical capital level, k.

The first-order conditions of workers is then unchanged, and gives

$$\beta\alpha A\left(\frac{k}{h_i}\right)^{1-\alpha} = \frac{h_i^{\Gamma}}{\delta_i}$$

Using this, we obtain the physical capital to human capital ratio of the low type workers

$$\frac{k}{h_1} = \left[(1-\alpha)(1-\beta) A \mu^{-1} \left((1-\rho) + \rho \left(\frac{\delta_2}{\delta_1} \right)^{\frac{\alpha\gamma}{1+\gamma-\alpha\gamma}} \right) \right]^{\frac{1}{\alpha}}$$

The important result here is that

$$\frac{\partial \frac{k}{h_1}}{\partial \rho} > 0$$

That is, as the fraction of high type workers in the population increases, the physical to human capital ratio at which low type workers are employed also increases. Intuitively, when there are more high type workers, firms increase their investments anticipating to match with these workers sometime. However, because of random matching (the frictional nature of the economy), they often match with less skilled (low type) workers, and end up bringing a greater amount of physical capital to the relationship, increasing the output, and hence the wages that these workers receive. This result contrasts sharply with the Walrasian economy where this ratio is either constant or decreasing.

Moreover, given the wage determination rule (52), W_1, the earnings of low type workers (as well as W_2 the earnings of high type workers) are increasing in ρ. *That is, as the average quality of workers in the population increases, the earnings of all workers increase.*

This is the sense in which the model with frictional assignment of workers to firms features social increasing returns.

It is also straightforward to check that $W_1'(h_1)$ is increasing in ρ. Therefore, human capital investment decisions of workers are strategic complements. Each worker wants to invest more when others do so. We will see next that this can lead to multiple equilibria: in one equilibrium, firms and workers invest a lot in physical and human capital, whereas in the other one there is low investment.

7.3 Multiplicity of Equilibria

It is also straightforward to see that if instead of continuous choices, firms have some "nonconvexity" in their choices, multiple equilibria are possible because of the two-sided investments.

Firms choose between two technologies $\begin{array}{l} \nearrow k_1 \\ \searrow k_2 \end{array}$

Except for this, assume exactly the same structure is above, in particular assume that $W_i = \beta y_{ij}$, and all workers have cost δ. Let $\theta \equiv 1/\left(1 - \alpha + \Gamma\right)$.

When workers expect all firms to choose k_1, is optimal for them to choose

$$h_1 = \left[\beta\alpha\delta A k_1^{1-\alpha}\right]^{\theta},$$

whereas when all firms invest $k_2 > k_1$, workers choose

$$h_2 = \left[\beta\alpha\delta A k_2^{1-\alpha}\right]^{\theta} > h_1.$$

In turn, when workers invest h_2, k_2 is relatively more profitable for the firms because of the complementarity between physical and human capital.

Consequently, we have the following result

Theorem 15 *Suppose that*

$$\begin{gathered}(1-\beta)Ak_2^{1-\alpha}\left[\alpha\beta\delta Ak_2^{1-\alpha}\right]^{\alpha\theta}-\mu k_2\\ >(1-\beta)Ak_1^{1-\alpha}\left[\alpha\beta\delta Ak_2^{1-\alpha}\right]^{\alpha\theta}-\mu k_1\\ \text{and}\\ (1-\beta)Ak_2^{1-\alpha}\left[\alpha\beta\delta Ak_1^{1-\alpha}\right]^{\alpha\theta}-\mu k_2\\ <(1-\beta)Ak_1^{1-\alpha}\left[\alpha\beta\delta Ak_1^{1-\alpha}\right]^{\alpha\theta}-\mu k_1\end{gathered}$$

then there exist two symmetric Nash Equilibria, in one firms choose k_1 and workers choose h_1, in the other firms choose k_2 and workers choose h_2.

We can also note that these equilibria are Pareto ranked, and the equilibrium with low investment has "underinvestment". So because of coordination failure, it is possible for all agents to the low level of investment.

Digression: coordination failure in competitive models. Is that possible?

And other important point to notice that there is multiplicity of equilibria here without scale effects. This contrasts with the Diamond Coconut model for example, where scale effects were at the root of the multiplicity.

7.4 Frictions and Overinvestment

The above analysis suggests that there will generally be underinvestment because of holdup problems created by search frictions. This is in fact not the case. Depending on the exact specification, overinvestment is also possible.

To illustrate this, take the above model, but modify the above model with two levels of investment, such that after the first match workers and firms can separate and look for another partner at some cost $\varepsilon > 0$ but small. This is in fact the model analyzed in Acemoglu (QJE, 1996).

All of the results discussed above are true when ε is sufficiently large, and all of the results in which workers are homogeneous and investment $k \in \mathbb{R}_+$ are true for all $\varepsilon > 0$ as well.

However, with the two levels of investments, k_1 and k_2, the limit $\varepsilon \to 0$ gives very different results.

Here I sketch the analysis. Assume that the lower level of investment $k_1 < k_2$ is optimal (i.e., would be chosen by a social planner).

It can be shown very easily that when all firms undertake this level of investment, all workers will choose human capital $h_1 = \left[\beta\alpha\delta A k_1^{1-\alpha}\right]^\theta$ and this is an equilibrium.

However, now imagine a situation in which all firms choose k_2 and all workers choose $h_2 = \left[\beta\alpha\delta A k_2^{1-\alpha}\right]^\theta$. If a firm deviates and chooses k_1, then bargaining with parameter β would imply that wages will be $\beta\times$output:

$$\beta k_1^{1-\alpha}\left[\beta\alpha\delta A k_2^{1-\alpha}\right]^\theta,$$

but when workers know that they can leave this firm, and by paying a small cost ε they can meet with a firm with capital k_2, their outside option is

$$\beta k_2^{1-\alpha}\left[\beta\alpha\delta A k_2^{1-\alpha}\right]^\theta > \beta k_1^{1-\alpha}\left[\beta\alpha\delta A k_2^{1-\alpha}\right]^\theta,$$

so the firm cannot simply pay a fraction β of the output; if it wants to keep the worker, it will be forced to pay this higher wage, making the lower but socially-optimal level of investment less profitable. Because of this, it is possible to support over-investment in equilibrium.

8 Burdett-Judd Model of Price Dispersion

So far we have analyzed search models without an explicit distribution of wages out of which agents sample. The Rothschild critique above suggests that this may be difficult to obtain. In fact, there are many models which to generate explicit wage distributions, and thus can be thought of as microfoundations for the standard McCall search model. We will do this in a number of different ways.

I start with the very creative model by Burdett and Judd (Econometrica, 1983) adapted to the labor market framework. We will see later that there are certain features of this model that do not make it ideal for the labor market. A key feature of this model and next one we will discuss is that search is now **directed**, in the sense that at least some of the workers will be able to decide which types (levels) of wages to seek.

For all practical purposes, we can focus on a static economy, but do not need to. The economy consists of a large countable number of workers, N and a large countable number of firms, M. Let $\mu = N/M$ be the number of workers per firm. All agents are risk neutral, and discounts the future at the rate δ. Firms can produce 1 unit of output per worker, in other words, each firm has access to a constant returns to scale production technology of the form

$$y = l,$$

where l is the number workers employed. Once a worker agrees with the firm, they produce forever.

If unemployed, workers receive zero wages.

The game is as follows:

- All firms simultaneously post a wage w, generating a wage distribution denoted by $F(w)$.
- Each worker decides how many wages from the wage distribution to sample, denoted by n_j for worker j.
- Each worker applies to the highest wage he has sampled.
- All firms accept all workers and employ them at the promised wage.

We will derive an equilibrium in which there is a non-degenerate distribution of wages $F(w)$.

Let us first take the sampling decisions of workers as given. This can be thought of as a discrete random variable with distribution $\langle q_n \rangle_{n=1}^{\infty}$.

Rather than analyze the full equilibrium, I'm going to sketch the stationary equilibrium, where the aggregate wage distribution is constant over time. From what we have seen so far, we know that workers will use an optimal stopping rule: If $\max_{n_t} w_{n_t} > R$, accept a job and work forever, otherwise search again.

Let $\Pi(w)$ denote the discounted net present value of profits for a firm offering wage w at time 0. I'm going to focus on an equilibrium where each firm always offers the same wage, which is without loss of any generality in a stationary equilibrium.

Then

$$\Pi(w) = \pi(w) + \beta\Pi^c(w),$$

where $\pi(w)$ is profits made from workers hired now,

$$\pi(w) = \frac{1-w}{1-\beta} l_0$$

with l_0 as the number of workers hired by this firm in the initial period, and $\Pi^c(w)$ is the continuation value, which is unaffected by what happens in this period, since the firm is small.

We can define a "partial" equilibrium as follows:

Definition 5 *Given* $\left(\langle q_n\rangle_{n=1}^{\infty}, R\right)$ *a partial equilibrium is a pair* $\langle F(\cdot), \pi\rangle$, *where* $F(w)$ *is the distribution of wages, such that*

1. $\pi(w) = \pi \quad \forall w \in$ *support of* $F(w)$
2. $\pi(w) \leq \pi \quad \forall w.$

Why is this the right definition of equilibrium?

Now we can prove:

Lemma 3 *If* $q_1 = 1$, *then* $F(w)$ *is degenerate at* $w = 0$ *(Diamond's Paradox).*

Proof. (sketch): Suppose R is used as the reservation wage by the workers. Then $w > R$ is dominated for firms $\Longrightarrow w = R$. Suppose $R > 0$, then set $w = R - \varepsilon$, for ε small enough it is optimal for workers to accept this wage $\Longrightarrow R > 0$ not an equilibrium. That $R = 0$ is an equilibrium is straightforward to check. ■

But more importantly:

Lemma 4 *If* $(\langle q_n \rangle_{n=1}, R)$ *is s.t.* $q_1 \neq 1$*, then* $F(\cdot)$ *is either degenerate at* 1 *or has connected support and is continuous.*

Proof. (sketch):

1. **continuous:** Suppose at $\hat{w}$ the distribution function of wages is discontinuous, so that $F(\hat{w}+) > F(\hat{w}-)$. Since $q_1 \neq 1$, the probability that there will exist a worker observing two wages exactly equal to $\hat{w}$, $\text{Prob}(n = 2, w_1 = w_2 = \hat{w}) > 0$. This implies that $\pi(\hat{w} + \varepsilon) > \pi(\hat{w})$. Hence a contradiction. $F(w)$ must be continuous.

