

# Capítulo Ocho

---

---

## EMPLEO, DESEMPLEO Y SALARIOS

---

---

Al asumir el equilibrio, podemos incluso ir tan lejos como para abstraernos de los empresarios y simplemente considerar que, en un cierto modo, los servicios productivos se están intercambiando directamente los unos por los otros.

- Leon Walras, *Elementos de Economía Pura* (1874)

Está claro que los productos no pueden salir al mercado y hacer intercambios por su propia cuenta. Por lo tanto, tenemos que recurrir a sus ... dueños ... quienes no deben... apropiarse del producto del otro, y entregar el suyo propio, excepto por las vías del mutuo acuerdo.

- Karl Marx, *Capital, I* (1867)

EN LA MAÑANA DEL 5 de enero, 1914, un mecánico virtualmente desconocido convertido en productor de automóviles, llamado Henry Ford, conmocionó a sus colegas y competidores al anunciar que él pagaría a sus trabajadores un mínimo de cinco dólares por una jornada de ocho horas, acortando de esta manera la jornada laboral y aumentando en más del doble la tarifa de pago por hora para la vasta mayoría de sus empleados.<sup>1</sup> Ford no estaba respondiendo a una provisión insuficiente de trabajo: un reportero que llegaba esa mañana para la conferencia de prensa en la cual se haría el anuncio observó una fila de varios cientos de trabajadores que buscaban empleo. En las semanas que siguieron al anuncio, la fila fuera de los portones aumentó a más de doce mil, casi tantos como los que trabajaban adentro. Asombrosamente, las ganancias aumentaron, respaldadas por un aumento de más del doble en el rendimiento por hora de trabajo de producción. Ford iba a convertirse en una palabra familiar en todo el mundo, y *Fordismo* en un enfoque peculiarmente americano a las relaciones de trabajo.

Para aquellos empleados con suerte, quienes habían estado en el lugar correcto en el momento correcto, los hechos fundamentales de la vida laboral dentro de la planta cambiaron hasta ser irreconocibles. La fuerza laboral de Ford había promediado los 13.623 empleados el año anterior. Durante el curso de ese año, 50.448 empleados habían salido de la empresa, la mayoría había renunciado; 8.490 habían sido despedidos. Al año siguiente del anuncio, el empleo había crecido en un tercio, pero el número de quienes renunciaban había disminuido a un décimo de su nivel anterior, y solamente veintisiete empleados habían sido despedidos. Está claro que cambios de esta magnitud no se pueden explicar por variaciones cíclicas de la oferta y la demanda en el mercado de trabajo local. Parece poco probable que Ford haya doblado el salario para atraer a mejores trabajadores o para retener a aquéllos en quienes la compañía había invertido en un costoso entrenamiento; un superintendente Ford se jactó de que “dos días son . . . suficiente tiempo para convertir en un moldeador de panales de primera clase a un hombre que nunca en su vida había visto una banca de moldurado.” Saber exactamente por qué Ford había hecho esto sigue siendo un misterio. Más importante aún, el éxito de su juego es un rompecabezas, pues no es consistente con la visión Walrasiana de que la maximización de los beneficios implica que uno les pague a sus empleados un salario igual a su precio de oferta (su siguiente mejor alternativa).

---

<sup>1</sup> Este recuento se basa en Raff (1988).

En el marco neoclásico, los servicios productivos literalmente no se intercambian directamente los unos por los otros como lo sugiriera Walras en los *Elementos*. Pero su extravagancia no está tan lejos de la realidad: una empresa es simplemente un conjunto de producción factible dado por las tecnologías disponibles, presidida por un gerente. El gerente selecciona la mezcla de insumos y producción que maximicen la riqueza del dueño, comprando insumos y vendiendo la producción en los mercados a precios determinados exógenamente. Es fácil ver por qué la jornada a cinco dólares de Ford no iba a tener sentido en este modelo.

Existen tres ingredientes básicos de un modelo más adecuado. El primero es el discernimiento de Ronald Coase (1992:717), que se menciona al cierre del capítulo 7, el cual dice que “lo que negociamos en el mercado no son, como los economistas a menudo suponen, entidades físicas sino los derechos para realizar ciertas acciones.” El segundo ingrediente es la trivialidad de Marx, de que el intercambio exige que los propietarios de los servicios productivos interactúen cara a cara. El tercer ingrediente es el descubrimiento que hiciera Henry Ford, de que los empleados pueden reciprocarse una buena paga con un gran esfuerzo de trabajo.

Marx fue el primero en recalcar el hecho de que el contrato de empleo no se relacionaba con cosas tales como la cantidad o la calidad del trabajo realizado; más bien, el contrato especificaba las horas durante las cuales el empleado acordaba someterse a la autoridad del empleador. De acuerdo con Marx (1973:275), el suministro actual de esfuerzos del empleado al proceso de producción no se aseguraba mediante un contrato sino que más bien era una “apropiación de trabajo por capital” que “solamente por un uso erróneo podía . . . haber sido denominado como algún tipo de intercambio.” Anticipándose a Ford (por no mencionar los desarrollos en la teoría económica a finales del siglo veinte), Marx (1967:544) señaló que un aumento en el salario pudiese reducir el costo del trabajo. Al igual que Marx, Coase puso de relieve el rol central de la autoridad en las relaciones contractuales de la empresa: “[o]bsérvese el carácter del contrato que con el cual se compromete un factor al que se le emplea en una empresa. . . . [E]l factor . . . acuerda obedecer las instrucciones del empresario a cambio de una cierta

remuneración.” De hecho, Coase (1937:387, 389) *definió* a la empresa por su estructura política:

Si un trabajador cambia del departamento Y al departamento X, él no va debido a un cambio en el precio sino porque se le ha ordenado hacerlo... la marca distintiva de la empresa es la supresión del mecanismo de precios.

Coase buscó entender por qué existen en sí las empresas y qué es lo que determina el alcance de lo que él llamaba (prestando la frase de Dennis Robertson) estas “islas de poder consciente en este océano de cooperación inconsciente.”

Herbert Simon (1951) propuso el primer modelo de empresa en conformidad con estas líneas. Él representó el contrato de empleo como un intercambio en el cual los empleados transfieren la autoridad sobre sus tareas y trabajos al empleador a cambio de un salario. Simon subrayó la ventaja de este arreglo para el empleador en vista de la inevitable incertidumbre sobre las tareas que se requerirían a lo largo del curso del contrato y, por lo tanto, el alto costo de llegar a un acuerdo sobre una especificación contractual completa de las actividades a ser llevadas a cabo. Yo me referiré al enfoque que estos disímiles autores iniciaron como el modelo Marx-Coase-Simon de relaciones laborales. Una característica de la interacción empleador-empleado en este enfoque es que las preferencias sociales – especialmente los motivos de reciprocidad y justicia – juegan una parte importante en la determinación de los resultados.

## LA RELACIÓN DE EMPLEO

El modelo de mercado laboral y la relación de trabajo a seguir es una variante de lo que se pudiera denominar la *regulación de esfuerzos* o el *modelo de disciplina de trabajo* que se basa en la *renovación contingente*. (Explicaré más adelante por qué pienso que la denominación común de “modelo de salario de eficiencia” es engañosa).

*El Problema.* No se puede contratar el esfuerzo de trabajo ya que el empleador conoce la información que se relaciona con el esfuerzo de un empleado en el mejor de los casos de manera muy imperfecta y ésta no es verificable (no es admisible en un tribunal). Incluso si la información fuese verificable, un contrato para pagarle a un

empleado de acuerdo a una señal de esfuerzo muy ruidosa expondría al trabajador a un nivel de riesgo subjetivamente muy costoso. Y aún así, el esfuerzo de trabajo es un argumento en la función de producción del empleador. Se pudiera evitar el problema si la persona quien hace el trabajo, como Robinson Crusoe, fuese también el demandante residual del resultado emergente, como sería el caso si fuese factible implementar contratos óptimos para la producción en equipo del tipo como el que se modeló en el capítulo 4. Pero, por los motivos allí explicados, este tipo de contrato también expondría al empleado a un nivel de riesgo inaceptable. Niveles de producción individuales también harían del trabajador un demandante residual sobre sus esfuerzos, pero las economías de escala generalmente hacen que la producción en equipo sea una necesidad. (Para capturar estas economías de escala asuma que involucrarse en la producción a cualquier nivel requiere de una unidad de capital, y que este requerimiento hace que la producción individual no sea rentable).

Sea  $e \in [0, 1]$  sea el esfuerzo por hora de trabajo (podiera simplemente ser la fracción de la hora durante cual el trabajador “trabaja” contrario a “no trabaja”). La producción por período es

$$y = y(be) + \varepsilon \quad \text{con } y' > 0 \quad \text{y con } y'' < 0 \quad (8.1)$$

donde  $b$  es el número de “horas” contratadas al trabajador (asumiendo sea sólo una “hora” por trabajador, por lo que  $b$  también es el número de trabajadores idénticos contratados), y  $\varepsilon$  es un término de error con media cero. El resultado se puede contratar, pero los niveles de rendimiento de trabajadores particulares no se pueden deducir de los niveles de producción debido a la naturaleza de producción en equipo y al término aleatorio en la función de producción.

Lo que sigue resume la interacción empleador – empleado. El principal (el empleador) conoce el máximo esfuerzo que realizará el agente (el trabajador),  $e(w, m; z)$ , dado un nivel de salario  $w$  y un nivel de monitoreo  $m$ , con una posición de reserva  $z$  del trabajador la cual se determina exógenamente (los argumentos de una función a la derecha del punto y coma son exógenos). Al inicio de un período, el empleador selecciona (con el objeto de maximizar los beneficios) y anuncia: una probabilidad de

terminación  $t(e, m) \in [0, 1]$  donde  $t_e < 0$  y  $t_m > 0$  sobre los rangos económicamente relevantes; un nivel de salario,  $w$ ; y un nivel de monitoreo por hora de trabajo contratado  $m$ . Los insumos, tanto de salario como de monitoreo, se miden en las mismas unidades desde producción por período. Luego del anuncio del empleador en cuanto a su estrategia para incentivar los niveles de esfuerzo, y por lo tanto con conocimiento de ello, el trabajador selecciona un nivel de esfuerzo  $e$  con el objeto de maximizar el valor presente de su utilidad vitalicia. Al final del período se le paga al trabajador y éste experimenta la utilidad en la cual incurre como resultado del esfuerzo que realiza y el salario que recibe, y se le renueva o se le termina el empleo, esto último ocurre con una probabilidad  $t(e, m)$ . Si al trabajador lo despiden del empleo (o le terminan el contrato), éste obtiene el valor presente de su utilidad vitalicia igual a  $z$  y se le reemplaza con un trabajador idéntico proveniente de un grupo de desempleados. Si el trabajador conserva el trabajo, esta misma interacción se llevará a cabo nuevamente durante el siguiente período; por lo tanto, la interacción es estacionaria (o sea, no varía en el tiempo).

El esquema de terminación  $t(e, m)$  es crucial para el funcionamiento del modelo. Un esquema simplificado de terminación se podría basar en la idea de que durante cualquier período existe una probabilidad,  $\eta(m)$ , de que el empleador “vea” al trabajador, en cuyo caso el empleador sabrá con certeza si el trabajador está o no trabajando. Supongamos que si no hay monitoreo el empleador no ve al trabajador, por lo tanto  $\eta(0) = 0$ , y que  $\eta' > 0$ . Esto daría un esquema de terminación  $t = \eta(m)(1 - e)$ , de donde se puede apreciar que  $t(0, \eta(m)) = \eta(m)$  y  $t(e, \eta(0)) = 0$ . Lo que es esencial para el modelo es que, para los niveles positivos de monitoreo, un aumento de esfuerzo reduce la probabilidad de terminación:  $t_e = -\eta(m)$ . De manera similar, un aumento en el monitoreo aumenta el efecto marginal de trabajar más duro para evitar la terminación de empleo:  $t_{em} = -\eta'(m)$ .

