A Game Theoretical View on Efficiency Wage Theories*

Dennis Wesselbaum†

The Kiel Institute for the World Economy and EABCN

February 8, 2010

Abstract

The efficiency wage theory developed by Akerlof (1982) assumes observability of effort and the ability of firm and worker to commit on their effort/wage decisions. We show that, from a game theoretical point of view, we have to understand the firm/worker relationship as a repeated *Prisoner's dilemma*. Therefore, cooperation is per se not a (subgame perfect) Nash equilibrium and hence the Akerlof (1982) theory is based upon an implicit assumption of cooperation, which can not be implemented w.l.o.g.. In addition, we find that this approach is a special case of the Shapiro and Stiglitz (1984) approach and hence unify the two approaches.

Keywords: Efficiency Wage, Prisoner's Dilemma, Repeated Game, Subgame Perfect Nash Equilibrium.
JEL classification: C72, C73, J41.

* I would like to thank seminar participants at the Kiel Institute for highly valuable comments.

† Contact: dennis.wesselbaum@ifw-kiel.de

1 Introduction

In order to explain equilibrium unemployment, nobel laureate George A. Akerlof developed the fair wage theory based on the sociological partial gift exchange approach (see Akerlof (1982)). This theory explains equilibrium unemployment by the fact that the wage is above the market clearing wage. The reason for this non-neoclassical phenomenon is the fact that firms tend to pay a higher wage in order to ensure that workers provide a desired amount of effort. In contrast to Akerlof's idea, Shapiro and Stiglitz (1984) suggested a different type of efficiency wage model. While in Akerlof's model effort is observable and commitable, Shapiro and Stiglitz assume that effort is not observable. Consistently, a worker has an incentive to shirk, viz. the worker might not provide effort. Under equilibrium full employment, this would urge the firm to pay a wage that is above the market clearing wage, in order to create some punishment for shirking. As a consequence, equilibrium full employment is not feasible. We infer that both theories work along the same dimension, namely increasing the wage above the market clearing wage. However, there are several challenges afflicted with Akerlof's idea in a dynamic context. Especially, the assumption of commitable effort is crucial.[1] Let us consider a dynamic version based upon the Akerlof approach.[2] After the firm and the worker have somehow matched, the firm maximizes its profits and determines the optimal wage/effort decision. The worker receives the wage and provides the desired amount of effort. Akerlof motivates this i.a. with the sociological gift exchange, i.e. workers have "sentiments" for the firm. However, the other side of the coin is that whenever there is sentiment, there is also the possibility of being discouraged by firm's decisions.

We find that the implicit assumption of Akerlof (1982) of cooperation being a (subgame perfect) Nash equilibrium of the corresponding game does not hold w.l.o.g.. Furthermore, we establish conditions for the (subgame) perfectness of the cooperation strategy for the entire time path of the game. In addition, we find that this approach is a special case of the Shapiro and Stiglitz (1984) approach and hence unify the two approaches and provide a game theoretical foundation of efficiency wages.

2 The Repeated Prisoner's Dilemma

We consider an infinitely repeated *Prisoner's dilemma*, i.e. the horizon of the game T is unknown to both players and they both expect the game to be played for a long period of time. Let us call our game B expressed by a simultaneous-move matrix. Furthermore, we assume that our game is time-independent and hence stationary. Let $\mathcal{P}$ be the set of players containing the finite set $\{1, 2, ..., n\}$. Our game starts in period $t = 0$ and is played every period. We assume that

[1] The observation of each worker is - in addition - hardly imaginable due to monitoring costs and negative psychological effects on the workers motivation.

[2] See e.g. Danthine and Donaldson (1990) or Danthine and Kurmann (2004).

every player has full information and that actions are revealed to all players before the next round. With this assumption, we enable players to condition their actions on the history of events up to point t. In order to avoid the problem of infiniteness of player's payoffs, we introduce discounting of future payoffs. We understand this discounting as a measure of time preference.[3] In the following, we will discuss the general mathematical background for later purpose. Let $U_{t,i}$ denote the utility function of player i over the outcomes of B in period t. In addition, let $\phi \in (0,1)$ be the time-independent discount rate. Consistently, if the payoffs are constant over B and t, we can write the stream of payoffs of player i as

$$\sum_{t=0}^{\infty} \phi^t U^{t,i}. \tag{1}$$

The average discounted value of the payoff stream is then given by

$$(1-\phi)\sum_{t=0}^{\infty} \phi^t \Gamma_i^t = (1-\phi^t)\hat{\Gamma}_i + \phi^t \check{\Gamma}_i, \tag{2}$$

where Γ_i^t is the constant payoff of player i for t periods and let $\hat{\Gamma}_i$ denote the constant payoff for the first t periods, while $\check{\Gamma}_i$ is the different payoff for the next t periods.[4] Furthermore, let $s_i^t = (s_i^0, s_i^1, ...)$ denote the history-dependent strategies. The strategy profile s is the n-tuple of indiviual strategies such that $s = (s_1, ..., s_n)$. Now, let us consider a repeated-game between players (i, j) with the two strategies C and N. A strategy for player i that ensures cooperation is: play C in the first period and in every consecutive period, iff player j always cooperated. However, play N, iff player j played N in the precedent period. Let player $i's$ repeated game strategy be $\tilde{s}_i = (\tilde{s}_i^0, \tilde{s}_i^1, ...)$. Consistently, in period t after history h^t[5]

$$\tilde{s}_i^t(h^t) = \begin{cases} C, & \text{iff } \ h^t = (C,C)^t, \\ N, & \text{otherwise.} \end{cases}$$

As an illustrative example, consider the following game presented in Table 1.