2. **connected:** Suppose $F(\cdot)$ is constant over an interval $[w_1, w_2] \Longrightarrow \pi(w_2 - \varepsilon) > \pi(w_2)$ as long as $w_2 - w_1 > \varepsilon$, since $w_2 - \varepsilon$ will be accepted by all workers accepting the wage w_2, but will have lower labor costs. Contradiction.

■

Given the structure, we can write

$$\pi(w \mid F(w)) = \begin{cases} (1-w)\mu \sum_{k=1}^{\infty} q_k k F(w))^{k-1} & w \geq R \\ 0 & w < R \end{cases}$$

This leased to the main theorem:

Theorem 16 *$\exists$ three types of equilibrium:*

1. *$q_1 = 1,\ R = 0,\ \pi = \mu$.*

2. *$q_1 = 0$, the labor market is competitive with "competitive" $w = 1$.*

3. *$0 < q_1 < 1$ and there is a nondegenerate wage distribution with $\bar{w} < 1$, $F(\cdot)$ is continuous over $[\underline{w}, \bar{w}]$, with*

$$\pi = \mu q_1(1 - \underline{w}) = \mu(1 - \bar{w}) \sum_{n=1}^{\infty} n q_n > 0,$$

and

$$(1 - w) \sum_{k=1}^{\infty} q_k k F(w))^{k-1} = q_1(1 - \underline{w})$$

for all $w \in [\underline{w}, \bar{w}]$. Why is this?

Therefore there is now **Equilibrium Wage Dispersion.**

Burdett and Judd also show that in equilibrium we will have only $q_1, q_2 > 0$, but I leave the proof of this to the next model, which also has a result like this.

The problem with the Burdett-Judd model is that firms have no capacity constraints, so on unemployment, and there is no entry margin, things we want for a model of the labor market. We now tackle this question.

9 Wage Dispersion in Search Equilibrium

I will now discuss the model of Acemoglu-Shimer (RES 2000), which can be viewed as a generalization of Burdett-Judd to an equilibrium labor market setting. One of the main purposes of this model is that it will link models of wage dispersion to models of directed search.

I will start with the simplest version, which has no capital and is static. We will see the elaborations below (which are quite important for the labor market context).

The description of the environment is as follows: there is a continuum 1 of risk-neutral workers and a much larger continuum $M \gg 1$ of risk-neutral firms. The sequence of events is as follows.

1. Each firm decides whether to create a job. If it creates a job, it chooses a level of capital $k \geq 0$ at cost rk and posts a wage w. I will start with the case in which $k = 1$ by assumption.

2. Each worker decides how many jobs to sample, her 'search intensity'. A worker who samples n jobs learns the wage offered by n randomly chosen active firms and applies to at most one of them. She pays a marginal search cost c_n for sampling the n^{th} job.

3. Each firm that receives at least one application chooses randomly one of the applicants, pays her the posted wage, and produces. The remaining applicants are unemployed, while active firms that receive no

applications are vacant. A worker's payoff is equal to her wage net of search costs if she is employed, and she loses her search cost otherwise. An active firm earns $f(k) - w - rk$ if it hires a worker, and loses its investment rk if it is vacant. An inactive firm earns nothing.

With the restriction to $k = 1$, I also impose $f(1) = 1$.

We will see later that it is straightforward to extend this model to a dynamic setting.

Exactly as in Burdett and Judd workers rationally anticipate the equilibrium wage distribution, but they do not know the wage offered by a particular firm. Because all firms appear to be identical, a worker randomly samples $n \geq 1$ of them. We assume that all workers costlessly observe one wage, $c_1 = 0$. The marginal cost of sampling the n^{th} job, c_n, is positive and increasing.

Each firm j makes a capital investment $k(j)$ and posts a wage rate $w(j)$. Let $x(j) \equiv (k(j), w(j))$ denote its strategy. We adopt the convention that an inactive firm chooses $x(j) = (0, 0)$. The strategies of all firms are summarized by a mapping $\mathbf{x} : [0, M] \rightarrow \mathbb{R}^2_+$.

For now, we can think of this mapping as

$$\mathbf{x} : [0, M] \rightarrow \mathbb{R}_+$$

since the wage is the only choice variable.

Each worker i chooses a sample size $n(i)$ and a preference function

$$p_n(i) : \mathbb{R}^n \mapsto \{0, 1, \ldots, n\}$$

over arbitrary n-tuples of wages. Thus $p_n(w_1, \ldots, w_n) = s \in \{1, \ldots, n\}$ tells us that a worker who observes wages $(w_1, \ldots, w_n)$ will apply for the job offering w_s. $p_n(w_1, \ldots, w_n) = 0$ denotes that the worker chooses to remain unemployed. Workers' strategies are summarized by the mapping $\mathbf{y} : [0, 1] \rightarrow \mathcal{F}$, where $\mathcal{F}$ represents the space of all preference functions with $n \geq 1$. From this we can extract each worker's sample size and preference function.

An equilibrium is straightforward to define as a $(\mathbf{x}^*, \mathbf{y}^*)$ such that each firm j chooses its capital-wage pair $x(j)$ to maximize its expected profit given the strategies of other players, $(\mathbf{x}^*, \mathbf{y}^*)$, and each worker i chooses her sample size and preference function to maximize her expected wage net of search costs given $(\mathbf{x}^*, \mathbf{y}^*)$.

Search frictions come from "coordination problems" in the sense that more than one worker may apply to a given job, which can at most employ one worker, so some of the applicants will remain unemployed. This is modeled with the assumption that workers use **anonymous matching strategies**, meaning that we cannot coordinate so that I apply to certain jobs while other workers apply to certain other jobs. This immediately implies an urn ball type technology (analogy to throwing a number of balls blindly into a number of urns).

As in all search models, it is convenient in what follows to define an equilibrium in terms of aggregate variables rather than individual strategies. Let $\pi(w)$ be the expected profit of a firm choosing wage w. Let V be the measure of active firms, and H denote the joint distribution of those firms' capital investments and wages, with support X. Also let H^w denote the distribution of wages with support $\mathcal{W}$. On the worker's side, let R_n denote the expected utility of a worker who chooses a sample size of n, and λ_n be the fraction of workers who sample $n \in \{1, 2, \ldots\}$ jobs, with $\Sigma_{n=1}^{\infty} \lambda_n = 1$. Let $\rho(w)$ denote the expected return of a worker who applies for a job offering wage w.

Let $G_n(w)$ denote the probability that a worker who samples n random wages, including one offering w, applies for the one offering w. This function is defined by integrating workers' preference function over the the wage distribution H^w:

$$G_n(w) = \int_{\mathcal{W}} \cdots \int_{\mathcal{W}} I(p_n(w, w_2, ..., w_n)) \, dH^w(w_2) \cdots dH^w(w_n), \tag{55}$$

where I is an indicator function, with $I(1) = 1$ and $I(p) = 0$ for all other p; so it takes value 1 if a worker would choose to apply for the first wage. Let $Q \equiv 1/V$ denote the ratio of workers to vacancies. Each firm therefore expects to be sampled by $Qn\lambda_n$ workers who sample n firms. Then the expected number of applications received by a firm offering a wage w is

$$q(w) = Q \sum_{n=1}^{\infty} n\lambda_n G_n(w). \tag{56}$$

Since each application decision is independent (the anonymous mixed strategy assumption), the firm receives no applications with probability

$$\exp(-q(w)).$$

Suppose that there were m workers. A particular worker applies to a firm offering wage w with probability $q(w)/m$, so the probability that the firm receives no applications is $(1 - q(w)/m)^m \approx \exp(-q(w))$. We have a continuum of workers here, so the last expression is exact.

The expected profit of a firm offering wage w is the probability that it receives an application, times its gross profit if it receives an application, $1 - w$, minus the sunk cost, r:

$$\pi(w) = \big(1 - \exp(-q(w))\big)(1 - w) - r. \tag{57}$$

Similarly, the probability that a worker applying for wage w is hired, is equal to the probability that a firm offering wage w hires a worker, divided by the expected number of workers applying for that wage, $q(w)$. Using this, the expected return to the application is

$$\rho(w) = \frac{1 - \exp\left(-q(w)\right)}{q(w)}\, w \tag{58}$$

Finally, the expected utility of a worker sampling $n \geq 1$ wages randomly drawn from H is

$$R_n = n \int_{\mathcal{W}} \rho(w) G_n(w)\, dH^w(w) - \sum_{i=1}^{n} c_i \tag{59}$$

For each of n wage draws, her expected payoff is equal to the probability that she applies for the wage, $G_n(w)$, multiplied by her expected return, $\rho(w)$, integrated over the density H^w. Summing over the n independent wage draws and subtracting search costs, we obtain the workers' expected utility.

So to review we have a model of "partially directed search", where workers can choose which jobs to apply to among those that they observe. The notation is:

V: measure of firms who are active. cost r

1: measure of workers

Output of one firm + one worker = 1.

$c(n)$: cost of locating n jobs: $c(1) = 0,\ c(2) = c$.

λ_n: prop. of workers locating n jobs.

$H^w(w)$: distribution function of wages over support $\mathcal{W}$

$G(w)$: endogenously determined prob. that a worker who has located two jobs one at wage w and the other at some wage randomly drawn from $H^w(w)$ will apply to wage w.

The problem is that as opposed to the Burdett-Judd model, it's not clear that workers should apply to the highest wage they sample (why not?).