*La Mejor Respuesta del Trabajador.* La función de utilidad por período del trabajador es

$$u = u(w, e) \tag{8.2}$$

donde  $u_w \geq 0$  y  $u_e \leq 0$  sobre los rangos económicamente relevantes. Esto no significa que los empleados preferirían no ofrecer esfuerzo alguno, sino más bien que cualquier resultado para el cual  $u_e > 0$  no puede ser una asignación de equilibrio, pues en este caso el empleado pudiera unilateralmente implementar un nivel de esfuerzo más alto, con lo cual aumentarían tanto las utilidades del empleador como su propio beneficio. El trabajador varía  $e$  para maximizar el valor presente de su utilidad esperada a lo largo de un horizonte infinito, dada una tasa de preferencia intertemporal  $i$ :

$$v = \frac{u(w,e) + (1-t(e))v + t(e)z}{1+i} \quad (8.3a)$$

o utilizando el supuesto de estacionariedad y reorganizando términos:

$$v = \frac{u(w,e) - iz}{i + t(e)} + z \quad (8.3b)$$

donde el primer término del lado derecho de la expresión reorganizada es la *renta de ejecución* que se introdujo en el capítulo anterior; en este caso, también se denomina una *renta de empleo*. Por lo tanto, tenemos: *valor presente del trabajo* = *renta de empleo* + *posición de reserva*. Dado este objetivo, el trabajador selecciona  $e$  con el objeto de establecer que

$$v_e = 0 \quad (8.4)$$

lo cual requiere que:

$$u_e = t_e(v - z) \quad (8.5)$$

Por lo tanto, el trabajador escogerá el nivel de esfuerzo que iguale el costo marginal del esfuerzo al beneficio marginal del esfuerzo. Si comenzamos desde un nivel bajo de  $e$ , el trabajador debería aumentar el esfuerzo hasta que la desutilidad marginal del esfuerzo apenas compense la ganancia marginal en el valor presente de utilidad ocasionado por la reducción asociada en la probabilidad de terminación. Las arriba mencionadas condiciones de primer orden (8.4 ó 8.5) definen la función de mejor respuesta del trabajador, la cual se muestra en la figura 8.1.

El siguiente ejemplo puede dar claridad sobre la función de mejor respuesta. Consideremos una persona para quien el salario es un “bien” y el trabajo es un “mal” cuya desutilidad depende no solamente del nivel de esfuerzo sino también de qué tan justa es la recompensa. Supongamos que la función de utilidad del empleado es

$$u = w - \frac{aw^f / w}{1 - e}$$

donde  $a$  es una constante positiva y  $w^f$  es una norma de salario exógena llamada “salario justo”. La desutilidad del esfuerzo representada por el segundo término aumenta en el esfuerzo (a una tasa creciente). Obsérvese también que es decreciente en el salario relativo al salario justo, lo que indica que el trabajo duro que se recompensa más justamente es menos oneroso que un esfuerzo menor a un salario que se considera no es justo. La motivación subyacente puede reflejar una variante de la función de preferencia recíproca que se introdujo en el capítulo 3: el empleado puede tomar la oferta de salario como un indicativo del tipo de empleador y experimentar una menor desutilidad en el esfuerzo al trabajar duro para un jefe generoso o justo.



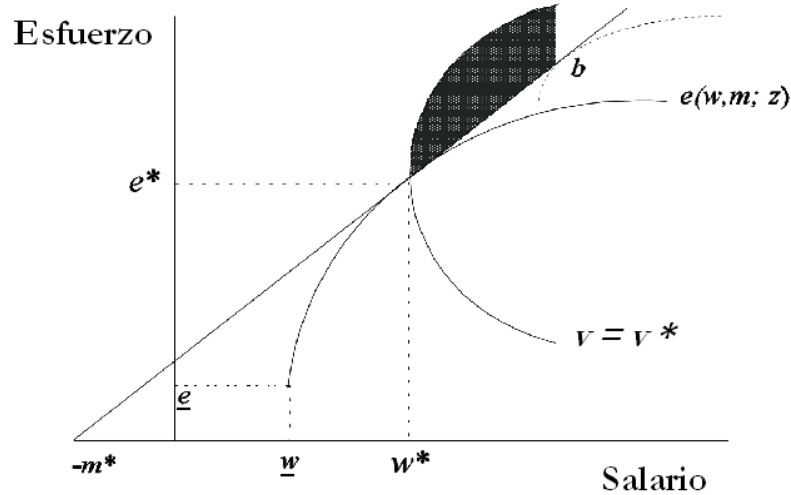


FIGURA 8.1 La función de mejor respuesta del empleado y la oferta óptima del empleador. El punto **a**, a saber, una oferta de salario de  $w^*$  y una respuesta de esfuerzo de  $e^*$ , cumplen, conjuntamente, las condiciones de primer orden de los problemas de optimización del empleado y del empleador. El nivel óptimo de monitoreo es  $m^*$ , cuya determinación no se muestra aquí (véase la figura 8.3). El punto **b** es uno de los óptimos de Pareto que conforman el locus de contrato eficiente (que no se muestra aquí). El área sombreada son los resultados que son Pareto superiores a **a**. La pendiente de la línea **ab** es  $e^*/(w^* + m^*)$ .

Supongamos que el empleador pueda, sin costo, observar al trabajador, pero la información no es verificable por lo que, como antes,  $e$  es no contractable, no hay monitoreo, y la función de terminación es simplemente  $t = 1 - e$ . La función de utilidad mencionada arriba implica que, para un salario finito, la desutilidad del esfuerzo se hace infinita a medida que  $e$  se acerca a 1, por lo que el empleado no escogerá  $e = 1$  y, como resultado, sabemos que  $t > 0$ . Asumamos que la posición de reserva del empleado se normaliza a cero y que la tasa de preferencia intertemporal también es cero (esta simplificación nos da una expresión matemática de forma cerrada para la función de mejor respuesta pero, obviamente, no es realista). Entonces, reescribiendo 8.3b, tenemos que

$$v = \frac{u(w, e)}{t(e)} = \frac{w - (aw^f/w)(1 - e)^{-1}}{1 - e}$$

y porque  $t_e = -1$ , podemos escribir la ecuación (8.5) para este caso como

$$-\frac{aw^f/w}{(1-e)^2} = -\frac{w - (aw^f/w)(1-e)^{-1}}{1-e}$$

esta función de mejor respuesta se puede escribir como una expresión explícita del esfuerzo del empleado (simplemente reorganizando los términos) como

$$e = 1 - \frac{2aw^f}{w^2}$$

Como era de esperarse, el nivel de esfuerzo es creciente y cóncavo en el salario, y decreciente en el nivel de salario justo. Podría ser instructivo para el lector derivar la misma función de mejor respuesta cuando se asumen preferencias convencionales (sin el motivo de equidad) tan sólo con eliminar el término  $(w^f/w)$  de la función de utilidad. Una comparación de las dos funciones de mejor respuesta muestra la importancia de las preferencias sociales.

Antes de continuar, es necesario hacer cuatro comentarios. Primero, debemos confirmar que la amenaza de despido implícita en el anuncio ex-ante de la función  $t(e, m)$  es creíble (es decir, será de interés para el empleador el llevarla a cabo ex-post – una vez que un trabajador perezoso haya sido detectado). ¿Por qué despediría un empleador a un trabajador solamente para contratar a otro trabajador idéntico? Asumamos que los empleados observan los niveles de esfuerzo el uno del otro y que cualquier terminación (del contrato) es de conocimiento común, entonces, si no se castigan los actos de falta o evasión de trabajo con un despido, los empleados cesarían de creer en el anuncio de terminación  $t(e, m)$ . Por lo tanto, es necesario despedir a los trabajadores que evaden el trabajo para sustentar la creencia de que la función de terminación es realmente vigente.<sup>2</sup> En segundo lugar, en un tratamiento más completo, el esquema  $t(e, m)$  (no solamente  $m$ ) sería diseñado por el empleador (si se despide o no a un trabajador puede depender, por

---

<sup>2</sup> El supuesto de que el juego es de conocimiento común y es estacionario significa que los trabajadores creerán que  $t(e, m)$  es vigente en cualquier caso. Sin embargo, al modelar un proceso dinámico mediante el cual los trabajadores aprenden la función de terminación como resultado de las terminaciones observadas agregaría una complicación substancial y poca claridad.

ejemplo, del costo que implique reclutar y entrenar a un reemplazo), pero hacer esto complica el modelo sin agregar mucha luz.

Tercero, el problema de optimización del horizonte infinito es simplemente una forma de obtener una función de mejor respuesta a partir del comportamiento del empleado; no es necesario describir el proceso de pensamiento del empleado. Puede que el empleado esté siguiendo una norma de trabajo (dictando un nivel dado de esfuerzo) que se desarrolla mediante el proceso de la actualización con base en la remuneración o pago que se describe en los capítulos 2 y 7. La ecuación 8.5 (la función de mejor respuesta) da la norma de trabajo que maximiza los pagos y por lo tanto la tendencia será adoptarla.

En cuarto lugar, uno podría preguntar ¿cómo es que el empleador conoce las funciones de mejor respuesta de los empleados? Así como el empleado pudiera llegar a una función de la mejor respuesta mediante métodos de prueba y error (con actualizaciones basadas en los pagos), el empleador puede llegar a obtener una estimación de las funciones de mejor respuesta variando la estrategia de disciplina laboral y observando los efectos sobre la producción total. Naturalmente, existen muchas circunstancias bajo las cuales este proceso de aprendizaje pudiera ser ineficiente o estar sesgado, pero asumiré que el empleador llega a una estimación exacta (recuerden: conocer la función de mejor respuesta no es lo mismo que ser capaz de escribir un contrato en  $e$  ya que  $e$  no es verificable).

*Maximización de Beneficios.* El empleador, quien se enfrenta a un mercado competitivo para su producto en el cual el precio dado es 1, varía  $m$ ,  $w$  y  $b$  para maximizar los beneficios esperados (asumimos que el empleador es neutral al riesgo).

$$\pi = y(he(w, m; z)) - (w + m)b \quad (8.7)$$

Las condiciones de primer orden para un máximo son,

$$\pi_h = y' e - (w + m) = 0 \quad (8.7a)$$

$$\pi_w = y' he_w - b = 0 \quad (8.7b)$$

$$\pi_m = y' he_m - b = 0 \quad (8.7c)$$

de las cuales podemos observar que la maximización de beneficios requiere que

$$e_w = \frac{e}{w+m} = e_m \quad (8.8a)$$

$$y' = \frac{w+m}{e} \quad (8.8b)$$

La primera condición requiere que el nivel promedio de esfuerzo por cada dólar gastado en trabajo sea igual al impacto marginal de las variaciones tanto en gastos de salario como en gastos de monitoreo. Esta es la llamada condición de Solow (debido a Robert Solow, quien fué el primero en derivar dicha condición) extendida para incluir monitoreo como un insumo. La otra condición de primer orden es análoga a la condición familiar para una maximización de beneficios, y es que el salario sea igual al producto marginal del trabajo. Al ser endógeno el esfuerzo, esta condición requiere que la productividad marginal del *esfuerzo* sea igual al costo de una unidad de *esfuerzo* (incluyendo el costo de monitoreo). Expresado equivalentemente como  $y' e^* = w^* + m^*$ , las condiciones de primer orden requieren que la productividad marginal del tiempo de trabajo (evaluado a los niveles que determina la condición de Solow) sea igual al costo por hora de una hora de trabajo, como se muestra en la figura 8.2.

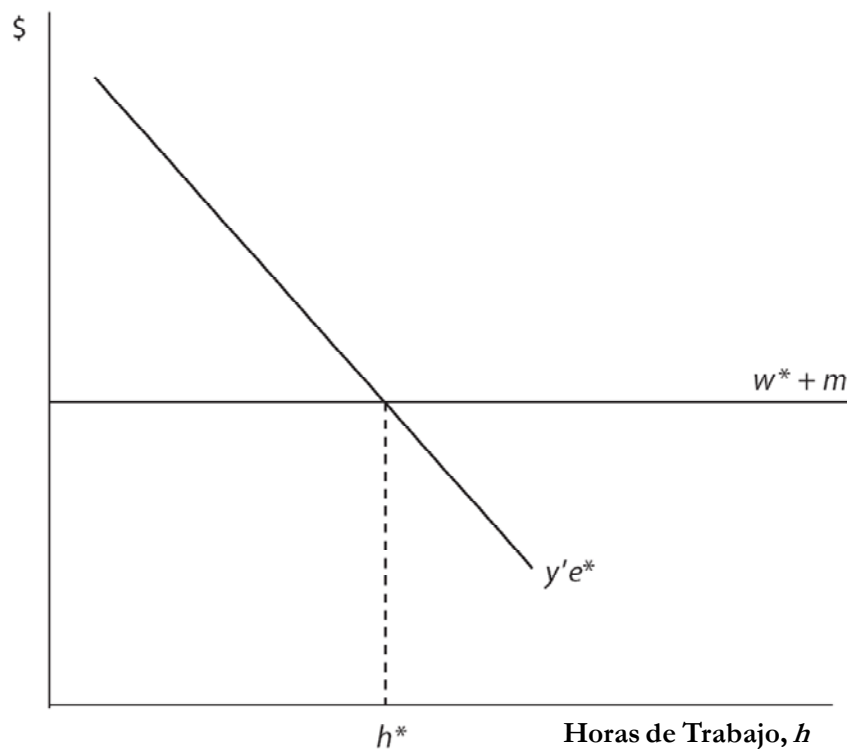


FIGURA 8.2 El nivel óptimo de contratación de la empresa. Obsérvese que:  $w^*$ ,  $m^*$ ,  $h^*$ , y  $e^*$  son la solución a las condiciones de primer orden en el texto.