Table 1: A Prisoner's Dilemma Example

i,j	C	N
C	1,1	-1,2
N	2,-1	0,0

If both players stick to their cooperation strategy, the payoff computation is straightforward, resulting 1. Now, let player i deviate from C in period t. For

[3] We introduce discounting, because worker and firm can not be sure how long the game will continue.

[4] $(1-\phi)\sum_{t=0}^{\infty}\phi^t\Gamma_i^t = (1-\phi)\left(\sum_{t=0}^{j-1}\phi^t\Gamma_i^t + \sum_{t=j}^{\infty}\phi^t\Gamma_i^t\right) = (1-\phi)\left(\frac{\hat{\Gamma}_i(1-\phi^t)}{1-\phi} + \frac{\check{\Gamma}_i\phi^t}{1-\phi}\right) = (1-\phi^t)\hat{\Gamma}_i + \phi^t\check{\Gamma}_i$, see Ratliff (1997).

[5] We assume that for any profile a it holds $h^0 = (a)^0$, such that $h^t = (C,C)^t$.

the first t periods, she receives 1 and for period $\bar{t}$, she receives 2, i.e. the payoff from (C, N). In any consecutive period $t > \bar{t}$, both players choose N, i.e. to not cooperate, and receive 0.[6] Using equation (2) yields that,

$$(1-\phi)\sum_{t=0}^{\infty}\phi^t\Gamma_i^t = (1-\phi^t)\hat{\Gamma}_i + \phi^t\left((1-\phi)\check{\Gamma}_i + \phi\tilde{\Gamma}_i\right), \tag{3}$$

if player i receives $\check{\Gamma}_i$ only for period t. Consistently, player $i's$ payoff from deviating is given by $1 - \phi^t(2\phi - 1)$. Some algebra yields that this cheating strategy is not profitable, as long as $\phi \geq \frac{1}{2}$. We have shown that cooperation is a Nash equilibrium of the game, if the time preference parameter is above a certain endogenously determined threshold.
As a final step, we show that cooperation is a subgame perfect Nash equilibrium. For this purpose, consider a subgame that starts in period $\hat{t}$ with history $h^{\hat{t}}$. The restriction $\hat{s}$ to the subgame $h^{\hat{t}}$ defines the strategy in this subgame. The restriction to this subgame is given by

$$\hat{s}_i^t(h^{\hat{t}}) = \begin{cases} C, & \text{iff } h^{\hat{t}} = (C,C)^{\hat{t}}, \\ N, & \text{otherwise.} \end{cases}$$

Now, we can identify two classes of histories (i) both players have choosen to cooperate for the entire game, and (ii) at least one player cheated in at least one of the previous periods. Then, for class (i) subgames the restriction reduces to the game strategy derived in (3), because the history up to this subgame has to read as $h^t = h^{\hat{t}} = (C,C)^t = (C,C)^{\hat{t}}$. Since in (3) $\tilde{s}_i^t$ is a Nash equilibrium, the restriction $\hat{s}_i^t$ is a Nash equilibrium strategy profile in (i), iff the condition for the discount factor holds.
In the second class, and by assumption, because one player has choosen to non-cooperate, both players choose to non-cooperate in the ongoing subgame. It can be shown, that such a strategy is also a subgame perfect equilibrium and consistently, that for any subgame the restriction of $\tilde{s}_i^t$ is a Nash equilibrium for that subgame, iff $\phi \geq \frac{1}{2}$. Therefore, $\tilde{s}_i^t$ is also a subgame perfect Nash equilibrium of the repeated game.

3 The Efficiency Wage Pendant

We have shown that in a repeated *Prisoner's dilemma* cooperation is a (subgame perfect) Nash equilibrium, iff the discount rate is above a certain threshold. In the following, we have to show that the efficiency wage theory - from a game theoretic viewpoint - yields such a *Prisoner's dilemma*. Which variables do we have to consider? The payoffs contain the worker's effort $e > 0$, since the utility

[6] Here, we assume that the punishment for deviating is playing N for any consecutive period. One might assume different punishment strategies, but since we consider a firm/worker relationship cheating should result in separation and hence there should be no way back to rebuild the relationship.

function in Akerlof (1982) has the form $u(e_n, e, w, \epsilon)$, where ϵ represents the worker's taste and e_n are the norms of effort. Moreover, effort is a decision variable for the worker and disregarding effort would lead to a distortion of our results. In addition, we consider the wage $w > 0$ - also present in the worker's utility function - to be some additional payment in case of the achievement of predetermined goals.[7] For the firm, only output y is considered. We assume that $w > e$, i.e. that the effort - given in terms of its share of the wage - is always smaller than the wage. Similarly, we assume $y > w$, such that the production process generates profits. Using these variables yields the following game-matrix presented in Table 2.