Definition 6 *An equilibrium consists of a measure V of active firms, $H^w(w)$ over $\mathcal{W}$, $\pi(w), \rho(w)$ and search intensity decisions such that:*

1. *(a)* $\forall w \in \mathcal{W},\ \forall w',\ \pi(w) \geq \pi(w')$

1. **Definition 7** *(a)* $\forall w \in \mathcal{W},\ \pi(w) = r$

   (b) A worker applies to w' over w iff $\rho(w') \geq \rho(w)$

   (c) n maximizes R_i.

So an equilibrium we will have Zero-Profit for firms, or mathematically

$$\forall w \in \mathcal{W},\ r = \left(1 - \rho^{-q(w)}\right)(1 - w) \tag{60}$$

Now use (60) to substitute for $q(w)$ in $\rho(w)$:

$$\tilde{\rho}(w \mid r) = \frac{rw}{(1-w)\left\{\log(1-w) - \log(1-w-r)\right\}}$$

This function gives the return to worker who applies at a job at wage w making gross profits r or net zero profits. Then by definition,

$$\begin{array}{ll} \tilde{\rho}(w \mid r) = \rho(w) & \forall\, w \in \mathcal{W} \\ \tilde{\rho}(w \mid r) \leq \rho(w) & \forall\, w \notin \mathcal{W} \end{array}$$

What simplifies the analysis is the observation that $\tilde{\rho}(w \mid r)$ is strictly quasi-concave with maximum at w^*.

Lemma 5 *In equilibrium, a worker always applies to the highest wage she locates; i.e., over the wage distribution support $\mathcal{W}$, $\rho(w)$ is strictly increasing.*

Proof. Suppose $\exists\, w_1 < w_2 \in \mathcal{W}$ and $\rho(w_1) > \rho(w_2)$.

$$\Longrightarrow G(w_1) \geq G(w_2) \Longrightarrow q(w_1) \geq q(w_2) \Longrightarrow \rho(w_1) < \rho(w_2),$$

a contradiction.

Next: $\rho(w)$ cannot be constant over a range since $\rho(w) = \tilde{\rho}(w \mid r)$ and $\tilde{\rho}(w \mid r)$ strictly increasing over $(-\infty, w^*]$. Also since $\rho(\cdot)$ weakly increasing over $\mathcal{W}$, $\mathcal{W} \subseteq (-\infty, w^*]$. ■

We also have:

Proof.

Lemma 6 *If* $\lambda_1 < 1$, *then* $\mathcal{W} \subseteq [0, \bar{w}] \cup \{w^*\}$. $H^w(\cdot)$ *is continuous over* $[0, \bar{w}]$ *and may have an atom at* w^*.

■

The proof of this lemma is very similar to the equivalent lemma for Burdett-Judd, with the only difference that now there can be an atom (mass point) at w^* (why is that?). Also there can be a gap between $\bar{w}$ and w^*, so the entire support is not connected (why is that?). Finally, you should check that you can prove that the lower support has to be 0.

The next lemma shows that in fact an equilibrium not everybody will sample two or more wages.

Lemma 7 *In any equilibrium,* $\lambda_1 > 0$.

Proof. Suppose $\lambda_1 = 0$, then Lemma 6 $\Longrightarrow 0 \in \mathcal{W}$ but $G(0) = 0 \Longrightarrow \pi(0) = 0 < r$. A contradiction. ■

Given this lemma, the rest is straightforward. The next result shows a natural decreasing returns to scale in fixed sample size search:

Lemma 8 *In equilibrium,* $R_2 - R_1 \geq R_{n+1} - R_n \ \forall n \geq 2$.

Now, combining Lemmas 5 and 6 $\Longrightarrow G(w) = \hat{G}(w)$ over $[0, \bar{w}]$, which enormously simplifies the analysis.

This whole thing then leads to:

Theorem 17 *In equilibrium,* $\lambda_1 + \lambda_2 = 1$.

That means, all workers will either sample one or two wages. What is the intuition?

Once we have this, the characterization of equilibrium follows naturally.

Theorem 18 *There always exists an equilibrium with* $\lambda_1 = 1$, *and no search,* $w = 0$ *and* $Q = -\log(1-r)$.

The basic intuition here is that the search intensity decisions of workers are strategic complements; if nobody else search is, firms do not differentiate their wages, so no point in searching. However, there is also a degree of strategic substitutability in search decisions: if everybody searches, then all firms will offer very high wages, so no point for me to search, and this explains the results that $\lambda_1 > 0$.

Now an equilibrium with search features can be characterized. First, for notational convenience define

$$\lambda_1 = \lambda, \lambda_2 = 1 - \lambda.$$

Then we have

$$\begin{aligned} \pi(w) &= (1 - \rho^{-q(\lambda + 2G(w)(1-\lambda))})(1 - w) \\ \pi(w) &= r \end{aligned}$$

This implies

$$G(w) = \frac{\log(1-w) - \log(1-w-r) - q\lambda}{2q(1-\lambda)}$$

Since $w = 0 \in \mathcal{W}$, and $\pi(0) = r$, we also have $\pi(0) = 1 - e^{-q\lambda} = r$, which gives us the type the so the labor market as

$$Q = \frac{-\log(1-r)}{\lambda}$$

where recall that $Q \equiv 1/V$.

Now substitute this into $G(w)$ to obtain an explicit form solution for the distribution of wages.

$$G(w) = \frac{\lambda}{2(1-\lambda)} \left(\frac{\log(1-w-r) - \log(1-w)}{\log(1-r)} - 1 \right)$$

The comparative statics here are sensible:

$\lambda \downarrow$, $G(w)$ shifts up. $\longrightarrow$ More search $\Longrightarrow$ Higher wages.

Also let μ be the mass point at w^*. Then

$$G(w^*) = 1 - \mu/2$$

(why is that?)

Therefore:

$$\begin{aligned} \pi(w^*) &= (1-\rho^{-q(\lambda+(1-\lambda)(1-\mu))})(1-w^*) = r \\ \Longrightarrow & \mu = \max\left\{ \frac{1}{1-\lambda} \left(1 - \lambda \left(\frac{\log(1-w^*-r) - \log(1-w^*)}{\log(1-r)} \right) \right); 0 \right\} \end{aligned}$$

Also by construction $G(\bar{w}) = 1 - \mu$

$$\Longrightarrow \bar{w} = 1 - \frac{r}{1-(1-r)^{(2(1-\mu)+\lambda(2\mu-1))/\lambda}}$$

Finally, $R_2 = R_1$

$$\int_0^{\bar{w}} G(w)(1-G(w))\tilde{\rho}'(w \mid r) + \mu(1-\mu)\left[\tilde{\rho}(w^* \mid r) - \tilde{\rho}(\bar{w} \mid r)\right] = c,$$

which will have a solution for c sufficiently low, and thus a search equilibrium always exists when search costs are sufficiently low.

Important Remark: $\lambda > \underline{\lambda}$. This implies that even if $c = 0$, all workers will *not* search.

So what is happening at the limit $c \longrightarrow 0$?

It can be shown that $G(w) \longrightarrow w^*$.

What is this wage w^*?

This is exactly the wage that satisfies the Hosios condition and achieves constrained efficiency.

This seems mysterious, but we will see that it's not.

One way of uncovering the mystery is that this is also the wage that would have happened if all workers observed two or more wages (clearly this has to be exogenously, since Lemma 7 above shows that workers will not like to do this)—thus this model is somehow limiting to the equilibrium of an economy in which all workers observe sufficiently many wages. We will come back to this next.

But before we do this, what happens if we now reintroduce the capital decisions (i.e., relax $k = 1$)?

Answer: very similar equilibrium to before, but now there is capital/technology dispersion as well as wage dispersion. Why is this?

Also, what happens if we repeat this economy overtime, so that workers lose their jobs, and find new ones etc.?

Answer: exactly the same structure as here. Why is that?

Important: the dynamic version of this model is very close to the Mc-Call model, except that some workers are observing two wages from the distribution, so wage dispersion will continue to exist in equilibrium.

10 Burdett-Mortensen Model of Search and Wage Dispersion

Now let's think of another application of the Burdett-Judd model to a dynamic labor market, which will also be another application of wage posting (and also emphasizes the differences between the two types of "wage posting").

In the models of dispersion we analyzed so far, the crucial ingredient was that some agents end up with two offers in their hands. So far, we thought of this as reflecting a Stigler-type fixed-sample size search by workers. How else can this happen in the labor market?

One possibility is because employed workers also get wage offers. We will see that this will also lead to the wage dispersion, but through a very different mechanism.

In particular, imagine the following environment.

There is a continuum 1 of workers, who can be employed or unemployed.

Think of the workers as in the McCall sequential search world, observing wages from a given distribution (except that, imagine we are in continuous time, so workers see a wage at some flow rate). Moreover assume that both employed and unemployed workers receive wage offers at the flow rate p. An employed worker who receives an offer can leave his job and immediately start at the new job if he so wishes. The important assumption is that the future distribution of job offers and rate of job offers is unaffected by whether the worker is employed or not (this is not an assumption in the original Burdett-Mortensen model, but it simplifies life).

There is a continuum $m < 1$ of firms. They post wages and their wage offers are seen by a worker at the flow rate q. Thus p and q are exactly as in our standard search model, except that they are not what "matching" probabilities but flow rates of a worker seeing a wage, and a wage being seen by a worker. I do not consider free entry, and for simplicity, I will take p and q as exogenous.

Unemployed workers receive a benefit of $b < 1$. Employed workers produce output equal to 1, and there is no disutility of work.

There is exogenous separation at the rate s, and also potentially endogenous separation if workers receive a better wage offer.

Both workers and firms are risk-neutral and discount the future at the rate r.

Wage posting corresponds to a promise by the firm to employ a worker at some prespecified wage untill the job is destroyed exogenously. Workers observe promised wages before making their decisions. Let us denote the offered wage distribution by $F(w)$, and let us restrict attention to steady states, assuming that this distribution is stationary.