Debido a que  $h$  no aparece en la función de mejor respuesta del trabajador, se puede describir el proceso de maximización de beneficios de manera secuencial, como sigue: primero, el empleador resuelve el problema de disciplina laboral, y selecciona  $m$  y  $w$  para satisfacer la ecuación (8.8a). Luego, substituyendo  $e^*$  y  $w^*$  de la ecuación (8.8a) en la ecuación (8.8b), determina cuántas horas de trabajo debe contratar. Finalmente, substituyendo  $e^*$ ,  $w^*$ ,  $m^*$ , y  $h^*$  en la ecuación (8.7), determina si este plan de producción es suficientemente rentable como para llevarlo a cabo, dados los usos alternativos que tiene el capital requerido.

A objeto de ilustrar el contrato de equilibrio, regresemos al ejemplo de arriba. Recordemos que  $m = 0$ . Utilizando la función de mejor respuesta (8.6), se establecerá el salario de tal manera que satisfaga la ecuación (8.8a) o

$$\frac{e}{w} = \frac{1 - 2aw^f/w^2}{w} = \frac{4aw^f}{w^3} = e_w$$

obteniendo el nivel de salario óptimo  $w^*$  como

$$w^* = (6aw^f)^{1/2}$$

Si  $a = 1$  y si  $w^f = 6$ , entonces ofrecer el salario justo es óptimo para el empleador. La respuesta del esfuerzo óptimo del empleado a la oferta del salario óptimo del empleador se encuentra substituyendo este valor de  $w^*$  en la ecuación (8.6), lo que da como resultado  $e^* = 2/3$ . No obstante, si  $w^f = 24$ , el salario que maximiza el beneficio es la mitad del salario justo. Para salarios justos menores a 6, resulta óptimo para el empleador exceder la norma del salario justo.

*Elección de la Tecnología.* Consideremos ahora una función de producción más general con un insumo diferente al trabajo de  $y(k, E)$ , donde  $k$  es el insumo no laboral por período,  $E = be$  es el insumo total de esfuerzo y, como antes, la función es creciente y cóncava en sus argumentos. Supongamos que las variaciones en  $k$  están asociadas con arreglos espaciales que están en desacuerdo y otras medidas del proceso de producción que afectan la facilidad para monitorear el proceso de trabajo. Por ejemplo, procesos altamente intensivos en capital, tales como las líneas de ensamblaje de las cuales Henry Ford fue pionero, puedan funcionar “al paso de las máquinas,” facilitan en mayor grado la identificación de los trabajadores que hacen un nivel de esfuerzo bajo. Para reflejar este hecho, la función de terminación es ahora  $t = t(e, m, k)$ ; un ejemplo pudiera ser  $t = \eta(m, k)(1 - e)$ , donde  $\eta(m, k)$  es la probabilidad de que se detecte a un empleado que no está trabajando. Como antes,  $\eta(\cdot)$  es creciente en  $m$ . Si  $\eta$  es creciente en  $k$  (como lo sugiere el ejemplo de la línea de ensamblaje), entonces  $t_{ek} < 0$ ; debido a que una tecnología más intensiva en  $k$  facilita el proceso de monitoreo, se aumenta el efecto (negativo) del esfuerzo sobre la probabilidad de terminación. En este caso podemos decir que el proceso de producción  $k$ -intensivo es más “transparente” desde el punto de vista del monitor y que un proceso menos  $k$ -intensivo es más “oscuro.” Casos contrarios también existen. Lo importante no es el signo de  $t_{ek}$  sino el hecho de que la elección de la tecnología en general afectará de alguna manera la facilidad de monitoreo, a saber  $t_{ek} \neq 0$ .

¿Cuál será el efecto de las variaciones en  $k$  sobre la función de mejor respuesta del empleado? Si utilizamos la nueva función de terminación y si diferenciamos totalmente la ecuación (8.5) con respecto a  $k$  y  $e$ , tenemos que

$$\frac{de}{dk} = \frac{t_{ek}}{u_{ee} - t_{ee}}$$

lo que, al utilizar la condición de segundo orden para el problema de maximización del empleado, nos muestra que  $de / dk$  toma el signo de  $-t_{ek}$ . Por lo tanto, si las tecnologías  $k$ -intensivas son más transparentes, aumentos en  $k$  desplazan hacia arriba la función de mejor respuesta (al aumentar el beneficio marginal del empleado al aumentar su esfuerzo). Este efecto se reflejará en la elección del nivel de  $k$  que maximiza los beneficios. Si dejamos que  $\rho$  sea el precio de alquiler por período de una unidad de  $k$ , y si diferenciamos parcialmente la función de beneficios (al utilizar la función aumentada de producción) con respecto a  $k$ , tenemos ahora la siguiente condición adicional de primer orden:

$$\pi_k = y_k + e_k h y_E - \rho \quad (8.7 \text{ d})$$

Por lo tanto, la elección del insumo  $k$  no igualará el precio de alquiler del componente de entrada  $k$  a su productividad marginal sino a su productividad marginal *más* su efecto sobre el suministro de esfuerzo multiplicado por la productividad marginal del esfuerzo. La presencia de este “efecto de la disciplina de trabajo” sobre la elección de tecnología significa que, en general, no se tiene que

$$\frac{y_k}{y_E} = \frac{\rho}{\mu}$$

donde  $\mu = (w + m) / e$  es el costo de una unidad de esfuerzo. Como resultado, la tasa marginal de sustitución en la producción (la pendiente de una isocuanta) no será igual factoría relación de precios de los factores en equilibrio competitivo. La razón es que los factores de producción son valorados no solamente por su contribución a la producción sino también por sus efectos sobre el ambiente de disciplina laboral. (El monitoreo es un caso puro de tal insumo, pues no aparece en absoluto en la función de producción).

Durante el resto del capítulo, por motivos de simplicidad, ignoraré el componente no laboral,  $\kappa$ .

### CARACTERÍSTICAS DE LA TRANSACCIÓN DE EQUILIBRIO

Los valores de  $e$ ,  $b$ ,  $e$ , y  $m$  que satisfacen las ecuaciones (8.5) y (8.8) constituyen la transacción de equilibrio, a saber, una mejor respuesta mutua por parte del empleador y del empleado. Obsérvense cinco cosas en cuanto al equilibrio.

Primero, los trabajadores generalmente enfrentan restricciones de cantidad. En general, la restricción de participación no obliga (no se cumple con igualdad), es decir,  $v^* > \bar{z}$ . Esto implica que el mercado laboral no se vacía: trabajadores idénticos que reciben  $\bar{z}$  preferirían estar empleados y recibir  $v$  pero no son capaces de conseguirlo. Aquellos trabajadores incapaces de hacer una transacción que les permita obtener  $v$  están *limitados por la cantidad*, incapaces de comprar o vender tanto como quisieran a los términos de intercambio existentes.

Segundo, el intercambio resultante  $(e^*, n^*)$  es *Pareto ineficiente*. Este debe ser el caso ya que, a estos valores, las condiciones de primer orden del empleador y del empleado requieren que

$$\begin{aligned} v_e = 0 & \quad \text{pero } \pi_e > 0 \\ \text{y} & \\ v_n > 0 & \quad \text{pero } \pi_n = 0; \end{aligned} \tag{8.9}$$

y, por lo tanto, existen algunos valores  $(\Delta e, \Delta n)$  (suficientemente pequeños) tales que

$$v(e^* + \Delta e, n^* + \Delta n) > v(e^*, n^*)$$

y

$$\pi(e^* + \Delta e, n^* + \Delta n, \dots) > \pi(e^*, n^*)$$



Por lo tanto, existe un pequeño aumento en el esfuerzo que acompañado por un pequeño aumento en el salario sería consistente con un mejoramiento en el sentido de Pareto. Dado que el empleador ha escogido no solamente  $w$  sino también  $m$  para maximizar los beneficios, una demostración análoga muestra que una leve disminución en el monitoreo y un leve aumento en el esfuerzo también implican un mejoramiento en el sentido de Pareto.<sup>3</sup>

Tercero, *trabajo no productivo* así como otros insumos no productivos serán contratados en un equilibrio competitivo. Un ejemplo de factores de producción no productivos son aquellos que se utilizan para monitorear a los trabajadores. Estos inputs no aparecen en la función de producción pero son contratados por la empresa para maximizar beneficios pues ellos contribuyen en forma alternativa con el cumplimiento de los objetivos de la empresa. Sabemos que factores de producción puramente no productivos serán contratados porque  $m = 0$ ,  $t_e = 0$  (sin monitoreo, el trabajar más duro no altera la probabilidad de terminación) por lo que  $e(w, 0; z) = \underline{e}$  (el nivel de esfuerzo de reserva es escogido). Si asumimos que  $e = \underline{e}$  los beneficios no se maximizan, lo cual implica que  $m^* > 0$ , el nivel óptimo de monitoreo debe ser positivo.

Cuarto, el equilibrio competitivo es *técnicamente ineficiente*: existe una asignación alternativa en la cual se obtiene el mismo nivel de producción utilizando menos de un insumo y no más de cualquiera otro (esto define la ineficiencia técnica). Suponga que al empleador se le exigiera (por parte de una Entidad omnipotente) que eleve el salario en  $\Delta w$  y se le instruyera disminuir el nivel de monitoreo en una cantidad  $\Delta m$ , justo lo suficiente para restaurar el nivel de esfuerzo al nivel de equilibrio, entonces,

$$e(w^*, m^*; z) = e(w^* + \Delta w, m^* - \Delta m, z) \quad (8.10)$$

Si la Entidad también estipula que las horas de trabajo continúen siendo las mismas, la producción se mantendrá inalterada. Pero se ha disminuido uno de los

---

<sup>3</sup> Al enfoque de disciplina laboral se le denomina, a veces, el modelo del “salario de eficiencia” ya que Leibenstein (1957) y otros contribuyentes tempranos a esta literatura sugirieron que para tomar en cuenta los efectos de nutrición, esfuerzo variable y similares, se debería medir el trabajo en “unidades de eficiencia” y no en horas. La costumbre no cambió pero es un término erróneo ya que (en contraste al modelo Walrasiano) los equilibrios que describe el modelo son técnicamente ineficientes (ver más abajo) y Pareto ineficientes.

componentes, el monitoreo: la reducción en el monitoreo, dada por  $\Delta m$ , libera recursos que pueden ser utilizados productivamente. Por lo tanto, el equilibrio competitivo  $(w^*, m^*, b^*)$  es técnicamente ineficiente en el sentido estándar que se definió anteriormente. Este caso se ilustra en la figura 8.3.

Las razones para la ineficiencia son instructivas. Las estrategias de ejecución combinan típicamente el monitoreo (que tiene un costo de oportunidad social ya que utiliza recursos con usos alternativos, por ejemplo, el trabajo del monitor, o los recursos que se requieren para producir el equipo de vigilancia) con una renta de ejecución, en este caso  $v - \zeta$  (que es una transferencia pura y, en consecuencia, no implica un costo de oportunidad social). Por lo tanto, y dado que tanto el monitoreo como el salario implican un costo para el empleador, mientras que solamente el monitoreo es socialmente costoso, tenemos un caso estándar en el cual los costos marginales privados son diferentes a los costos marginales sociales, produciendo como resultado, una típica falla de mercado. Por lo tanto, desde un punto de vista de eficiencia social, las estrategias de disciplina laboral que se determinan de manera competitiva hacen un uso generalmente exagerado del monitoreo y sub-utilizan las rentas de ejecución. Más zanahorias y menos garrote podrían influir sobre una mejora técnica eficiente. Obsérvese que si se asocian tecnologías más intensivas en capital con procesos más transparente de producción (como en el ejemplo anterior), la misma demostración es válida para los bienes de capital: se podrían hacer mejoras en técnicas eficientes (con respecto a la transacción de equilibrio competitivo) elevando los salarios y disminuyendo el consumo de capital.

