

Efficiency wages: Variants and implications

Wages affect productivity and non-wage costs; this carries important labor market and policy implications

Keywords: efficiency wages, selection wages, turnover, morale, discipline

ELEVATOR PITCH

Higher wages increase labor costs but also improve the productivity of the labor force in several ways. If firms take this into account and set their wages accordingly, the resulting wages could fail to adjust demand and supply but may induce phenomena like over-education, discrimination, regional wage differentials, and a tendency for larger firms to pay higher wages. All these phenomena are quantitatively important and well-established empirically. Efficiency wage theory provides an integrated theoretical explanation rather than a sundry list of reasons, and offers an efficiency argument for progressive income taxation.

KEY FINDINGS

Pros

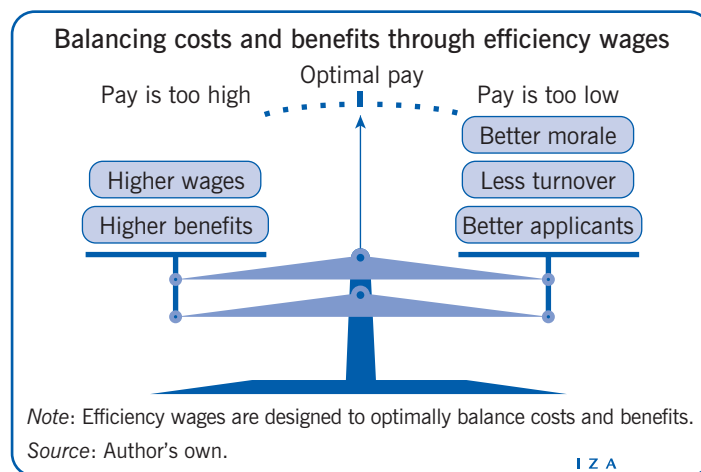
- + Efficiency wage theory can provide a unified explanation for some key labor market pay and employment tendencies.
- + According to efficiency wage theory, progressive income taxation can be used to reduce pre-tax wage inequality.
- + Firms could be discouraged from employing efficiency wages by taxation strategies and labor legislation.
- + Collective bargaining practices may help to curtail the use of efficiency wages.

Cons

- It is difficult to pinpoint selected efficiency wage effects empirically because they affect wage setting jointly with other influences.
- Efficiency wage theory may complement other theoretical approaches to understand how labor markets function, but such integration is not sufficiently developed.
- The relationship between efficiency wage setting, collective bargaining, and labor market institutions is insufficiently explored.
- Due to efficiency wage setting, wages in a free market do not always reflect the actual costs of producing a good or service, leading to distorted consumer choices and unwarranted differences in income.

AUTHOR'S MAIN MESSAGE

Labor markets are usually treated like commodity markets, where quality is well-defined and demand and supply are equated by price. However, wages, unlike prices, serve several further functions that affect profitability beyond attracting workers. With higher wages, more demanding hiring standards can be implemented and better workers hired, labor turnover and associated turnover costs can be reduced, and better worker attitudes can be kindled. The efficiency wage approach leads to a unified understanding of some important empirical phenomena and entails some unorthodox policy conclusions, such as the idea that progressive taxation enhances efficiency.



MOTIVATION

Efficiency wage theory focuses on aspects of firms' wage-setting practices that affect the efficiency of their workforce. In particular, wages may be raised beyond the level necessary to attract enough applicants to fill a firm's vacancies. Such higher pay would increase the efficiency of the firm's workforce through various channels. The most important channels are the following:

- *Selection.* By offering higher wages, more applicants will be available to choose from, and more promising applicants will apply. This permits the firm to implement more demanding hiring standards and end up with a more productive labor force.
- *Turnover.* Labor turnover is reduced when higher wages are offered. This brings down turnover costs (i.e. costs for advertising an opening, selecting among the applicants, and training the new hires).
- *Discipline.* The higher the wage, the more the workers will tend to avoid behavior that may trigger dismissal. This may improve work discipline.
- *Morale.* Workers who are paid well generally produce better quality work.

Such increases in efficiency may make it worthwhile for the firm to raise its wage offers above what is necessary to fill its vacancies. Conversely, lower wages may limit a firm's realistic scope for selecting workers according to expected productivity and lead to higher turnover, less discipline, and reduced morale. These negative effects may prevent a firm from lowering its wages, even if a sufficient number of applicants would have been available at a lower wage offer.