First let's look at the search behavior of an unemployed worker. As usual, the worker is solving a straightforward dynamic programming problem, and his search behavior will be characterized by a reservation wage. Moreover, in this case the reservation wage is easy to pin down. Since there is no disutility and accepting a job does not reduce the future opportunities, an unemployed worker will accept all wages

$$w \geq b$$

Let's now look at the behavior of an employed worker, currently working at the wage w_0. By the same reasoning, this worker will take any job that offers

$$w \geq w_0$$

Therefore, firms get workers from other firms that have lower wages and lose workers to exogenous separation and to firms that offer higher wages.

Now with this structure, using exactly the same arguments as we have seen so far, we can immediately establish that there will exist a continuous connected wage distribution over some range $[b, \bar{w}]$.

Why? (I will do this informally, since this repeats what we have already done, and also not distinguishing between offered and actual wage distributions, see below)

First, it is easy to check that $\bar{w} \leq 1$. If $\bar{w} > 1$, the firm would make negative profits. This implies that employing a (one more) worker is always profitable.

Suppose that the wage distribution is not continuous, meaning that there is an atom at some point w'. Then it is a more profitable to offer a wage of $w' + \varepsilon$ than w' for ε sufficiently small, since with positive probability a worker will end up with two wages of w', thus accepting each with probability 1/2. A wage of $w' + \varepsilon$ wins the worker for sure in this case.

Suppose that the wage distribution is not connected, so that there is zero mass in some range (w', w''). Then all firms offering w'' can cut their wages to $w' + \varepsilon$, and receive the same number of workers.

The lower support has to be $\underline{w} = b$. Suppose not, i.e., suppose $\underline{w} > b$. Then firms offering $\underline{w}$ can cut their wages without losing any workers.

Let's now look at the differential equations determining the number of workers employed in each firm and workers in unemployment.

Unemployment dynamics are given by

$$\dot{u} = s(1-u) - pu$$

since workers receive wage offers at the rate p, and all of them take their offers.

Therefore, steady state unemployment is fixed by technology as

$$u = \frac{s}{s+p}$$

However, employment rate of firms is endogenous. Imagine that the equilibrium wage distribution is given by $G(\tilde{w})$ and the offered wage distribution is $F(\tilde{w})$. Let us continue to restrict attention to steady states, where both of those are stationary. It is important that these two are not the same (why?).

Then the level of employment of a firm offering wage w (now and forever) follows the law of motion

$$\dot{N}(w) = q(u + (1-u)G(w)) - (s + p(1 - F(w)))N(w)$$

where the explanation is intuitive; the offer of this firm is seen by a worker at the flow rate q, and if he is unemployed, which has probability u, he takes it, and otherwise he is employed at some wage distribution G. His wage is lower than the offered wage with probability $G(w)$, in which case he takes the job.

The outflow is explained similarly, bearing in mind that now what is relevant is not the actual wage distribution but the offered wage distribution $F(w)$.

To find the steady state, we need to set $\dot{N}(w) = 0$, which implies

$$N(w) = \frac{q(u + (1-u)G(w))}{(s + p(1 - F(w)))} \tag{61}$$

Moreover, we have a similar law of motion for the distribution function $G(w)$. In particular, the total fraction of workers employed and getting paid the wage of less than or equal to w is

$$(1-u)G(w)$$

The outflow of workers from this group is equal to

$$[s + p(1 - F(w))](1-u)G(w)$$

by the same reasoning as above.

The inflow of workers into the status of employed and being paid a wage less than w only come from unemployment (when a worker upgrades from the wage w' to $w'' \in (w', w]$, this does not change $G(w)$. Hence the inflow is

$$pF(w)u,$$

which is the measure of unemployed workers receiving an offer, pu, times the probability that this offer is less than w.

Equating the outflow and the inflow, we obtain the cumulative density function of actual wages as

$$G(w) = \frac{pF(w)\,u}{[s + p(1 - F(w))](1 - u)}$$

and using the steady-state unemployment rate:

$$G(w) = \frac{psF(w)}{p[s + p(1 - F(w))]} \tag{62}$$

The important thing to note is that

$$G(w) < F(w),$$

meaning that the fraction of jobs in the equilibrium wage distribution below wage w is always lower than the fraction of offers below w, so that F first-order stochastically dominates G.

Stated differently, this means that low wages have a lower probability of being accepted and, once accepted, a lower probability of surviving. Thus equilibrium wages are **positively selected** from offered wages.

Now combining (62) this with (61), we obtain

$$\begin{aligned} N(w) &= \frac{q\left(\frac{s}{s+p} + \frac{p}{s+p}\frac{psF(w)}{p[s+p(1-F(w))]}\right)}{(s + p(1 - F(w)))} \\ &= \frac{psq}{(s + p(1 - F(w)))^2} \end{aligned}$$

Thus we now have to solve for $F(w)$ only, or for $G(w)$ only.

In equilibrium, all firms have to make equal profits, which means equal discounted profits. This is a complicated problem, since a new firm accumulates workers slowly. Rather than solve this problem, let us look at the limit where $r \to 0$. This basically means that we can simply focus on state state, and equal discounted profits is equivalent to equal profits in the steady state.

The profits of a firm offering wage w (when the offer wage distribution is given by F), is

$$\pi(w) = (1-w) N(w)$$

In other words, an equilibrium satisfies

$$\pi(w) = \bar{\pi} \text{ for all } w \in \text{supp} F,$$

where $\bar{\pi}$ is also determined as part of the equilibrium.

Now solving these equations:

$$\pi(w) = (1-w) \frac{psq}{(s + p(1-F(w)))^2} = \bar{\pi}$$

Inverting this:

$$F(w) = 1 - \sqrt{\frac{(1-w)sq}{p\bar{\pi}}} + \frac{s}{p}$$

over the support of F.

Moreover, we know that $w = b$ is in the support of F, and $F(b) = 0$, and this implies

$$0 = 1 - \sqrt{\frac{(1-b)sq}{p\bar{\pi}}} + \frac{s}{p}$$

or

$$\bar{\pi} = \frac{(1-b)psq}{(p+s)^2}$$

Now substituting, we have

$$F(w) = 1 - \sqrt{\frac{(1-w)(p+s)^2}{(1-b)p^2}} + \frac{s}{p},$$

which is a well-behaved distribution function that is increasing everywhere.

Moreover, since $F(\bar{w}) = 1$, we also obtain that

$$\bar{w} < 1,$$

so even the highest wages less than the full marginal product of the worker.

From here, observed wage distribution is quite easy to calculate.

Why is there wage dispersion? The answer is similar to the models we have already seen.

The Burdett-Mortensen model is sometimes interpreted as a model of "monopsony". The reasoning is that firms do not face a flat labor supply curve, but can increase their "labor supply" by increasing their wages (they attract more workers and avoid losing workers). The reason why wages are lower than full marginal product is argued to be this monopsony power. This is in fact quite **misleading**. In the baseline search model, there is both monopoly power and monopsony power (there is bilateral monopoly, that's why there is bargaining!). In fact, we saw that in the Diamond's Paradox, even though in equilibrium firms do not face an upward-sloping labor supply, they have full "monopsony" power, and can hold workers down to their reservation utility (unemployment benefit or zero).

The nice thing about the Burdett-Mortensen model is that it generates a wage dispersion together with employer-size dispersion, and it matches a very well-known stylized fact that larger employers pay higher wages. The typical interpretation for this is that larger employers attract higher-quality workers or somehow workers have greater bargaining power against such employers. Burdett-Mortensen turn this on its head; they argued that it is not that larger employers pay higher wages, it is that employers that pay higher wages become larger in equilibrium! This is quite nice.

Nevertheless, there are theoretical objections to the Burdett-Mortensen model. The most important is that it is not optimal for firms to post wages. Instead they should post "contracts" that make workers pay upfront and receive their full marginal products. If this is not possible (because negative wages or bonding contracts are not allowed), they can make workers receive a low-wage early on, and increase their wages later. Why is this?

Finally, note that even though there is wage posting here, there isn't *directed search*. In the next lecture, we will see that directed search is the essence. The previous model of wage dispersion in fact may have mimicked directed search, for reasons we are going to see more clearly soon.

11 Wage Posting and Directed Search

We have so far seen a number of models which generate non-degenerate wage distributions. These models have the potential of leading to a coherence general equilibrium model with sequential search. The difficulty with these models, however, is that they are rather difficult to work with. What I will do next is to show that the equilibrium version of the Burdett-Judd model

(Acemoglu-Shimer, RES 2000) in the limit with small costs of search becomes very tractable, in fact leading to a new class of models which are very useful for the analysis of labor market equilibria. These models are sometimes referred to **competitive search** models, but I prefer to emphasize the two underlying assumptions: wage posting and directed search, so I will refer to them as **directed search** models. Hopefully, you'll see below why I emphasize the directed search element and why this is the right and this is.

The key observation is that in the Acemoglu-Shimer model, as costs of sampling wages go down, the equilibrium wage distribution tends to a model where workers observe sufficiently many wages, and decide which wage to apply to. This is essentially the type of model we will now analyze.

To bring out the most important points, I start from the economic environment of the search and investment model. Recall that in this model there are ex ante investments by firms, and bilateral search to form productive partnerships. In particular, recall that production requires 1 firm - 1 worker, with access to the production function $f(k)$, where k is capital for worker chosen before the matching stage by the firm. We assume

$$f' > 0, \quad f'' < 0$$

I continue to denote the rate of time preference by r, and the rate of separation due to the destruction of capital by s.

I am going to think of search frictions as equivalent to coordination frictions as in the more micro model of the last lecture. In particular, if there are an average of q workers per vacancy of a certain type then the flow rate of match for workers is $\mu(q)$, which is assumed to be continuously differentiable with $\mu' < 0$. Similarly, the flow rate of matching for a vacancy is $\eta(q) \equiv q\mu(q)$, where I am purposefully using the notation little q to distinguish this from the capital Q before which referred to the economy-wide queue length, whereas q it's specific to a type of job.