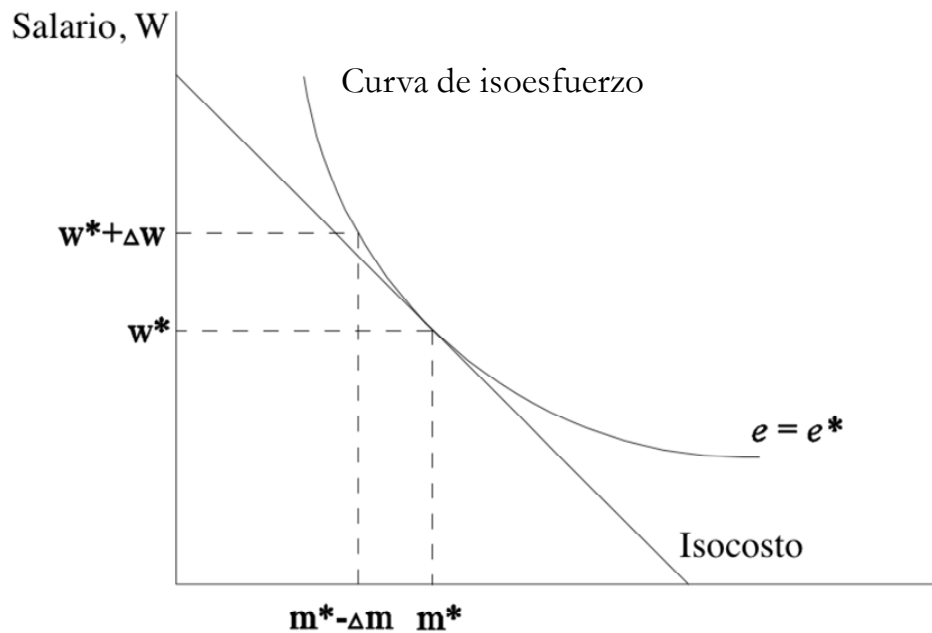


FIGURA 8.3 El nivel de monitoreo de maximización de beneficios es técnicamente ineficiente.

Quinto; la transacción de equilibrio también se caracterizará por *comodidades en el lugar de trabajo que son subóptimas en el sentido de Pareto*, tales como un horario flexible, un ambiente de trabajo respetuoso y seguro, y similares. En el modelo Walrasiano estándar, el empleador está limitado por la decisión de suministro laboral de los trabajadores (restricción de participación) y, por tal razón, el empleador se ve motivado a proveer comodidades en el lugar de trabajo como una manera de disminuir los costos de trabajo: un trabajo cuyo ambiente laboral sea más amable para el trabajador puede atraer potenciales empleados con salarios más bajos. Dado que la restricción de participación es la propia utilidad del trabajador en la siguiente mejor alternativa que tenga a disposición, el empleador maximizará los beneficios evaluando la importancia de las comodidades en el lugar de trabajo (en relación con otros argumentos de la función de utilidad del trabajador) exactamente como lo hace el trabajador. ¿Se tiene este mismo resultado cuando el esfuerzo no es contratable? Ya veremos que no se tiene.

Supongamos que la función de utilidad del empleado se amplía para incluir una medida de las comodidades de trabajo que se ofrecen (por hora de trabajo),  $\alpha$

$$u = (w, \alpha, e)$$

con  $u_\alpha > 0$  sobre el rango relevante, y que una unidad de las comodidades que el empleador debe proveer cuesta  $p$  por hora de trabajo contratada. Entonces tenemos un nuevo valor presente del trabajo  $v(e, w, \alpha, z)$ , así como una nueva función de mejor respuesta  $e(w, m, \alpha, z)$  y una condición adicional de primer orden para el empleador

$$\pi_\alpha = y' h e_\alpha - h p = 0 \quad (8.7e)$$

Esta condición requiere que el ingreso marginal producto de las comodidades (el primer término) sea igual al costo marginal (y promedio) de proveer dichas comodidades. Es claro que el empleador tendrá en cuenta, en alguna medida, las preferencias de los trabajadores sobre las comodidades laborales ya que  $e_\alpha > 0$ ; el tener un trabajo más amigable motivará al trabajador a realizar un mayor esfuerzo (el valor del trabajo para el trabajador será ahora mayor).

¿Tendrá el empleador en cuenta *lo suficientemente* las preferencias del empleado? La respuesta es no. Las comodidades en el lugar de trabajo no se diferencian de los salarios en este modelo; éstas son valoradas por el trabajador pero para el empleador son costosas de proveer. Ya hemos visto que la oferta que hace el empleador para maximizar sus beneficios  $(w^*, e^*)$  será Pareto inferior a alguna otra combinación de  $e$  y  $w$  caracterizada por pequeños aumentos en ambos. El mismo razonamiento aplica para las condiciones de trabajo: porque en el punto de equilibrio competitivo  $(e^*, w^*, \alpha^*, m^*)$

$$\pi_\alpha = 0 \quad \text{y} \quad v_\alpha > 0 \quad (8.11)$$

$$v_e = 0 \quad \text{y} \quad \pi_e > 0 \quad (8.12)$$

por lo que una leve mejoría en las comodidades en el lugar de trabajo, acompañada de un pequeño aumento en el esfuerzo, inducirán una mejora en el sentido de Pareto.

¿Qué es lo que explica la diferencia entre el tratamiento Walrasiano y el tratamiento post-Walrasiano? En el primero, la restricción de participación es activa (se

cumple con igualdad), por lo que en equilibrio, el locus de isobeneficios de la empresa es tangente al locus de indiferencia de los trabajadores, lo que da como resultado un contrato de trabajo óptimo en el sentido de Pareto. En el segundo caso, la restricción de participación no es activa (se cumple con desigualdad) y, en su lugar, la empresa es retringida por la función de mejor respuesta del empleado. Dado que la función de mejor respuesta no coincide con la restricción de participación, el segundo equilibrio no es un óptimo de Pareto.

## EL MERCADO DE TRABAJO EN EQUILIBRIO GENERAL

La relación de empleo en una empresa está inmersa en un sistema de mercado compuesto por muchas de estas empresas y otros actores. Para estudiar esto, supongamos que existe un número suficientemente grande de empresas idénticas que utilizan trabajo como se mencionó anteriormente y que los mercados relevantes son perfectamente competitivos en el sentido de que no existen barreras para entrar o salir. Si los beneficios de la empresa (ingresos netos menos el costo de oportunidad del capital) son positivos habrá incentivos para que nuevas empresas entren; si los beneficios son negativos, éstos inducen a las empresas a salirse del mercado. Por lo tanto, el número de empresas en equilibrio se determina mediante las condiciones de primer orden arriba mencionadas y la *condición de cero beneficios*:

$$\pi = y(h e(w, m, z)) - (w + m)b - \delta = 0 \quad (8.13)$$

donde  $\delta$  es el costo dado por período de insumos fijos (la unidad de capital) y  $b, e, m, w$  satisfacen las condiciones de primer orden arriba mencionadas. Obsérvese que  $z$ , (la única variable en la ecuación 8.13 a parte de  $\delta$  que no se determina por las condiciones de primer orden mencionadas anteriormente), es ahora endógena. ¿Cómo se determina  $z$ ?

*La Posición de Reserva del Trabajador.* Para algunos valores de  $e$  y de  $w$ ,  $v(e, w) = z$  por lo que el trabajador es indiferente entre su trabajo – o sea, proveer un esfuerzo  $e$  y recibir un salario  $w$  – y su siguiente mejor alternativa, a saber  $z$ . La restricción de participación

del trabajador se satisface como una igualdad. Podemos ver, de la ecuación (8.5), que en este caso se tiene que obtener  $u_e = 0$  (el nivel de esfuerzo escogido cuando la renta del empleo es cero es aquél para el cual la desutilidad del trabajo es cero). Por lo tanto la utilidad de la transacción  $(\underline{e}, \underline{w})$  es el equivalente por período de  $\underline{z}$  ó  $u(\underline{e}, \underline{w}) = i\underline{z}$ . El nivel de trabajo  $\underline{e}$  es por lo tanto la cantidad de trabajo por hora que un trabajador escogería para llevar a cabo en ausencia de cualquier estrategia de incentivos por parte del empleador.

Pero, ¿qué es  $\underline{z}$ ? Ya que se asume que el trabajo es idéntico, el salario esperado por el trabajador en un empleo alternativo debe ser igual al salario en su trabajo actual, por lo que el costo de ser despedido es la reducción en el bienestar experimentado durante la temporada que dure desempleado. Un trabajador despedido pasará el siguiente período sin empleo, y recibirá un beneficio de desempleo (u otras transferencias que reemplacen el ingreso) igual a  $b$  mientras no realiza ningún trabajo (y presumiblemente iniciará la búsqueda de un trabajo). Por lo tanto, el desempleado experimenta una utilidad por período dada por  $u(b, 0)$  que puede reflejar la utilidad del ocio, el estigma social de no tener trabajo, y similares. Al final de cada período existe una probabilidad  $\lambda$  de que el trabajador desempleado encuentre trabajo y con ello salga del grupo de desempleados; Así, la duración esperada de estar desempleado es  $1/\lambda$ . Por lo tanto,

$$\begin{aligned} z &= \frac{u(b,0) + \lambda v + (1-\lambda)z}{1+i} \\ &= \frac{u(b,0) + \lambda v}{i + \lambda} \end{aligned}$$

Esta es la posición de reserva definida de la misma forma que el valor presente del trabajo. Vemos que  $d\underline{z}/d\lambda > 0$  si  $v - \underline{z} > 0$ , lo que requiere que  $iv - u(b, 0) > 0$ . Esto significa que un aumento en la probabilidad de dejar el grupo de desempleados mejora la posición de reserva del trabajador, siempre y cuando los beneficios por período de tener un trabajo ( $iv$ ) excedan los beneficios por período de no tener un empleo ( $u(b, 0)$ ).

*Estática Comparativa.* Recordemos que  $(w + m)/e = \mu$  es el *costo de una unidad de esfuerzo*. Debido a que un aumento en la posición de reserva del empleado desplaza su función de mejor respuesta hacia la derecha, se puede mostrar fácilmente que  $d\mu / dz > 0$ , lo que significa que el costo por unidad de esfuerzo varía con  $z$ , como era de esperarse. Como consecuencia, los beneficios también varían inversamente con  $z$ , a saber  $d\pi / dz < 0$ .

En equilibrio general competitivo la posición de reserva del trabajador ( $z$ ) debe ser tal que los niveles de  $e$ ,  $m$ ,  $h$ , y  $w$ , escogidos por las empresas y los trabajadores para maximizar los beneficios y las utilidades satisfagan la condición de cero beneficios. Es la entrada y salida de empresas motivada por beneficios positivos o negativos y los efectos agregados de empleo resultantes los que hacen que este nivel de equilibrio sea  $z$ . El proceso es el siguiente. Con  $n$  empresas que producen, donde cada una de ellas utiliza  $h$  tal y como es definida por las condición de primer orden (8.7), el empleo total  $H$  es definido por  $nh \equiv H$ , donde yo normalizo el suministro de trabajo a unidad, por lo que  $H$  es la *tasa de empleo agregado*. La probabilidad de salir del grupo de desempleados varía con el nivel de empleo, por lo que

$$\lambda = \lambda(H, \dots) \quad \text{con } \lambda' > 0$$

de donde sabemos, dado  $(dx / d\lambda > 0)$ , que

$$z = z(H, \dots) \quad \text{con } z' > 0$$

por lo que la posición de reserva del trabajador mejora cuando aumenta la tasa de empleo, como era de esperarse. Supongamos ahora que el número de empresas es tal que  $\pi > 0$ , lo que motiva la entrada de empresas adicionales. El empleo adicional resultante eleva  $H$ , lo que eleva  $z$  y a su vez eleva el costo por unidad de esfuerzo. La entrada de empresas continúa hasta que la ecuación (8.13) se cumpla, con lo cual se determina el nivel agregado de empleo de equilibrio  $H$  así como  $z(H)$ .<sup>4</sup>

<sup>4</sup> El equilibrio existirá siempre y cuando sean posibles beneficios positivos cuando  $H = 0$  y los beneficios sean negativos si la demanda es igual a la oferta de trabajo, una condición suficiente para esto es que  $z(1) > y(h) - \delta$ . Ya que  $\pi(H)$  es monotónica, el equilibrio es único.