Table 2: The Efficiency Wage Game

F/W	C	N
C	(y-w),w-e	-w,w
N	y,-e	0,0

If firm and worker cooperate, the firm receives the output and has to pay the "extra" wage, while the worker receives the wage and provides effort (which can not be used for leisure). If the firm cooperates, but the worker chooses to non-cooperate, the firm has to pay the wage (since it committed on paying the wage) but receives no output.[8] The same considerations hold vice versa for the case (N, C). If both players choose to non-cooperate, the firm receives no output and the worker can spend this effort for leisure, hence generating some amount of utility.

If we now apply the methodology introduced in the precedent section, we infer that this is indeed a *Prisoner's dilemma*, since the static (subgame perfect) Nash equilibrium is (N, N), while $0 < (y - w)$ and $0 < w - e$.

Therefore, we can set up two propositions, such that cooperation is a (subgame perfect) Nash equilibrium in the infinitely repeated game, i.e. that firm and worker choose to cooperate over the entire game.

Proposition 1

The firm will cooperate, iff

$$y \geq \frac{w}{\phi}. \tag{4}$$

Proof

See the Appendix.

Proposition 2

The worker will cooperate, iff

$$w \geq \frac{e}{\phi}. \tag{5}$$

[7]Those goals are written down in the contract between firm and worker and can be considered as performance pay (see Lemieux et al. (2007)).

[8]For the sake of simplicity, we assume an extreme case. However, this assumption leaves our qualitative results unaffected.

Proof
See the Appendix.
We have shown that the efficiency wage theory - from a game theoretic view - is in fact a *Prisoner's dilemma* and, consistently, we established conditions, such that cooperation is a (subgame perfect) Nash equilibrium.

4 Final Remarks

We have shown that from a game theoretic viewpoint, the Akerlof approach is based upon an implicit assumption of commitability of effort. However, since the game between firm and worker is a *Prisoner's dilemma*, cooperation is per se no (subgame perfect) Nash equilibrium. We develop conditions for which cooperation is in fact a Nash equilibrium and consistently show that the Akerlof (1982) efficiency wage theory is nested within the Shapiro and Stiglitz (1984) theory. Moreover, we can understand the former approach as a special case of the general game between firm and worker, i.e. the latter. With this game theoretic approach, we are able to unify these two ideas, often viewed as disparate.

Appendix

Proof of Proposition 1

The firm will cooperate, iff

$$y \geq \frac{w}{\phi}. \tag{6}$$

Consider the game presented in Table 2. If we apply (3) to this problem, we initially obtain

$$(1-\phi^t)(y-w) + \phi^t \left((1-\phi)y\right). \tag{7}$$

Consistently, the profit of cheating is given by

$$y - w + \phi^t(w - \phi y). \tag{8}$$

However, if cheating should not be profitable, the stream of profits has to be smaller than the profit from cooperation, i.e.

$$y - w + \phi^t(w - \phi y) \leq y - w. \tag{9}$$

Applying some algebra yields the condition

$$y \geq \frac{w}{\phi}, \tag{10}$$

q.e.d.

Proof of Proposition 2

The worker will cooperate, iff

$$w \geq \frac{e}{\phi}. \tag{11}$$

The problem for the worker is solved analogously to the firms problem. Therefore, the initial condition looks as follows

$$(1-\phi^t)(w-e) + \phi^t \left((1-\phi)w\right). \tag{12}$$

Some rearranging gives

$$w - e + \phi^t e - \phi^{t+1} w. \tag{13}$$

The condition for non-profitability of cheating is given by

$$w - e + \phi^t e - \phi^{t+1} w \leq w - e, \tag{14}$$

such that

$$w \geq \frac{e}{\phi}. \tag{15}$$

q.e.d.

References

Akerlof, George A.. 1982. "Labor Contracts as Partial Gift Exchange." *Quarterly Journals of Economics*, (97(4): 543-569.

Danthine, Jean-Pierre and Andre Kurmann. 2004. "Fair Wages in a New Keynesian Model of the Business Cycle." *Review of Economic Dynamics*, 7(1): 107-142.

Danthine, Jean-Pierre and John B. Donaldson. 1990. "Efficiency wages and the business cycle puzzle." *European Economic Review*, 34(7): 1275-1301.

Lemieux, Thomas, W. Bentley MacLeod, and Daniel Parent. 2007. "Performance Pay and Wage Inequality." NBER Working Paper, No. 13128.

Ratliff, Jim. 1997. Lecture Slides to the Graduate-Level Course on Game Theory. Available via the following link: http://www.virtualperfection.com/gametheory/.

Shapiro, Carl and Joseph E. Stiglitz. 1984. "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review*, 74(3): 433-444.