So this might seem somewhat strange; workers know what the various wages are, but conditional on applying to a job they may not get it; but this is sensible when there is no (centralized) coordination in the economy, because too many other people may be applying specifically to that job. The urn ball technology captured is in a very specific way, and in particular, we had

$$\eta(q) = 1 - \exp(-q) \text{ and } \mu(q) = \frac{1 - \exp(-q)}{q}$$

The technology here generalizes that.

As explained above, first all firms post wages w and also choose their capital k.

Workers observe **all** wages and then choose which job to seek. (they do not care about capital stocks).

Now more specifically let $q(w)$ be the ratio of workers seeking wage w to firms offering w. then $\mu(q(w))$ is flow rate of workers getting a job with wage w and $\eta(q(w))$ is flow rate of firms filling their jobs.

What equilibrium concept should we use here? Thinking about it intuitively, it's clear that we should ensure that workers apply to jobs that maximize utility and anticipate queue lengths at various wages rationally. This is straightforward.

The harder part is for firms. Firms should choose wages and investment to maximize profits, anticipating queue lengths at wages not offered in equilibrium. The last part is very important and corresponds to **Subgame perfection**. This is obviously important, since we have a dynamic economy, and you can see what will go wrong if we didn't impose subgame perfection.

Before we go further, let us first write the Bellman Equations, which are intuitive and standard for the firm (again imposing steady state throughout):

$$\begin{aligned} rJ^V(w,k) &= \eta(q(w))(J^F(w,k) - J^V(w,k)) - sJ^V(w,k) \\ rJ^F(w,k) &= f(k) - w - sJ^F(w,k) \end{aligned}$$

implying a simple equation for the value of firm

$$J^V(w,k) = \frac{\eta(f(k)-w)}{(r+s)(r+s+\eta)}$$

which we will use below.

The value of an employed worker is also simple:

$$rJ^E(w) = w + s(J^U - J^E(w))$$

What is slightly more involved is the value for unemployed worker.

Recall that unemployed workers take an important action: they decide which job to seek. Let $J^U(w)$ be the value of an unemployed worker when seeking wage w.

$$\underbrace{rJ^U(w)}_{\text{utility of applying to wage } w} = \mu(q(w))[J^E(w) - \underbrace{J^U}_{\substack{\text{maximal utility} \\ \text{of unemployment}}}]$$

where I have suppressed unemployment benefits without loss of any generality.

So what is J^U? Clearly:

$$J^U = \max_{w \in \mathcal{W}} J^U(w)$$

where $\mathcal{W}$ is the support of the equilibrium wage distribution.

Now this already builds in the requirement that w maximizes $J^U(w)$.

Also it is clear that w, k should maximize $J^V(w, k)$.

But what are the $q(w)$'s?

If we did not impose subgame perfection, then we could have crazy $q(w)$'s. Instead, firms would have to anticipate what workers would do if they deviate and create a new wage distribution.

So off-the-equilibrium path $q(w)$ should satisfy

$$\mu(q(w)) \left[J^E(w) - J^U\right] = rJ^U$$

or if $J^E(w) - J^U < rJ^U$, then $q(w) = 0$.

To define an equilibrium more formally, let an allocation be a tuple $\langle \mathcal{W}, Q, K, J^U \rangle$, where $\mathcal{W}$ is the support of the wage distribution, $Q : \mathcal{W} \to \mathbb{R}$ is a queue length function, $K : \mathcal{W} \rightrightarrows \mathbb{R}$ is a capital choice correspondence, and $J^U \in \mathbb{R}$ is the equilibrium utility of unemployed workers.

Definition 8 *A directed search equilibrium satisfies*

1. *For all* $w \in \mathcal{W}$ *and* $k \in K(w)$, $J^V(w,k) = 0$.

2. *For all* k *and for all* w, $J^V(w,k) \leq 0$.

3. $J^U = \sup_{w \in \mathcal{W}} J^U(w)$.

4. $Q(w)$ *s.t.* $\forall w$, $J^U \geq J^U(w)$, *and* $Q(w) \geq 0$, *with complementary slackness.*

In words, the first condition requires firms to makes zero profits when they choose equilibrium wages and corresponding capital stocks. The second requires that for all other capital stock and wage combinations, profits are nonpositive. The third condition defines J^U as the maximal utility that an unemployed worker can get. The fourth condition is the most important one. It defines queue lengths to be such that workers are indifferent between applying to available jobs, or if they cannot be made indifferent, nobody applies to a particular job (thus the **complementary slackness** part is very important). This builds in the notion of **subgame perfection.**

Now we have

Theorem 19 ***(Acemoglu and Shimer)*** *Equilibrium* k, w, q *maximize* $\frac{\mu(q)w}{r+s+\mu}$ $(= rJ^U)$ *subject to* $\eta(q)\frac{(f(k)-w)}{r+s+\eta(q)} = (r+s)k$. *And conversely, any solution to this maximization problem can be supported as an equilibrium.*

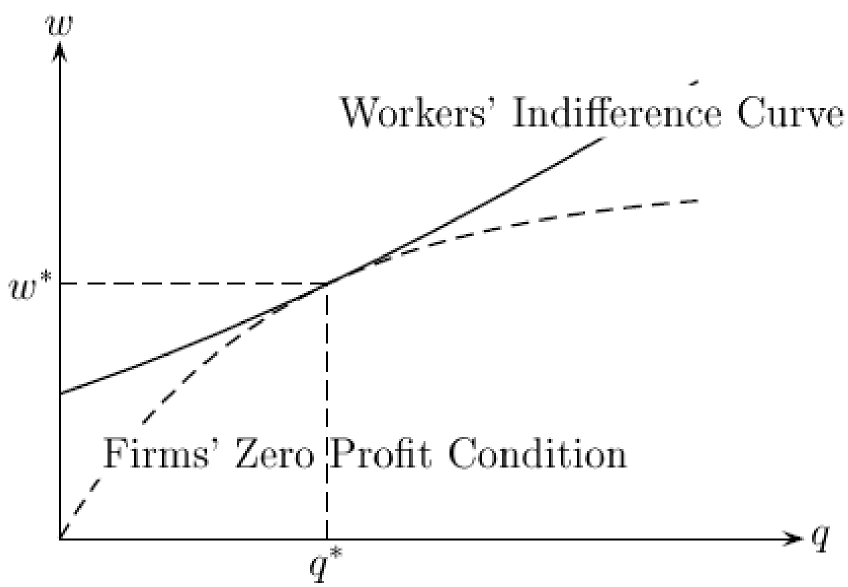
w
Workers' Indifference Curve
w*
Firms' Zero Profit Condition
q
q*

Figure 2:

Basically what this theorem says is that the equilibrium will be such that the utility of an unemployed worker is maximized subject to zero profit.

Proof. (sketch) Suppose not. Take k', w', q' which fails to maximize the above program. Then another firm can offer k'', w'' where (k^*, w^*, q^*) is the solution and $w'' = w^* - \varepsilon$. For ε small enough workers prefer k'', w'' to k', w', so $q'' > q^*$, which implies that k'', w'' makes positive profits, proving that (k', w', q) can't be an equilibrium. ■

This theorem is very useful because it tells us that all we have to do is to solve the program:

$$\max \frac{\mu(q)w}{r+s+\mu(q)}$$
$$\text{s.t.} \quad \frac{\eta(q)(f(k)-w)}{r+s+\eta(q)} = (r+s)k$$

Is this a convex problem?

No, but let's assume differentiablity (which we have so far), then first order conditions are necessary.

Forming the Lagrangian with multiplier λ

$$\frac{\eta(q)f'(k)}{r+s+\eta(q)} = r+s \tag{63}$$

$$\frac{\mu(q)}{r+s+\mu(q)} - \frac{\lambda\eta(q)}{r+s+\eta(q)} = 0 \tag{64}$$

and

$$\frac{(r+s)\mu'(q)}{(r+s+\mu(q))^2} + \lambda\left(\frac{(r+s)\eta'(q)(f(k)-w)}{(r+s+\eta(q))^2}\right) = 0 \tag{65}$$

Now (63) is identical to (46) above, which was

$$\frac{Q^S\mu(Q^S)f'(k^S)}{r+s+Q^S\mu(Q^S)} = r+s$$

implies that, denoting the capital labor ratio in the wage posting equilibrium by k^{wp},

$$k^{wp} = k^S$$

Therefore, with wage posting, capital investments are always efficient.

Why is this? You might think this is because there is no more holdup problem, and this is essentially true, but the intuition is a bit more subtle. In fact, there is something like hold-up because firms that invest more in equilibrium prefer to pay higher wages, but despite this the efficient level of investment results. The reason is that the higher wages that they pay is exactly offset with the higher probability that they will attract workers, so net returns are not subject to hold-up.

Next we have

$$\lambda = \frac{r + s + \eta(q)}{(r + s + \mu(q))q}$$

and substitute this into (iii), and used at zero profit constraints to solve for

$$w = f(k) - \frac{(r + s)(r + s + \eta(q))}{\eta(q)}k$$

Then we have:

$$\eta' \frac{q^2 f(k)}{r + s} + \left[r + s + \mu + \mu' q - q^2 \mu'\right] k = 0$$

which is identical to (47). We have therefore established:

Theorem 20 *The directed search equilibrium of the search and investment model is constrained efficient.*

Therefore, the equilibrium is constrained efficient! (note uniqueness is not guaranteed, but neither was it in the social optimum)

Thus, wage posting decentralizes the efficient allocation as the unique equilibrium.

How can we understand this efficiency better?