## MEJORAS PARETIANAS NEGOCIADAS

Si mejoras en el sentido de Pareto son técnicamente factibles con respecto al equilibrio competitivo (los puntos en la lente sombreada de la figura 8.1), ¿Por qué no se realizan? ¿Por qué el trabajador y el empleador no llegan a un acuerdo sobre un nivel levemente más alto tanto de salario como de esfuerzo? ¿O por qué no llegan a un acuerdo sobre una mejora en las comodidades del trabajo, como un horario flexible, y un mayor nivel de esfuerzo? La respuesta es que un acuerdo así no se puede hacer cumplir. El salario y las horas son contratables, pero el nivel de esfuerzo no lo es. Tales acuerdos son técnicamente factibles pero no son factibles desde el punto del comportamiento de los agentes dada la estructura de información del problema y las instituciones que definen la interacción. Por lo tanto, el hecho de que  $e^*$ ,  $w^*$ ,  $m^*$ ,  $h^*$  sea ineficiente en el sentido de Pareto no nos dice si dicha ineficiencia se puede eliminar o atenuar la ineficiencia con contratos alternativos factibles u otros arreglos institucionales. Si el trabajador fuera el único empleado, entonces la propiedad de los activos de la empresa podría transferirse al trabajador quien, como un productor autoempleado y demandante residual del flujo de ingresos que resultan de sus esfuerzos, se podría dispensar de la necesidad de monitoreo. Como veremos en el capítulo 9, los beneficios de la elección óptima de esfuerzo en ausencia de monitoreo cuando el trabajador es el mismo propietario pudieran ser suficientes como para pagar al antiguo empleador un retorno fijo sobre los activos, suficiente para compensarle por la pérdida de dichos activos. Pero incluso en ausencia de economías de escala que permiten esta solución tipo Robinson Crusoe, la misma no sería factible si el empleado fuera adverso al riesgo o estuviera limitado en el crédito, como veremos en el capítulo 9. En estos casos, el trabajador pudiera preferir continuar trabajando bajo los contratos descritos anteriormente en lugar de ser un trabajador-propietario, incluso si *se le otorgaran* los activos.

Consideremos otra posible solución institucional. Supongamos que el equipo de trabajo está organizado como un sindicato y puede negociar con el empleador. Debido a que los miembros del equipo de trabajo son idénticos, el sindicato simplemente implementa las decisiones unánimes de sus miembros. Más aún, suponga que la información de los miembros del equipo de trabajo sobre las otras acciones de trabajo que ellos mismos realizan les permite utilizar el monitoreo de sus propios compañeros



para implementar un nivel común de esfuerzo de trabajo. Esto significa que la transacción ya no está restringida por las funciones individuales de mejor respuesta de los trabajadores. Los resultados de equilibrio pueden entonces incluir pares  $\{w, e\}$  por encima de la función de mejor respuesta si un acuerdo entre el sindicato y el empleador puede ser garantizado. Asumamos que el resto de la interacción tal y como se describió anteriormente permanece inalterada. En particular, la función de terminación del empleador sigue vigente y el empleador determina el nivel de empleo en la forma usual, a saber, igualando el ingreso marginal producto del esfuerzo con su costo efectivo. Naturalmente el sindicato puede negociar sobre los tipos de monitoreo, la función de terminación, y el nivel de empleo, pero el introducir estas complicaciones no iluminaría el punto principal en lo que sigue.

Al caer en cuenta de la posibilidad de una mejora en el sentido de Pareto con respecto al equilibrio competitivo  $\{w^*, e^*\}$ , tanto el empleador como el trabajador prometen ofrecer respectivamente  $\{w^+ > w^*\}$  y  $\{e^+ > e^*\}$ , donde  $\{w^+, e^+\}$  es una mejora en el sentido de Pareto con respecto a  $\{w^*, e^*\}$ , el equilibrio no cooperativo de Nash descrito anteriormente. Dicha mejora puede ser cualquier pareja  $\{w^+, e^+\}$  en la lente de mejoras en el sentido de Pareto de la figura 8.1. Las dos partes están por lo tanto comprometidas en una interacción de negociaciones en la cual el conjunto de negociación es la totalidad de la lente de mejoras en el sentido de Pareto y la frontera de negociación es el locus de contratos eficientes. La posición de reserva en este problema de negociación no es que el empleador y el equipo de trabajo se rehúsen a negociar del todo, sino más bien que ellos negocien en el nivel Pareto-inferior no cooperativo de  $\{w^*, e^*\}$ . El problema de la negociación, con pagos por un período, se ilustra en la figura 8.4.

Si se pudiera establecer un acuerdo vinculante para implementar las dos ofertas  $w^+$  y  $e^+$  entonces esperaríamos que un resultado como  $\{w^+, e^+\}$  fuese bastante común, al menos en aquellas situaciones en las que los empleados pueden participar en el monitoreo de los compañeros y negociar colectivamente con sus empleadores. No obstante, un acuerdo como  $\{w^+, e^+\}$  podría ser imposible de lograr a través de un acuerdo vinculante. Por ejemplo, si existen otros factores no observables que afecten el resultado de la producción, el empleador no puede detectar violaciones a lo acordado

sobre los niveles de esfuerzo al observar el nivel agregado de producto , lo cual generalmente sucede.

Si no se puede hacer un contrato vinculante para ejecutar el acuerdo, el empleador y el sindicato podrían implementar un resultado Pareto-superior adoptando estrategias de cooperación condicional (un buen ejemplo de la ley del Tali3n): cada uno implementa la parte que le corresponde en la transacci3n que genera una mejora en el sentido de Pareto  $\{w^+, e^+\}$  siempre y cuando el otro haga lo mismo pero deja de hacerlo, pasando al resultado Pareto inferior  $\{w^*, e^*\}$  si el otro incumple. Aunque estas estrategias suenen un poco abstractas, a menudo se observan variantes de las mismas. No es extra3o que los sindicatos amenacen con “trabajar de acuerdo con el reglamento” – a saber, acometer solamente las tareas que son requeridas expl3citamente por contrato, mientras que los empleadores a menudo condicionan pagos de salarios m3s altos a cambios en las reglas de trabajo que implementen niveles de esfuerzo m3s altos.

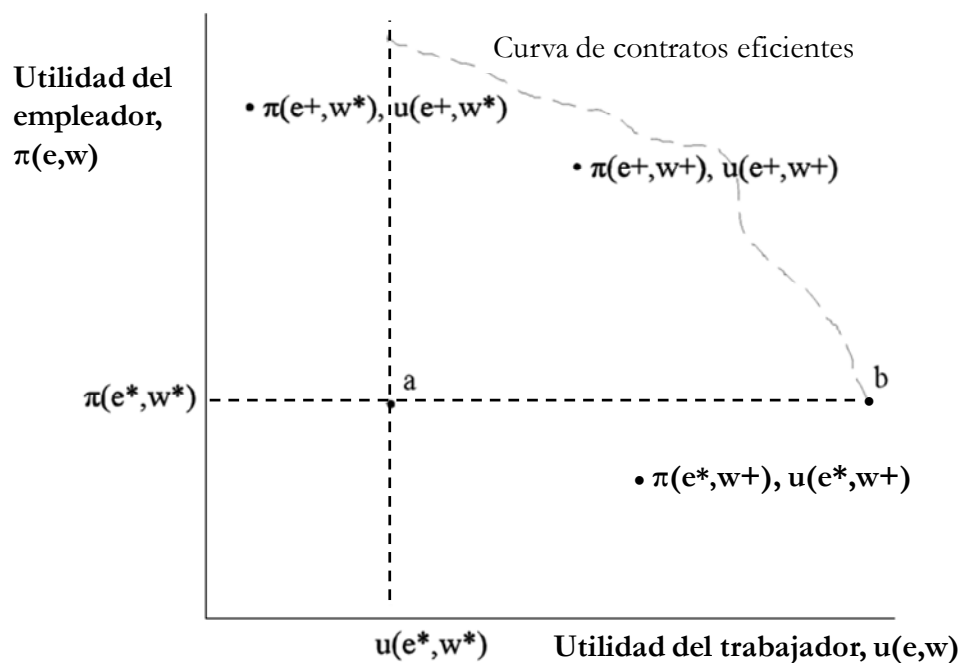


FIGURA 8.4 El problema de negociación del empleador y del sindicato: los pagos por período. Nota: El conjunto de negociación es el área limitada por los pagos de la interacción no cooperativa y el locus de contratos eficientes. Si las estrategias disponibles fueran  $w^+$  (sin restricción) y  $w^*$  para el empleador, y fueran  $e^+$  (sin restricción) y  $e^*$  para el empleado, el juego es un dilema de los prisioneros. El punto **a** es el equilibrio del juego no cooperativo (que corresponde con el punto **a** en la figura 8.1) mientras que el punto **b** es un punto sobre el locus de contratos eficientes (que corresponde con el punto **b** en la figura 8.1).

Supongamos que el conjunto de acciones para cada uno está restringido a solamente  $\{e^*, e^+\}$  para el sindicato y a  $\{w^*, w^+\}$  para el empleador y sus conjuntos de estrategias son jugar incondicionalmente los valores de equilibrio no cooperativo, o las estrategias de cooperación condicional (la ley del Tali3n) descritas anteriormente. Pudiera haber muchas otras estrategias pero estos conjuntos truncados de estrategias son suficientes para mostrar lo que me interesa. Usando la notaci3n anterior, los valores presentes de los pagos esperados de los dos negociadores aparecen en la tabla 8.1. Consideremos el pago esperado para el sindicato de adoptar incondicionalmente  $e^*$  si la empresa ha ofrecido el pago condicional  $w^+$ . En el primer per3odo, cada uno de los trabajadores obtiene la utilidad por per3odo de recibir un alto pago por poco trabajo  $u(e^*, w^+)$  y posteriormente es despedido con probabilidad  $t(e^*)$ , recibiendo como resultado su posici3n de reserva  $z$ , y permanece con probabilidad  $(1 - t(e^*))$ , pero de ah3 en adelante recibe los pagos correspondientes al equilibrio no cooperativo (porque la empresa deja de cooperar en

respuesta a la acción  $e^*$  del sindicato). El resto de los pagos pueden ser interpretados en forma similar (tabla 8.1).

TABLA 8.1

Valor presente de los pagos esperados en el juego repetido de negociación

	Empleador	$w^+$ <i>Condicional</i>	$w^*$ <i>Incondicional</i>
Sindicato			
$e^+$ <i>Condicional</i>		$v^+ = \frac{u(e^+, w^+) - iz}{i + t(e^+)} + z$	$\frac{u(e^+, w^*) + (1 - t(e^+))v^* + t(e^+)z}{1 + i}$
		$\pi^+ = \frac{\pi(w^+, e^+)}{i}$	$\frac{\pi(e^+, w^*) + \pi^*}{1 + i}$
$e^*$ <i>Incondicional</i>		$\frac{u(e^*, w^+) + (1 - t(e^*))v^* + t(e^*)z}{1 + i}$	$v^* = v(e^*, w^*, z)$
		$\frac{\pi(e^*, w^+) + \pi^*}{i}$	$\pi^* = \frac{\pi(w^*, e^*)}{i}$

Sabemos que  $v^+ > v^*$ , por lo que para un  $i$  suficientemente bajo se puede demostrar fácilmente que la acción  $e^+$  puede ser la mejor respuesta a  $w^+$  la acción en el contrato condicional. La ganancia de un período para el empleado, hecha posible por el resultado del alto pago que recibe por hacer poco esfuerzo  $\{e^*, w^+\}$  es más que compensada por la diferencia entre  $v^+$  y  $v^*$  (así como por la mayor probabilidad de que un miembro del sindicato que juegue  $e^*$  sea despedido al final del primer período y, por lo tanto, reciba  $z$ ). De manera similar, para un  $i$  suficientemente bajo, el  $w^+$  será la mejor respuesta a  $e^+$  la acción en el contrato condicional. Por lo tanto, el resultado  $\{w^+, e^+\}$  se puede implementar bajo ciertas condiciones. Naturalmente, si la ganancia obtenida en un periodo por no cooperar (abandonar el acuerdo), es lo suficientemente grande, o si la probabilidad de mantener el trabajo lo suficientemente pequeña, o la tasa de preferencia por el tiempo suficientemente grande, no se podrá implementar el resultado cooperativo como un equilibrio de Nash.