Acemoglu-Shimer consider a number of different economies

1. Wage posting but no directed search. Clearly, in this case things are very bad, and we get the Diamond paradox.

2. An economy where firms choose their own capital level, and then "post a bargaining parameter β" and upon matching, the firm and the worker Nash bargain with this parameter. It can be shown that if there is no capital choice, this economy will lead to an equilibrium in which all firms post the Hosios β, and constrained efficiency is achieved. But if there is a capital choice, and the only thing workers observe are the posted β's, then in equilibrium all firms offer the Hosios β, but there is under investment because of the hold-up problem.

3. An economy where firms choose their own capital level and workers apply to firms observing these capital levels, and then they bargain according to some exogenously given parameter β. In this case, the equilibrium is inefficient and may have under or overinvestment. If the value of β is at the Hosios value, then the equilibrium will be constrained efficienct.

4. An economy where firms choose their own capital level and post β, and workers observe both k and β, then always constrained efficiency.

So what do we learn? What is important is directed search, and especially the ability to direct search towards higher capital intensity firms. With wage posting, those are the high-wage firms, hence the objective is achieved. But the same outcome is also obtained if β is at the Hosios level, and workers observe capital levels.

Next, one might wonder whether an economy in which workers know/observe all of the wages offered in equilibrium is too extreme (especially given our motivation of doing away with a Walrasian auctioneer). A more plausible economy may be one where workers observe a finite number of wages.

Interestingly, we do not need all workers to observe all the wages as the model with a non-degenerate wage distribution in the last lecture illustrated.

Theorem 21 *Suppose each worker observes (can apply to) at least two of the firms among the continuum of active firms, then the efficient allocation is an equilibrium of the search and investment model with directed search and wage posting.*

Proof. (sketch) Suppose all firms are offering (q^{wp}, w^{wp}, k^{wp}). Now consider a deviation to some other (w', k'). Any worker who observes (w', k') has also observed another firm offering (w^{wp}, k^{wp}). Since (w^{wp}, k^{wp}) maximizes worker utility, he will apply to this in preference of

$$(w', k') \Longrightarrow q(w') = 0.$$

Consequently, all firms will be happy to offer (w^{wp}, k^{wp}) and they will each be tracked the queue length of q^{wp}. ■

What is the intuition? **Effectively Bertrand Competition.** Each firm knows that it will effectively be competing with another firm offering the best possible deal to the worker, even though differently from the standard Bertrand model, it does not know which particular firm this will be. Nevertheless, the Bertrand reasoning forces each firm to go to the allocation that is best for the workers.

Note that this theorem is not stated as an "if and only if" theorem. In particular, when each worker only observes two wages, there can be other “non-efficient" equilibria. This is again left for homework:

Theorem for Homework *When each worker observes two wages, there can exist non-efficient equilibria.*

This last theorem not withstanding, the conclusion of this analysis is that relatively little information is required for wage posting to decentralize the efficient allocation.

12 Risk Aversion in Search Equilibrium

The tools we developed so far can also be used to analyze general equilibrium search with risk aversion. Let us focus on the one-period model with wage posting. This can again be extended to the dynamic version, but explicit form solutions are possible only under constant absolute risk aversion (see Acemoglu-Shimer, JPE 1999)

Measure 1 workers; and they all have utility $u(c)$ where the consumption of individual i is

$$C_i = A_i + y_i - \tau_i$$

where A_i is the non-labor income of individual, y_i is his labor income, equal to the wage w that he applies it obtains if he's employed, and equal to the unemployment benefit z when unemployed. Finally, τ_i is equal to the taxes paid by this individual. u is increasing, concave and differentiable.

Let us start with a homogeneous economy where $A_i = A_0$ and $\tau_i = \tau$ for all i.

We also assume that firms are risk-neutral, which is not chill for example because workers may hold a balanced mutual fund. I will onlypresent the analysis for the static economy here.

Timing of events:

- Firms decide to enter, buy capital $k > 0$ (as before irreversible,) and post a wage w

- Workers observe all wage offers and decide which wage to seek (apply to).

As before, if on average there are q times as many workers seeking wage w as firms offering w, then workers get a job with prob. $\mu(q)$.

Firms fill their vacancies with prob. $\eta(q) \equiv q\mu(q)$, with our standard assumptions, $\mu'(q) < 0$ and $\eta'(q) > 0$

As before, let an allocation be $\langle \mathcal{W}, Q, K, U \rangle$, where $\mathcal{W}$ is the support of the wage distribution, $Q : \mathcal{W} \to \mathbb{R}$ is a queue length function, $K : \mathcal{W} \rightrightarrows \mathbb{R}$ is a capital choice correspondence, and $U \in \mathbb{R}$ is the equilibrium utility of unemployed workers.

Definition 9 *An allocation is an equilibrium iff*

1. $\forall w \in \mathcal{W}$ *and* $k \in K(w)$, $\eta(Q(w))(f(k) - w) - k = 0.$
2. $\forall w, k,\ \eta(Q(w))(f(k) - w) - k \leq 0.$
3. $U = \sup_{w \in \mathcal{W}} \mu(Q(w))u(A + w) + (1 - \mu(Q(w))u(A + z)$
4. $Q(w)$ *s.t.* $\forall w,\ U \geq \mu(Q(w))u(A + w) + (1 - \mu(Q(w)))u(A + z)$ *and* $Q(w) \geq 0$, *with complementary slackness.*

- $\Longrightarrow$ As before type of subgame perfection on beliefs about queue lengths after a deviation.

Characterization of equilibrium is similar to before

Theorem 22 $(\mathcal{W}, Q, K, U)$ *an equilibrium if and only if* $\forall\ w^* \in \mathcal{W},\ q^* \in Q(w^*), k^* \in K(w^*)$

$$(w^*,\ q^*, k^*) \in \arg\max \mu(q)u(A + w) + (1 - \mu(q))u(A + z)$$

s.t.

$$\eta(q)(f(k) - w) \geq 0.$$

In words, every equilibrium maximizes worker utility subject to zero profits, as proved before in the context of the risk-neutral model.

The analysis is similar to before. Profit maximization implies an even simpler condition (because the environment is static)

$$\eta(q^*) f'(k^*) = 1$$

Zero profits gives

$$\eta(q^*)(f(k^*) - w^*) = k^*$$

Now combining these two:

$$w^* = f(k^*) - k^* f'(k^*),$$

which you will notice is exactly the neoclassical wages equal to marginal product condition. Why is that?

Finally, combining this with, $\eta(q^*) f'(k^*) = 1$, we can derive a relation in the (q, w) space which corresponds to the zero-profits and profit maximization constraints that an equilibrium has to satisfy.

An equilibrium is then a tangency point between the indifference curves of homogeneous workers and this profit-maximization constraint, as we had in the risk-neutral model of Acemoglu-Shimer (IER, 1999):

The equilibrium can be depicted and analyzed diagrammatically.

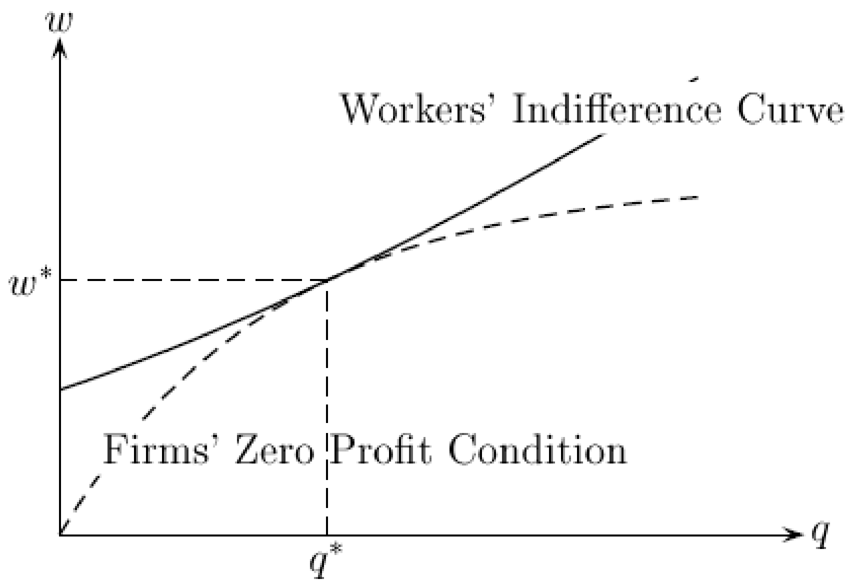

Figure 3:

Notice again that uniqueness **not** guaranteed.

What makes this attractive is that comparative statics can also be done in a simple way, exploiting "revealed preference" or single crossing.

For example, we have a change such that all workers become more risk-averse, i.e., and the utility function becomes more concave, what happens to equilibrium?

We can show that as risk-aversion increases, then we have $w \downarrow, q \downarrow, k \downarrow$.

Why? Indifference curves become everywhere steeper, the causing the tangency point to shift to the left. Unambiguous despite the fact that equilibrium may not be unique.

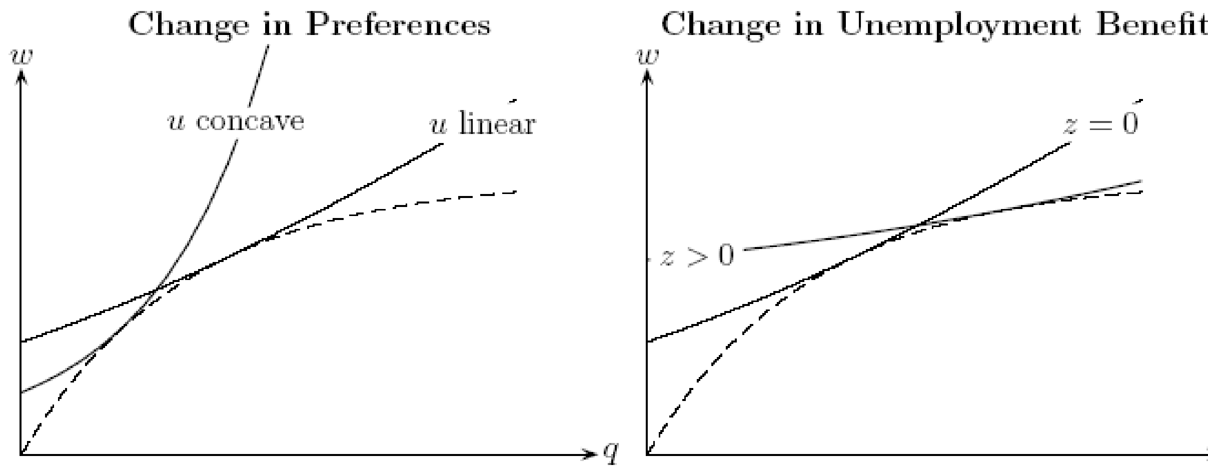

Figure 4:

Essentially, comparative static result unambiguous because u_1-curve single-crosses u_2-curve.

Intuition: "Market Insurance." Workers are more risk-averse, so firms offer insurance by creating low-wage but easier to get jobs. Capital falls because once jobs are easier to get for workers, vacancies remain open for longer (with higher probability), so capital is unused for longer, reducing investment. Summarizing this:

Theorem 23 *Consider a change from utility function u_1 to u_2 where u_2 is a strictly concave transformation of u_1. Then if (k_1, w_1, q_1) is any equilibrium* *with preferences* u_1 *and* (k_2, w_2, q_2) *is any equilibrium with preferences* u_2*, then* $k_2 < k_1$*,* $w_2 < w_1$ *and* $q_2 < q_1$*.*

Similarly, what happens when the unemployment benefits z increases from z_1 to z_2?

Theorem 24 *Consider a change from unemployment benefits* z_1 *to* $z_2 > z_1$*. Then if* (k_1, w_1, q_1) *is any equilibrium with benefits* z_1 *and* (k_2, w_2, q_2) *is any equilibrium with benefits* z_2*, then* $k_2 > k_1$*,* $w_2 > w_1$ *and* $q_2 > q_1$*.*

Proof. (sketch) By revealed preference

$$\begin{aligned} \mu(q_1)(u(A + w_1) - u(A + z_1)) &\geq \mu(q_2)(u(A + w_2) - u(A + z_1)) \\ \mu(q_2)(u(A + w_2) - u(A + z_2)) &\geq \mu(q_1)(u(A + w_1) - u(A + z_2) \end{aligned}$$

Multiply through and simplify

$$\begin{aligned} & (u(A + z_1) - u(A + z_2))(u(A + w_1) - u(A + w_2)) \geq 0 \\ \Longrightarrow \quad & z_1 \leq z_2 \Longleftrightarrow w_1 \leq w_2. \end{aligned}$$

All inequalities strict since all curves smooth. ■

What happens when there is heterogeneity?

Suppose that there are $s = 1, 2, \ldots, S$ types of workers, where type s has utility function u_s, after-tax asset level A_s, and unemployment benefit z_s. Let U now be a vector in $\mathbb{R}^S$, and assume, for simplicity. Then:

Theorem 25 *There always exists an equilibrium. If* $\{\mathcal{K}, \mathcal{W}, Q, U\}$ *is an equilibrium, then any* $k_s^* \in \mathcal{K}$*,*$w_s^* \in \mathcal{W}$*, and* $q_s^* = Q(w_s^*)$*, solves*

$$U_s = \max_{k,w,q} \mu(q) u_s(A_s + w) + (1 - \mu(q))\, u(A_s + z_s)$$

subject to $\eta(q)(f(k) - w) - k = 0$ *for some* $s = 1, 2, .., S$*. If* $\{k_s^*, w_s^*, q_s^*\}$ *solves the above program for some* s*, then there exists an equilibrium* $\{\mathcal{K}, \mathcal{W}, Q, U\}$ *such that* $k_s^* \in \mathcal{K}$*,* $w_s^* \in \mathcal{W}$*, and* $q_s^* = Q(w_s^*)$*.*

The important result here is that any triple $\{k_s^*, w_s^*, q_s^*\}$ that is part of an equilibrium maximizes the utility of one group of workers, subject to firms making zero profits. The market *endogenously* segments into S different sub-markets, each catering to the preferences of one type of worker, and receiving applications only from that type.

The efficiency and output-maximization implications of this model are also interesting. First, supposed that $u(\cdot)$ is linear. Then $z = \tau = 0$ maximizes output. In particular, we have

Theorem 26 *Suppose that* u *is linear, then* $z = \tau = 0$ *maximizes output.*

Proof. (sketch) The equilibrium solves $\max \mu(q)w$ subject to $q\mu(q)(f(k) - w) = k$. Substituting for w we obtain:

$$\mu(q)f(k) - k/q \equiv y(k, q),$$

which is net output, thus is maximized by equilibrium choices. ■

But an immediate corollary is that if $u(\cdot)$ is strictly concave, than the equilbrium with $z = \tau = 0$ does *not* maximize output.

Theorem 27 *Suppose that u is strictly concave, then $z = \tau = 0$ does not maximize output.*

This is an immediate corollary of the previous theorems.

Theorem 28 *Let u be an arbitrary concave utility function, q^e be the output-maximizing level of queue length and let*

$$z^e \equiv \frac{u(A_0 - \tau^e + w^e) - u(A_0 - \tau^e + z^e)}{u'(A_0 - \tau^e + w^e)}$$

and the balanced-budget condition

$$\tau^e = (1 - \mu(q^e))z^e$$

then the economy with unemployment benefit z^e achieves an equilibrium with q^e and the maximum output.

The following figure gives the intuition:

But this is not "optimal", since when workers are risk averse, maximizing output is not necessarily the right objective. Optimal unemployment benefits, z^o, should maximize ex ante utility. Interestingly, this could be greater or less than the efficient level of unemployment benefits, z^e, which maximizes output. What is the intuition for this?

13 Moral Hazard and Optimal Unemployment Insurance

The previous model incorporated risk aversion into a model of search (and investment). An implicit form of moral hazard was present, that when work-

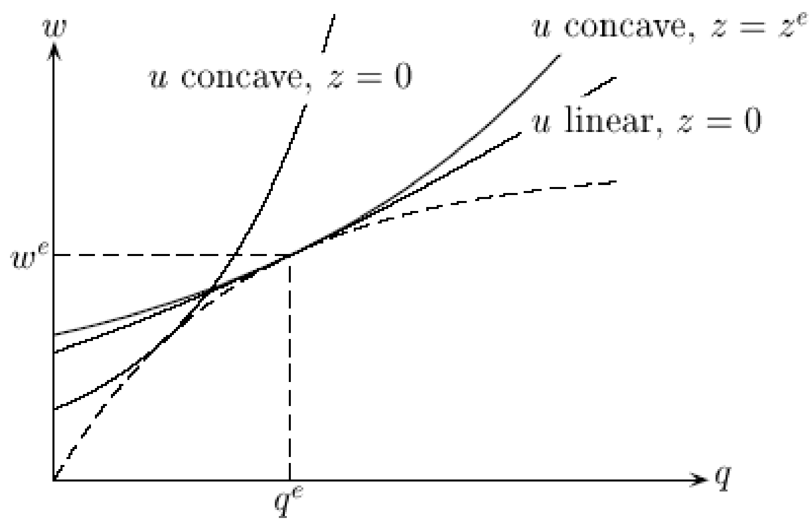

Figure 5:

ers had access to better insurance, they became more choosy in their job application decisions, which in equilibrium changed the wage distribution.

A more standard approach in the literature directly focuses on moral hazard decisions. This approach was first developed by Shavell and Weiss' classic paper in 1979. Here I will present a slight generalization of their approach based on a more recent paper by Hopenhayn and Nicolini (JPE, 1997). The important focus here will be on thinking of the dynamics of optimal unemployment insurance. For example, as opposed to the previous model where unemployment insurance levels were constant, in practice, unemployment benefits run out after a certain duration of unemployment. Is this optimal? Could the unemployment insurance systems used in practice be improved?

The model incorporates moral hazard regarding search effort (but there are no application decisions). Since the firm side is left implicit, it is essentially a partial equilibrium model. The preferences of the agent are

$$E\sum_{t=0}^{\infty}\beta^t\left[u(c_t)-a_t\right]$$

where $c_t \in \mathbb{R}$ is consumption and $a_t \in A$ is search effort, which lead to a probability of finding a job $p_t = p(a_t)$. All jobs are homogeneous and pay w (the feature that rules out the application margin). We naturally assume that

$$p'' < 0,\ p' > 0.$$

We also assume that the individual has zero income when unemployed and does not have access to any savings or borrowing opportunities. This last assumption is crucial and simplifies the analysis by allowing the unemployment insurance authority to directly control the consumption level of the individual. Otherwise, there will be an additional constraint which determines the optimal consumption path of the individual.

Let s_j be state at time j

$$\begin{array}{lcl} s_j = 0 & \rightarrow & \text{unemployed} \\ s_j = 1 & \rightarrow & \text{employed} \end{array}$$

The important object will be the history of the agent up to time t, which is denoted by $h_t = \{s_j\}_{j<t}$. Let $\mathcal{H}_t$ be this set of all such histories.

A general insurance contract can be represented as a mapping

$$\tau : \mathcal{H}_t \longrightarrow A \times \mathbb{R}$$

where the first element of the mapping is a_t, the "recommended search effort" and the second element z_t is the transfer to the worker, which will directly determine his consumption, since he has no access to an outside source of consumption and no savings opportunities.

Let $V_0(\tau)$ be the expected discounted utility at $t = 0$ associated with contract τ, and to prepare for setting up the dual of this problem, let $C_0(\tau)$ be the expected cost (net transfers) to the agent.

Now the optimal contract choice can be set up as

$$\max\ V_0(\tau)$$

s.t.

IC (incentive compatibility constraints) – if any

$$C_0(\tau) \leq \underline{C}$$

The last constraint for example may require the total cost to be equal to zero, i.e., all benefits to be financed by some type of payroll taxes or other taxation. E.g., budget balance as in the previous model.