Esta estructura – la mejora negociada a partir de una interacción no cooperativa hasta un resultado cooperativo que sea Pareto superior – captura hechos empíricos importantes en cuanto a las relaciones de trabajo. A menudo uno encuentra lugares de trabajo cooperativos y no cooperativos (y a menudo muy conflictivos) no solamente dentro de la misma industria sino incluso en diferentes unidades de producción de una misma empresa.<sup>5</sup> Típicamente, es más probable que las grandes empresas con fuerzas de trabajo sindicalizadas implementen la solución cooperativa. Los mercados laborales en dichas firmas son denominados como el *mercado laboral primario*, caracterizado por empleos de largo plazo, carreras de promoción bien definidas, despidos infrecuentes por justa causa, y repartición de ganancias debido a la cooperación entre los trabajadores y los dueños. Otros sectores de la economía (caracterizados a menudo por empleos inseguros, ascensos o carreras de promoción cortas, y salarios bajos) implementan la solución no cooperativa y conforman el *mercado laboral secundario*. Estas diferencias se ilustran en el ejemplo de negociación que se presentó anteriormente mediante  $t(e^+) < t(e^*)$  y las ganancias compartidas provenientes de la cooperación  $\pi^+ > \pi^*$  y  $v^+ > v^*$ . Obsérvese que, de acuerdo con esta interpretación, una forma pura del modelo de disciplina laboral aplica para el mercado laboral secundario, mientras que un modelo híbrido –de mejora negociada a partir de resultados ineficientes propios de la solución no cooperativa– describe mejor el mercado laboral primario.<sup>6</sup>

¿Cómo podría la estructura de una economía afectar la viabilidad del resultado cooperativo? Obsérvese que la probabilidad de terminación tiene el mismo efecto sobre la viabilidad del resultado cooperativo como lo tiene la tasa de preferencia por el tiempo. Si la probabilidad de ser despedido es alta ( $t$  es grande) será difícil sostener el resultado cooperativo como un equilibrio. Consideremos una función de terminación más realista en la cual la probabilidad de pérdida del trabajo es la suma de la probabilidad de terminación por una justa causa,  $t(e)$ , y la probabilidad  $\tau$  de terminación por otras razones (fluctuaciones en la demanda o reubicación de la empresa, por ejemplo). El modificar la

<sup>5</sup> El mismo modelo es ampliamente aplicable a otros tipos de situaciones de colaboración, tales como los matrimonios (Lundberg y Pollak 1993), en las cuales también se observan resultados cooperativos y no cooperativos.

<sup>6</sup> Una interpretación alternativa que ofrecen Bulow y Summers (1986) sostiene que los altos salarios del mercado laboral primario se explican por el hecho de que el modelo del “salario de eficiencia” aplica aquí pero no en el mercado laboral secundario que se caracteriza por vaciarse con salarios bajos. No obstante, su interpretación parece dudosa dados los altos niveles de desempleo involuntario entre grupos demográficos que típicamente buscan trabajo en el mercado laboral secundario.

tabla de pagos para tener en cuenta las terminaciones que no son por justa causa requeriría agregar una  $\tau$  a  $t(e)$  dondequiera que ésta última aparezca en la tabla 8.1. Queda entonces claro que altos niveles de terminaciones que no son por una justa causa hacen que sea más difícil de sostener el resultado cooperativo. Por lo tanto, es más probable que mejoras en el sentido de Pareto basadas en negociaciones con respecto a los niveles esfuerzo-salario se puedan sostener ahí donde coexisten dos instituciones: los sindicatos de trabajo con la capacidad de negociar con los empleados e implementar el monitoreo entre los compañeros, por un lado, y políticas macroeconómicas que suavicen la volatilidad de las fluctuaciones de la demanda agregada, por el otro. Esto es un ejemplo de lo que se denomina *complementariedad institucional*, una situación en la cual los efectos benéficos de una institución aumentan en presencia de la otra. (En el capítulo final me referiré nuevamente a las complementariedades institucionales). En lugares donde las complementariedades institucionales son fuertes esperaríamos ver la coexistencia tanto de negociaciones colectivas efectivas como de una efectiva estabilización macroeconómica o la ausencia de ambos.

### ¿POR QUÉ LAS EMPRESAS NO VENDEN TRABAJOS?

Es probable que existiera una forma más simple de lograr no solamente una *mejora* en el sentido de Pareto con respecto al resultado no cooperativo sino también de implementar un resultado Pareto *eficiente*. El resultado clave en la demostración anterior de la ineficiencia en el sentido de Pareto es que la restricción de participación del trabajador no es activa (no se tiene con igualdad) y, por tal motivo, el trabajador recibe lo que yo he denominado una renta de ejecución. ¿Pero, por qué debería ser así? ¿Acaso el empleador pasó por alto la oportunidad de aumentar sus beneficios? ¿No podría una empresa, la cual observa que el trabajador recibe un aumento substancial en el valor presente de su utilidad vitalicia esperada cuando tiene un trabajo asegurado, simplemente cobrar una tasa por garantizarle el trabajo (Carmichael 1985)? Si la empresa explotara esta oportunidad, el trabajador podría entonces pagarle a la empresa una suma inicial de  $w^* - \tilde{w}^*$  y, por tanto, habiendo pagado la tasa por el trabajo, sería simplemente indiferente si toma el trabajo pero, más importante aún, *no* sería indiferente ante la posibilidad de perderlo. Consideremos cómo puede funcionar esto.

La tasa por trabajo a ser considerada es una transferencia no retornable que se realiza una sola vez, la cual es requerida por el empleador como condición para garantizar el empleo (esto a veces se denomina erróneamente una fianza). Asumamos que la riqueza total del trabajador es  $v + K$  donde, igual que antes,  $v$  es el valor del trabajo y  $K$  es “la riqueza adicional”. Dado que el trabajador financia la tasa por trabajo con recursos provenientes de  $K$ , la tasa simplemente reduce la riqueza del trabajador. Asumo que esto no tiene efecto marginal alguno sobre el comportamiento del trabajador. Por esta razón, la función de mejor respuesta del trabajador  $e(w)$  no se ve afectada. Asumiré que el trabajador se ve efectivamente restringido (por consideraciones de procedimiento o reputación) de ser oportunista y adoptar una función de despido para aprovecharse de la tasa que recibe y aumentar el nivel de las nuevas contrataciones. El costo del trabajador por reducir sus otras riquezas en un dólar es equivalente a la reducción en un dólar de la riqueza  $v$ . Ya que  $h$  es el número de trabajadores contratados, también es el número de tasas de trabajo recibidos por la empresa. Para simplificar la exposición, me he abstraído completamente del monitoreo (el nivel de esfuerzo  $e$  es conocido por el empleador pero no es verificable).

El empleador varía  $h$ ,  $w$ , y  $B$  para maximizar

$$\pi = y(he(w)) - hw + iBh \quad (8.14)$$

sujeto a

$$v(e(w), w - iB) \geq \bar{z}$$

donde  $i$  es la tasa de retorno,  $B$  es la magnitud de la tasa, y  $v(\cdot)$  es el valor presente ex-ante del trabajo con una tasa  $B$ . El término  $w - iB$  es el salario neto, teniendo en cuenta el costo de oportunidad que representa para el empleado el privarse de los retornos que obtendría por  $iB$  sobre su riqueza.

El Lagrangiano asociado al problema de optimización está dado por

$$r = y(he(w)) - hw + iBh + \gamma \{v(e(w), w - iB) - \bar{z}\}$$

con las condiciones de primer orden:

$$r_w = y'be' - b + \gamma(v_w + v_e e') = 0 \quad (8.15a)$$

$$r_b = y'e - w + iB = 0 \quad (8.15b)$$

$$r_B = ib - i\gamma v_w = 0 \quad (8.15c)$$

$$r_\gamma = v - \bar{z} = 0 \quad (8.15d)$$

De (8.15b) podemos determinar el nivel de empleo como aquél que iguala el producto marginal del esfuerzo  $y'$  con el costo de una hora de trabajo  $(w - iB)$  por unidad de esfuerzo realizado por hora, o el costo de una unidad de esfuerzo, o

$$y' = \frac{w - iB}{e} \quad (8.16)$$

De la función Lagrangiana, podemos interpretar fácilmente a  $\gamma$  como el precio sombra de la restricción de participación  $y$ , esto es, dado (8.15c),

$$\gamma = -\frac{dr}{dz} = \frac{h}{v_w} \quad (8.17)$$

La ecuación (8.17) da el efecto sobre los beneficios de un cambio en la posición de reserva del trabajador, a saber, el aumento de salario necesario para satisfacer la restricción de participación del trabajador  $(1/v_w)$  veces el nivel de empleo. También podemos ver que para niveles positivos de empleo y de no saciedad  $(v_w > 0)$   $\gamma > 0$ , por lo que la restricción de participación es activa.

Eliminando  $b$  de (8.15a) y sustituyendo en el valor de  $\gamma$  tenemos que

$$y'e' - 1 = -1 - \frac{e'v_e}{v_w}$$

lo cual, reorganizando términos, da

$$y' = -\frac{v_e}{v_w} \quad (8.18)$$

Combinando las ecuaciones (8.16) y (8.18) tenemos que



$$\frac{w-ib}{e} = -\frac{v_e}{v_w} \quad (8.19)$$

lo cual requiere que el costo de una unidad de esfuerzo para la empresa (el lado izquierdo de 8.19) sea igual al (negativo de) la tasa marginal de sustitución entre salarios y esfuerzo en el locus de iso-valores-presentes del trabajador (el lado derecho).

El problema y su solución se pueden interpretar de la siguiente manera. En la figura 8.5, permitamos que el eje horizontal represente el salario recibido por el trabajador,  $w$ , así como el costo salarial en el que incurre el empleador,  $w - iB$ . El empleador identifica el locus iso- $v$  para el cual  $v = \bar{v}$  (porque sabe que la restricción de participación estará activa). A lo largo de este locus existe algún punto (**a**) que maximiza la pendiente de un rayo desde el origen o  $e/(w - iB)$  con lo cual se satisface la ecuación (8.19). Este punto, que está fuera de la función de mejor respuesta del trabajador, naturalmente no se puede obtener directamente: el ofrecer el salario  $w^* - iB^*$  no motivaría al trabajador para que suministre el nivel de esfuerzo  $e^*$ . La tasa de salario se determina entonces como aquella que es suficiente para motivar al trabajador para que suministre el nivel de esfuerzo  $e^*$ . Y el nivel de la tasa de empleo es aquel que implementa el punto **a**, satisfaciendo la restricción de participación con igualdad. Habiendo determinado  $w$ ,  $B$ , y  $e$ , se determina  $b^*$ , mediante la ecuación (8.15b). Posteriormente, el empleador verifica para si en el equilibrio podría ser de mayor beneficio el no contratar trabajo y /o si el trabajador está satisfecho. Si ninguno de ellos es el caso, se mantiene el supuesto de que la restricción de participación es activa y se ha identificado correctamente el máximo beneficio.

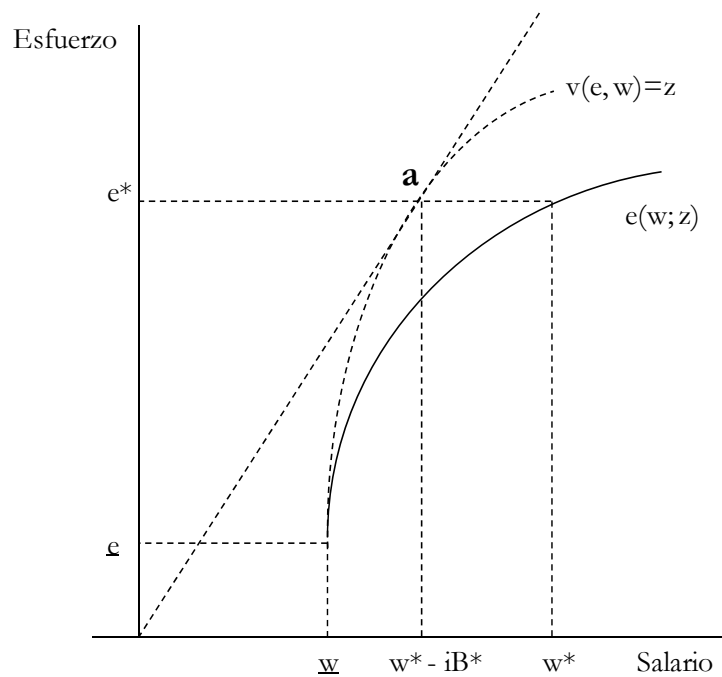


FIGURA 8.5 Las tasas óptimas que se pagan por el trabajo vacían el mercado laboral e implementan un óptimo de Pareto. El empleador identifica el punto **a** como la solución que maximiza  $e/w$ , el esfuerzo extraído del empleado por unidad de costo. Para implementar este resultado, el empleador ofrece el salario de  $w^*$  (a lo cual el empleado responde con  $e^*$ ) con una tasa por trabajo de  $B^*$ .