Instead of this problem, we can look at the dual problem

$$\min C(V) = C_0(\tau)$$

s.t. IC

$$V_0(\tau) \geq V$$

Let us start with the full information case where the social planner (the unemployment insurance authority) can directly monitor the search effort of the unemployed individual, so the individual has no choice but to choose the recommended search effort. This implies that there are no IC constraints.

Then, it is straightforward that full insurance is optimal, i.e., $c_t = c \; \forall \; t$, and the level of search effort will solve:

$$a^* = \arg\max_a p(a) \sum_{t=0}^{\infty} \beta^t [1 - p(a)]^t \left[\frac{u(c)}{1-\beta} - a \right]$$

The more interesting case is the one with imperfect information, where a is the private information of the individual, so he will only follow the recommended search effort if this is incentive compatible for him. In other words, as in all types of implementation or optimal policy problems, there is an "argmax" constraint on the maximization problem.

Suppose $V_0(\tau) = V$. Let us introduce some useful notation

$$\begin{array}{ll} V^e = V_1(\tau) & \text{if } s_1 = 1 \\ V^u = V_1(\tau) & \text{if } s_1 = 0. \end{array}$$

This implies that we can write the value of the individual as

$$V = u(c) - a + \beta \left\{ p(a) V^e + (1 - p(a)) V^u \right\}$$

Now the incentive compatibility constraints boil down to

$$\text{(IC)} \qquad a \in \arg\max_{a'} u(c) - a' + \beta \left\{ p(a') V^e + (1 - p(a')) V^u \right\}. \qquad (66)$$

Naturally, (66) defines a very high dimensional object. It basically requires a to be better than or as good as any other feasible choice in A. These kinds of constraints are very difficult to work with, so the literature usually takes **the first-order approach**, which is to represent (66) with the corresponding first-order condition of the agent, i.e.,

$$\beta p'(a)(V^e - V^u) = 1 \tag{67}$$

This may seem innocuous, but in many situations it leads to the **wrong** solution. One has to be very careful in using the first-order approach. In this case, the situation is not so bad, because the individual only has a single choice, and given V^e and V^u, his maximization problem is strictly concave, so the first-order condition (67) is necessary and sufficient for the individual's maximization problem. Nevertheless, this constraint itself, i.e., (67), is non-linear and non-convex, so some of the difficulties of designing optimal contracts carry over to this case.

The problem is further simplified by noting that after the individual finds a job, there is no further incentive problem, so after that point there will be full consumption smoothing, i.e.,

$$V^e = \frac{u(c^e)}{1-\beta} \qquad \text{for some } c^e. \tag{68}$$

This is equivalent to a per-period transfer $c^e - w$ to the agent. In other words, there may be negative or positive transfers to the agent after he finds a job. The level of these transfers will be a function of its history, i.e., when (after how many periods of unemployment) he has found a job.

Now let

$$W(V^e) = \frac{c^e - w}{1 - \beta}$$

be the discounted present value of the transfer from the principal to the agent. Inverting (68), we have

$$W(V^e) = \frac{-w + u^{-1}[(1-\beta)V^e]}{1 - \beta}$$

Differentiating this equation, we obtain an intuitive formula

$$W'(V^e) = \frac{1}{u'(c^e)},$$

which states that the cost of providing greater utility is the reciprocal of the marginal utility of consumption for the individual. When $u'(c^e)$ is high, providing more utility to the individual is relatively cheap. From the concavity of the individual's utility function, u, W is also seen to be a convex function (it is clearly increasing).

Now let $C(V)$ be the cost of providing utility V to an unemployed individual. It can be written in a recursive form as

$$C(V) = \min_{a, c^u, V^e, V^u} c^u + \beta \left\{ p(a) W(V^e) + [1 - p(a)] C(V^u) \right\}$$

subject to

$$u(c^u) - a + \beta \left\{ p(a) V^e + [1 - p(a)] V^u \right\} = V \tag{69}$$

$$\beta p'(a)(V^e - V^u) = 1 \tag{70}$$

where c^u is utility given to unemployed individual, (69) is the promise keeping constraint, which makes sure that the agent indeed receives utility V. (70) is the IC constraint using the first-order approach. Note that this formulation makes it clear that the social planner or the unemployment insurance authority is directly controlling consumption. Otherwise, there would be another constraint corresponding to the Euler equation of the individual for example.

Also, notice that this is a standard recursive equation, so time has been dropped and everything has been written recursively. This creates quite a bit of economy in terms of notation. Moreover, the existence of a function $C(V)$ can be again guaranteed using the contraction mapping theorem (Theorem 1).

An interesting question is whether $C(V)$ is convex. Recall that in the standard dynamic programming problems, concavity of the payoff function and the convexity of the constraints set were sufficient to establish concavity of the value function. Here we are dealing with a minimization problem, so the equivalent result would be convexity of the cost function. However, the constraint set is no longer convex, so the convexity of $C(V)$ is not guaranteed. This does not create a problem for the solution, but it implies that there may be a better policy than the one outlined above which would involve using lotteries.

Can you see why lotteries would improve the allocation in this case? Can you see how the problem should be formulated with lotteries?

Here, to simplify the analysis, let us ignore lotteries.

To make more progress, let us assign multiplier λ to (69) and η to (70). Then the first-order conditions (with respect to a, c^u, V^e and V^u) are

$$\beta p'(a)[W(V^e) - C(V^u)] - \lambda\left[\beta p'(a)(V^e - V^u) - 1\right] - \eta\beta p''(a)(V^e - V^u) = 0$$

$$1 - \lambda u'(c^u) = 0$$

$$\beta p(a)W'(V^e) - \lambda\beta p(a) - \eta\beta p'(a) = 0$$

$$\beta\left[1 - p(a)\right]C'(V^u) - \lambda\beta[1 - p(a)] + \eta\beta p'(a) = 0$$

The second first-order condition immediately implies

$$\lambda = 1/u'(c^u)$$

Now substituting this into the other conditions (and using constraint (70)), we have

$$p'(a)[W(V^e) - C(V^u)] = \eta p''(a)(V^e - V^u) \tag{71}$$

$$C'(V^u) = \frac{1}{u'(c^u)} - \eta\frac{p'(a)}{1 - p(a)} \tag{72}$$

$$W'(V^e) = \frac{1}{u'(c^e)} = \frac{1}{u'(c^u)} + \eta\frac{p'(a)}{p(a)} \tag{73}$$

In addition, we have the following envelope condition by differentiating the cost function with respect to V:

$$C'(V) = \frac{1}{u'(c^u)} = [1 - p(a)]C'(V^u) + p(a)W'(V^e) \tag{74}$$

We now have a key result of optimal unemployment insurance:

Theorem 29 *The unemployment benefit and thus unemployed consumption, c^u, is decreasing over time. In addition, if $C(V)$ is convex, then $V^u < V$.*

Proof. (sketch) From (72) and (73), we have that

$$W'(V^e) - C'(V^u) = \eta p'(a)\left[\frac{1}{1-p(a)} + \frac{1}{p(a)}\right].$$

Since $\eta > 0$ (see the paper, or think intuitively), this immediately implies

$$W'(V^e) > C'(V^u)$$

Now use the Envelope condition (74), which immediately implies

$$W'(V^e) > C'(V) > C'(V^u) \tag{75}$$

Let $\hat{c}^u$ be next period's consumption. Then we have

$$C'(V^u) = \frac{1}{u'(\hat{c}^u)},$$

which combined with (75) and (74) and the concavity of the utility function u immediately implies

$$\hat{c}^u < c^u$$

as claimed. Moreover, (75) also implies that $V^u < V$ as long as C is convex, completing the proof of the theorem. ■

What is the intuition? **Dynamic incentives**: the planner can give more efficient incentives by reducing consumption in the future.

A related question is what happens to the transfer/tax to employed workers. Is this a function of history?

Theorem 30 *The wage tax/subsidy is a function of history, h_t, i.e., it is not constant.*

Proof. (sketch) Let us revisit the envelope condition and rewrite it as

$$
\begin{aligned}
C'(V_t) &= [1-p(a_t)]C'(V_{t+1}) + p(a_t)W'(V_t^e). \\
C'(V_t) &= \sum_{i=0}^{T-1} \left\{ \prod_{j=0}^{i-1} (1-p(a_{t+j})) \right\} p(a_{t+i}) W'(V_{t+i}^e) \\
&\quad + \left\{ \prod_{j=0}^{T-1} (1-p(a_{t+j})) \right\} C'(V_{t+T}^u)
\end{aligned}
$$

Now to obtain a contradiction, suppose that $V_t^e = V^e$ for all t. From Theorem 29, V_t^u must eventually be decreasing (since consumption benefits are). Let the second term with $\prod_{j=0}^{T-1}$ be denoted by b_2. Since $C'(V_{t+T}^u)$ is bounded, so as $T \to \infty$, we have $b_2 \to 0$. Therefore,

$$
C'(V_t) = \sum_{i=0}^{\infty} \left\{ \prod_{j=0}^{i-1} (1-p(a_{t-1+j})) \right\} p(a_{t-1+i}) W'(V_{t+i}^e)
$$

Since, by hypothesis, $W'(V_{t+i}^e)$ is constant, we have

$$
\begin{aligned}
C'(V_t) &= W'(V^e) \sum_{i=0}^{\infty} \left\{ \prod_{j=0}^{i-1} (1-p(a_{t+j})) \right\} p(a_{t+i}) \\
&= W'(V^e),
\end{aligned}
$$

which contradicts (75), so V^e cannot be constant and therefore c^e cannot be constant. ■

Under further assumptions, it can be established that generally c^e is a decreasing sequence, which implies that Optimal unemployment insurance schemes should make use of employment taxes conditional on history as well as allow for decreasing benefits.

Can you see the intuition for why wage taxes/subsidies are non-constant? Can you relate this result to decreasing benefits?