El equilibrio con las tasas de trabajo óptimas es abismalmente diferente al del caso anterior. Primero, dado que la restricción de participación es activa, el trabajador es indiferente entre tomar o no el trabajo. El mercado laboral por lo tanto se vacía; no hay trabajadores que hayan sido empleados involuntariamente. Este resultado resalta una importante limitación de buscar vaciar el mercado laboral como objetivo de política: si los trabajos son suficientemente poco atractivos puede que no haya un exceso en demanda. Segundo, mientras que las rentas ex-ante (antes de aceptar el trabajo) son cero, las rentas ex-post son en realidad mayores a aquéllas del caso en el que no se paga la tasa de empleo (para un  $z$  dado, el salario óptimo es mayor, ya que se establece no solamente para motivar el esfuerzo sino también para incrementar el valor de la tasa que puede ser extraída del potencial trabajador). Tercero, la riqueza del trabajador se reduce y aumentan los beneficios del empleador. En equilibrio general, el efecto sería aumentar el número de empresas y el nivel de empleo (hasta que  $z$  haya subido lo suficiente para restaurar la condición de cero beneficios).

Este modelo parece predecir que los tasas por trabajo llegarían a ser comunes (y substanciales) y que, como resultado, el desempleo involuntario sería algo raro. Sin embargo, son pocas las empresas que les cobran una suma inicial por trabajar a sus trabajadores y, mientras existan substitutos para estos tasas de trabajo, tal como un período inicial de prueba con un salario bajo, serán realmente raros y de una magnitud remotamente cercana a la que haría que el trabajador fuera indiferente entre aceptar o no un trabajo.<sup>7</sup>

¿Significa esto que los empleadores simplemente no toman ventaja de una oportunidad que les pueda rendir beneficios? Es algo similar a encontrar en la calle un billete de 500 dólares, no es que nunca pase algo así, es solo que parece poco probable. ¿Por qué entonces no sucede? Decir que los potenciales empleados no tienen mucho dinero para pagar una tasa por el trabajo no es una respuesta convincente, ya que la riqueza limitada de un trabajador simplemente restringe las sumas que se podrían cobrar, pero no refuta la lógica del argumento, el cual seguiría prediciendo tasas óptimas y por ello vaciando los mercados laborales. Una explicación más convincente es que motivaciones positivas del empleado hacia el empleador son importantes para inducir un trabajo de alta calidad y altos niveles de esfuerzo, y el empleador tiene en cuenta estas motivaciones positivas a la hora de manejar la negociación más difícil posible. Hay dos tipos de evidencia que respaldan esta interpretación.

La primera, que los empleadores son reacios a recortar los salarios durante períodos de alto desempleo, aparentemente por razones de moral y motivación de los empleados. Truman Bewley (1999) trató de entender la reticencia de los empleadores a recortar los salarios en épocas de recesión y así tomar ventaja de la posición de reserva más baja de los empleados que implica dichos momentos. Sus extensas entrevistas mostraron que los empleadores le temen al efecto que las reducciones de salario tienen sobre la moral de los empleados. La razón por la cual los empleadores se abstienen de recortar los salarios durante una recesión se puede aplicar con igual o mayor fuerza al

---

<sup>7</sup> En los Estados Unidos, los trabajos que pagan rentas altas típicamente las ofrecen a los principiantes así como a los trabajadores titulares (experimentados), lo que pone en duda la interpretación de los salarios iniciales más bajos como tasas implícitas de trabajo. Las tasas implícitas de trabajo adoptan formas no monetarias, como cuando un empleador asegura un respaldo político de un potencial trabajador o exige favores sexuales.

hecho de que la mayoría no cobra nada por los trabajos, incluso si las rentas asociadas a los trabajos son substanciales.

Segundo, como vimos en el capítulo anterior, experimentos de mercados laborales muestran que las personas típicamente exhiben preferencias fuertemente recíprocas, y proveen altos niveles de esfuerzo en respuesta a las ofertas del “empleador” que se consideran son generosas (Fehr, Kirchsteiger, y Riedl 1998). Donde las tasas por trabajo estaban entre las posibles estrategias de los empleadores,, los sujetos experimentales las abandonaron ya que los beneficios caían como resultado de la respuesta negativa de los “empleados”.

Otra de las razones por las cuales las tasas de trabajo no son comunes podría ser porque el empleado potencial no confía en que los efectos de reputación o de debido proceso puedan ser suficientes para evitar que los empleadores despidan a los trabajadores sin justa causa para aumentar el número de tasas por trabajo cobradas.

El hecho de que el pago de tasas por trabajo no sean común se toma a veces como un indicativo de que el modelo de disciplina laboral mencionado arriba es de poca relevancia empírica, pero la evidencia mostrada anteriormente sugiere que, cuando el modelo se amplía para incluir los tipos de preferencias sociales descritos en el capítulo 3, el modelo *es* consistente con el hecho de que cobrar tasas por trabajo raro. La interpretación más plausible es que una oferta de trabajo que genere una renta de trabajo ex post puede ser considerada por un empleado como una señal de generosidad del empleador (si no viene acompañada del cobro de la tasa por trabajo) o simplemente como una estrategia para maximizar beneficios (si viene acompañada del cobro de la tasa por trabajo). Por lo tanto, cobrar una tasa por trabajar afecta las creencias del empleado en cuanto a las intenciones (o tipo) del empleador y el empleado responde de acuerdo a ello. Esta interpretación resalta la importancia que tienen las consideraciones de equidad y moral al establecer un salario, y motiva la designación de un empleo como una especie de intercambio de regalos (Akerlof 1982).

## DISCIPLINA LABORAL E INCENTIVOS: EVIDENCIA

Ya que existen muchos modelos plausibles de mercados laborales que han modificado los supuestos convencionales del modelo Walrasiano, debo explicar porqué me concentré en una clase particular de modelos. La razón principal es que el modelo de disciplina laboral basado en la renovación contingente es consistente con un número de hechos no controversiales sobre el funcionamiento de la economía (mientras que el modelo convencional no lo es).

El primero de ellos es la existencia de rentas substanciales de empleo en la mayoría de los trabajos. Uno no puede capturar los aspectos relacionados con los ingresos de estas rentas al comparar los niveles de ingreso de aquellos con y sin trabajo, ya que estos grupos difieren entre sí en muchos aspectos adicionales al simple status de empleo. La medida que teóricamente es más apropiada es la típica pérdida de los beneficios y ganancias del empleado por causa de una racha de desempleo exógenamente inducida, como la que pudiera existir debido al cierre de una planta. Henry Farber (2003:2), quien resumió la considerable literatura que existe sobre este tema, concluyó que los “trabajadores despedidos sufren períodos substanciales de desempleo y las ganancias en los empleos que tienen luego de haber sido despedidos son substancialmente menores que las ganancias que obtenían antes de haber sido despedidos.” Este recuento siguió siendo cierto incluso durante el período de fuerte demanda laboral de los años 1990s. El costo de la pérdida de trabajo en los Estados Unidos (una estimación del valor actual  $(v - z)$  utilizando un índice de preferencia de tiempo del 10 por ciento) puede estar entre la mitad y una-y-medio veces los ingresos anuales, dependiendo de la duración en el cargo.<sup>8</sup>

En términos subjetivos, las rentas pueden ser considerablemente mayores debido al estigma social y otras incomodidades que se asocian con estar sin trabajo. Un número de estudios han documentado la pérdida de bienestar subjetivo (medido a través de una serie de encuestas) que se asocia con la pérdida de un trabajo o el estar sin trabajo. Uno de los estudios (Winkelmann y Winkelmann 1998), utilizando datos de panel que permite hacer comparaciones de la misma persona en diferentes condiciones de empleo,

---

<sup>8</sup> Esta aproximación se basa en cálculos donde se utilizan las estimaciones de Farber. Ver también Burda y Mertens (2001).

descubrió que el efecto subjetivo de estar sin trabajo en sí era mucho mayor que el costo subjetivo de la pérdida de ingresos asociada.<sup>9</sup> Puede haber muchos trabajos en los cuales los empleados son indiferentes entre continuar empleados o ser despedidos, tal y como lo predice el modelo Walrasiano, pero la evidencia que indica que la mayoría de los empleados tienen una fuerte preferencia a seguir estando empleados es abrumadora.

Segundo, los salarios reales tienden a variar con el nivel de empleo (Bowles 1991, Blanchflower y Oswald 1994), como lo predice el modelo de regulación de esfuerzo. (En el modelo Walrasiano convencional, para un capital dado, el empleo adicional debe *disminuir* el producto marginal del trabajo y, en consecuencia, el salario).

Tercero, los empleadores dedican substanciales recursos en personal así como en otros recursos para monitorear los niveles de esfuerzo de sus empleados, gastos que serían inconsistentes con la maximización de beneficios si los niveles de esfuerzo fuesen invariantes o estuviesen sujetos a un contrato que puede ser cumplido sin costo alguno.

Cuarto, los niveles de esfuerzo en el trabajo parecieran ser bastante variables y raramente sujetos a contrato. Mientras que el pago basado en grupos es una práctica común, el pago por pieza individual es extremadamente raro fuera de la industria del vestido y del calzado (Petersen 1992). Más aún, cuando el sistema de pago para los instaladores de vidrios de automóviles en los Estados Unidos cambió de salarios por hora a tarifas por pieza, el rendimiento por trabajador aumentó en un quinto (Lazear 1996). De manera similar, los plantadores de árboles en Columbia Británica, que fueron asignados aleatoriamente para recibir una compensación por pieza, hicieron un trabajo 20 por ciento mejor que aquéllos a quienes se les había asignado aleatoriamente un salario fijo (Shearer 2001). Estas mayores respuestas ante mejoras en los incentivos sugieren que los empleados ejercen un control substancial sobre su esfuerzo de trabajo. En un estudio de granjeros quienes trabajaban bajo una variedad de contratos, se

---

<sup>9</sup> Blanchflower y Oswald (1994) estimaron en 60.000 dólares el ingreso adicional que se necesitaba para compensar el estar sin trabajo, pero dado que esto se basa en una comparación entre aquellos con y sin trabajo, condicional en un gran número de medidas demográficas y otras medidas, se puede estar sobreestimando el costo subjetivo de la pérdida de trabajo (el estar sin trabajo puede ser una de las muchas razones por las cuales los desempleados no son felices y es más probable que las personas hereditariamente infelices sean personas que no tienen trabajo).

hallaron incluso efectos mayores de incentivos por esfuerzo sobre la productividad. Laffont y Matoussi (1995) encontraron que la productividad de los Tunecinos que trabajaban como mano de obra contratada era la mitad de la productividad de aquéllos quienes trabajaban y eran los propios demandantes residuales (trabajo en familia). Más aún, las personas eran 50 por ciento más productivas cuando trabajaban bajo contratos de renta fija (y, en consecuencia, eran los únicos demandantes residuales sobre los resultados de su esfuerzo) que cuando trabajaban bajo los incentivos disgregados del esfuerzo cuando se trabaja en *aparcería*. En un estudio inusual, por el uso de datos longitudinales, en Filipinas se observó que:

[L]os trabajadores evidentemente suministran un mayor esfuerzo bajo un esquema de tarifa de pago por pieza o al trabajar en su propia granja en comparación con el empleo de salario por tiempo, como se refleja en el hecho de que ellos gastan aproximadamente un 10% más de su masa corporal, neto de consumo de calorías, cuando trabajan bajo el esquema de la tarifa de pago por pieza que en sus propias tierras. . . . [E]l mismo trabajador consume 23% (16%) por ciento más calorías por día cuando se encuentra empleado bajo un esquema de tarifa de pago por pieza (empleo en su propia granja) que cuando trabaja a cambio de un salario por tiempo. (Foster y Rosenzweig 1994:214).

Un estudio realizado en el Reino Unido, con observaciones individuales de las actividades de los empleados, reveló que el esfuerzo de trabajo respondía fuertemente a las condiciones macroeconómicas, tal y como se podría esperar con base en el modelo de regulación de esfuerzo, mayores niveles de desempleo motivaron una mayor intensidad en el trabajo (Schor 1988). Un estudio en los Estados Unidos utilizando series de tiempo encontró que la productividad laboral variaba fuertemente con el monto de las rentas de empleo, condicional en los movimientos en la relación capital-trabajo, el nivel de utilización de capacidad, y otras variables estándar en trabajos econométricos sobre la productividad (Bowles, Gordon y Weisskopf 1983). Otras evidencias sugieren que estos efectos del modelo de disciplina laboral son más fuertes en el mercado laboral secundario que en el mercado primario sindicalizado, y más fuertes aún en países donde los sindicatos de trabajo son más débiles.

La baja frecuencia de despidos por justa causa no es evidencia en contra de los modelos de disciplina laboral, pues una estrategia efectiva de disciplina podría resultar en la ausencia de despidos (como en Shapiro y Stiglitz 1984). Es más, incluso si la terminación por justa causa no hace parte explícitamente de la estrategia de disciplina

laboral del empleador – como sucede en los mercados de trabajo primarios de muchas economías de Europa – las evaluaciones que el empleador haga de los esfuerzos del trabajador se utilizan ampliamente para la elección de promociones o retiros, lo que reproduce efectivamente los efectos de terminación por justa causa como un incentivo al trabajo duro.

Algunos de estos hechos se pueden explicar por otros modelos de empleo como los de renta compartida o transacciones de activos específicos que se introducirán en el capítulo 10. Sospecho que un entendimiento adecuado de los mercados de trabajo y las relaciones de empleo podrá requerir de enfoques híbridos incluyendo otros modelos no Walrasianos que no se desarrollan aquí. Muchos de los hechos que se presentan más adelante son consistentes con más de uno de estos modelos.

## CONCLUSIÓN

El estímulo para gran parte del nuevo trabajo teórico sobre los mercados de trabajo vino de la insatisfacción con los aspectos microeconómicos de modelos macroeconómicos de empleo y desempleo agregado. Los macroeconomistas sobresalieron entre los primeros innovadores. Los modelos basados en contratos incompletos de esfuerzo u otros aspectos del intercambio laboral explicaron cómo un equilibrio competitivo podría exhibir desempleo involuntario, acortando así el vacío entre la teoría estándar y la observación empírica.

Durante el proceso, las teorías estándar del mercado laboral y de la empresa se transformaron substancialmente. Robert Solow (1990) resume la dirección del cambio en el título de su libro *El Mercado de Trabajo como Institución Social*, y Arthur Okun (1981) capturó el nuevo rol clave de la confianza y otras preferencias sociales al acuñar el término “el invisible apretón de manos”. La importancia de los motivos de reciprocidad y otras preferencias sociales para explicar por qué las empresas no venden los trabajos resalta la futilidad de simplemente introducir contratos incompletos como única modificación en un modelo Walrasiano. Los experimentos resumidos en la penúltima



sección del capítulo anterior sugieren que la no completitud contractual aumenta el rol de las preferencias sociales en la determinación de los resultados de equilibrio.

Son tres las implicaciones de las nuevas teorías que han recibido menos atención. La primera ya se había mencionado en la discusión de la elección que hace la empresa de los insumos de capital cuando el esfuerzo de trabajo no puede ser contratado. Si la dificultad para monitorear el esfuerzo en el trabajo difiere a lo largo de las tecnologías, la elección de la tecnología se verá influenciada por la naturaleza del problema de disciplina laboral. Por lo tanto, aspectos del ambiente de la disciplina laboral tales como las normas existentes, si los trabajadores despedidos tienen o no acceso a un seguro de desempleo, y otras influencias sobre la elección de esfuerzo por parte del trabajador, afectarán la rentabilidad de las tecnologías alternativas. Esta visión contrasta con el modelo estándar en el cual la elección de tecnología responde a la escasez de factores reflejada en los precios de los mismos. Esto también levanta dudas acerca de si la existencia de instituciones – la empresa convencional, por ejemplo – puede ser explicada por los requerimientos de tecnologías exógenamente dadas. Una visión más plausible es que las tecnologías y las instituciones evolucionan conjuntamente, y que cada una influye en el desarrollo de la otra.

He aquí un ejemplo. Cuando las empresas de camiones en los Estados Unidos instalaron computadores en las cabinas durante los años 1980s, éstas mejoraron ampliamente su capacidad para monitorear las acciones de los conductores (Baker y Hubbard 2000). Los dispositivos para registrar los viajes proveyeron a la compañía con información verificable sobre la velocidad, tiempo ocioso, y demás detalles sobre la operación del camión sobre los cuales había conflicto de intereses entre el conductor y la compañía. Por ejemplo, el costo de operar los camiones (que pagaba la compañía) era creciente y convexo en la velocidad del camión; los conductores preferían conducir a una velocidad mayor a la que minimizaba los costos y así tomarse descansos más largos. Los conductores que eran dueños de sus camiones eran demandantes residuales sobre sus ingresos, menos los costos, y en consecuencia internalizaban los costos de combustible y depreciación, lo cual resultó en ahorros significativos. Por tal razón, antes de que se introdujeran los dispositivos para registrar los viajes, los dueños-operadores tuvieron éxito al competir con las flotillas de las compañías en aquellos trayectos para los cuales

los conflictos de interés entre los conductores y las compañías eran particularmente fuertes.

Cuando se utilizaron los dispositivos para registrar los viajes, las compañías pudieron escribir contratos basados en la velocidad a la cual se manejaba el camión y proveer a los conductores con otros incentivos para actuar de acuerdo con los intereses de la compañía. A diferencia de otros computadores instalados en el vehículo (sistemas electrónicos de gestión de vehículos, o EVMSs por sus siglas en inglés), los dispositivos para registrar los viajes no proveyeron una mejoría en la coordinación entre los camioneros y los despachadores, ya que la información quedaba a disposición de la compañía solamente cuando el trayecto era completado. La única función de los dispositivos para registrar los viajes era la de mejorar la contractibilidad de aspectos referentes a los comportamientos de los conductores sobre los cuales había un conflicto de interés entre los conductores y la compañía. Al mejorar las oportunidades contractuales de las compañías, los dispositivos para registrar los viajes tuvieron dos efectos. Primero, conllevaron a una disminución significativa en la participación de mercado de los dueños-operadores. Segundo, los conductores en los camiones equipados con dispositivos para registrar los viajes manejaban más despacio. En contraste, la capacidad de los EVMSs para mejorar la coordinación entre los conductores y los despachadores bajó los costos pero no tuvo efecto especial alguno sobre la distribución de contratos o de propiedad en la industria.

En este caso, se escogió una tecnología porque aumentaba el conjunto de contratos factibles de manera tal que mejoraban los beneficios. Si las tecnologías son endógenas en este sentido, se hace difícil dar una definición precisa al término *costos de transacción*. En el modelo que se desarrolló arriba queda claro que los costos de monitoreo son costos de transacción. Sin embargo, la ecuación (8.7d) muestra que la predisposición a pagar por parte de la empresa para utilizar el insumo  $k$  se explica por la contribución que este insumo hace tanto a la producción como a la disciplina laboral. ¿Los costos de utilizar el bien  $k$  son costos de transacción? De ser así, ¿qué fracción del costo del insumo  $k$  debe ser asignada a costos de transacción y cuál a los costos de producción? Los costos de los dispositivos para registrar los viajes instalados en los camiones fueron casi totalmente costos de transacción. Pero ¿qué hay de los demás

EVMSs los cuales, al igual que los dispositivos para registrar los viajes, permitieron contratos con incentivos más fuertes para los conductores y también mejoraron en gran medida la coordinación? La misma ambigüedad surge con respecto al salario. Hemos visto que un aumento de salario acompañado por una disminución en el monitoreo podría sustentar el mismo nivel de esfuerzo laboral. Parecería extraño llamar la disminución en el monitoreo una disminución en los costos de transacción dado que el costo total de contratación de trabajo ha aumentado. En consecuencia, ¿son los salarios también costos de transacción? Estas ambigüedades en cuanto al significado del término parecen ineludibles y explican por qué hago poco uso aquí del esquema de costos de transacción.

Una segunda implicación de los nuevos modelos de mercado de trabajo es que, debido a que la relación de empleo persiste a lo largo de muchos años, el lugar de trabajo es un ambiente cultural donde evolucionan las preferencias y las creencias de los empleados. A este respecto, los lugares de trabajo no se diferencian de las escuelas o de los vecindarios, pues ejercen influencia acerca de quién conoce a quién, para hacer qué, y cuáles recompensas se asocian con cuáles comportamientos. Un ejemplo empírico sugerirá la importancia de estos efectos. A lo largo de un período de tres décadas, Melvin Kohn y sus colaboradores estudiaron la relación entre el cargo que tenemos en la estructura de autoridad de nuestro lugar de trabajo – dado como opuesto a tomar órdenes – y la valoración individual de auto orientación e independencia en sus hijos, así como la flexibilidad intelectual propia y la auto supervisión personal. Ellos llegaron a la conclusión que “la experiencia de la autodirección ocupacional tiene un efecto profundo sobre los valores, la orientación y el funcionamiento cognitivo de las personas.”<sup>10</sup> Su estudio en colaboración con Japón, los Estados Unidos y Polonia (Kohn, Naoi, Schoenbach, Schooler y Slomczynski 1990) rindieron hallazgos consistentes a lo largo de diferentes culturas: las personas que ejercen la auto orientación en el trabajo también valoran más la auto orientación en otras extensiones de su vida (incluyendo la crianza de los hijos y actividades de ocio) y son menos probables de exhibir fatalismo, desconfianza y auto súplica. Kohn y sus coautores (1983:142) razonan que “la estructura social afecta el funcionamiento psicológico individual principalmente al afectar las condiciones de las

---

<sup>10</sup> Ver Kohn (1969), Kohn, Naoi, Schoenbach, Schooler y Slomczynski (1990), Kohn y Schooler (1983), y Kohn (1990). La cita es de la página 967 del trabajo escrito en coautoría en 1990. Los estudios tienen en cuenta la posibilidad de que la personalidad esté afectando la estructura de trabajo y no al contrario.

propias vidas de las personas.” Kohn concluye, “La explicación simple que cuenta para virtualmente todo lo que sabemos sobre los efectos del trabajo en la personalidad. . . es que los procesos son directos: aprender del trabajo y expandir esas lecciones a las realidades fuera del trabajo” (Kohn 1990:59).

Ya que las dimensiones de personalidad mencionadas por Kohn son parte de las preferencias de los individuos que explican cómo ellos crían a sus hijos, en cuál tipo de actividades de ocio participan, y similares, esto constituye una evidencia fuerte de que las preferencias son endógenas en relación con la organización del lugar de trabajo.

Una tercera implicación es que las normas de salarios justos, ética de trabajo y demás preferencias sociales no son exógenas sino más bien evolucionan bajo la influencia de los salarios, esfuerzo en el trabajo y condiciones laborales actuales así como otras influencias externas al lugar de trabajo. Una discrepancia substancial entre la norma de salarios y el salario de equilibrio, por ejemplo, puede resultar en el desgaste de la norma o en una acción colectiva exitosa por parte de los empleados para mejorar su situación.

No sabemos qué es lo que Henry Ford tenía en mente cuando anunció la jornada de trabajo a cinco dólares. El hecho de que el rendimiento por hora del trabajador haya aumentado en más del doble con posterioridad al aumento de salario sugiere que el esfuerzo de los trabajadores aumentó substancialmente. (Ford aumentó el nivel de supervisión junto con el salario, por lo que no quedó duda de que disminuyó la probabilidad de tolerar trabajadores que perdieran tiempo durante su trabajo.) No sabemos si el aumento en el esfuerzo de los trabajadores fue una respuesta a la zanahoria de la supuesta generosidad de Ford (con lo que disminuye la desutilidad del esfuerzo, por ejemplo, en el lado izquierdo de la ecuación (8.5)) o al garrote que implica una supervisión más cercana y un aumento en las rentas de empleo (que aumenta en el lado derecho de la ecuación (8.5)).<sup>11</sup>

---

<sup>11</sup> Raff (1988) cree que el aumento en el componente de supervisión no es consistente con el modelo de disciplina laboral, pero parece que él asume (improbablemente) que la supervisión y el salario son substitutos más que complementos en la estrategia de disciplina laboral, contrario al razonamiento que se presenta aquí.