While standard demand–supply theory focuses on the idea that firms raise their wage offers if they need more applicants, or lower them if more applicants are available than needed, efficiency wage theory emphasizes the trade-offs between the costs and benefits involved in changing the wage level. The illustration on page 1 shows how the use of efficiency wages balances out costs (due to higher wages and fringe benefits) and benefits (resulting from better applicants, less turnover, and better morale). The market implications that may arise when firms set wages in this manner are wide-ranging and important. While profitable for the individual firms, such wage setting usually entails some inefficiencies in the functioning of the labor market. Efficiency wage theory suggests policies that reduce some of these inefficiencies.

DISCUSSION OF PROS AND CONS

The efficiency wage channels work jointly

The different channels through which the wage level (including fringe benefits) affects the productivity of a firm's labor force are not mutually exclusive; rather, they are used jointly and often reinforce each other. While the theoretical literature typically discusses the channels separately, it is, for practical purposes, useful to consider their joint overall effects. This is also appropriate because the different channels (such as selection, turnover, discipline, or morale) generate quite similar overall effects. This will be illustrated in the following with regard to the effect of progressive taxation on pre-tax wage inequality, the spillover effect of minimum wages, the size-wage effect, the congestion-wage effect,

and wage discrimination generated through the different channels. It will be shown that efficiency wages generate, overall, too much wage inequality in all these dimensions, and that progressive taxation may be used to counteract this tendency.

Wage inequality and progressive taxation

Consider the case that the taxation of labor income is made more progressive. A progressive taxation means that the marginal tax rate—the proportion of additional income earned that is subsequently collected as taxes—exceeds the average tax rate. An increase in the marginal tax rate (and possibly an increase in the tax-free allowance) renders it more costly for firms to make use of all the above mentioned channels, and thereby curb wage inequality:

- *Selection.* In order to obtain the same selection effect in the presence of a more progressive wage tax, firms must make higher offers to applicants, as a larger portion of any wage increase will be taxed away. The selection wage effect thus becomes less effective (and thereby more costly), and firms will make use of it to a lesser extent.
- *Turnover.* Likewise, the turnover wage effect will become less effective (and more costly), and firms will use it to a lesser extent.
- *Discipline.* In the same way, the discipline wage effect will become less potent (more costly), and firms will use it less.
- *Morale.* The effect on morale will be reduced in the same manner.

All in all, the pre-tax wages offered by firms are expected to *decline* if the wage tax or income tax becomes more progressive. This contrasts with the conclusion that emerges from standard demand–supply theory, which posits that an increase in progressivity reduces the benefits to the worker of acquiring additional qualification. Therefore, the supply of qualified workers is expected to decrease at a given wage level, while demand remains unchanged. Excess demand for these workers then rises (due to a reduced supply of qualified workers), and pre-tax wages increase in kind.

The conclusions that emerge from efficiency wage theory thus differ from those proposed by standard wage theory. Unfortunately, there are not many detailed empirical studies about such effects. One recent OECD study reports that the size of the tax wedge (the fraction of taxes and social contributions in gross wages) is correlated with greater equality in pre-tax earnings [1]. This corresponds with the observation that pre-tax wage inequality is very high in the US but quite low in the Scandinavian countries, which practice significantly more socio-economic redistribution, and is in line with predictions from efficiency wage theory. All in all, it is expected that the wages offered by firms *decline* if the progressivity of the wage tax or income tax is increased.

Efficiency wages and minimum wages

In standard demand–supply theory, the introduction of a minimum wage will eliminate some low-paying jobs and lift the wages of some other low-paying jobs to the minimum wage, but this will mostly leave wages above the minimum wage unaffected. In contrast,

and from the point of view of efficiency wage theory, an increase in the minimum wage will reduce the relative attractiveness of jobs that are paid slightly above the minimum wage and render jobs that were formerly paid below the minimum relatively more attractive. This makes a wage increase for wages that are already above the minimum a more effective instrument to increase firms' scope for selection, reduce employee turnover, and improve morale and discipline.

This suggests a spillover effect of a minimum wage: an increase in the minimum wage will raise the lowest wages as well as those that are just above the minimum level. Empirical findings support this prediction [2]. Such an interpretation is also in line with the observation that actual wages substantially exceed collectively set standard wages in countries with centralized wage setting [3]. In Germany, for example, many wage components such as extra pensions or a Christmas bonus (also termed the "13th salary"), are often offered as voluntary extras on top of collectively bargained wages.

The size-wage effect

Large employers pay significantly more than smaller employers for comparable workers. This is the size-wage effect, also referred to as the "firm-size wage effect," the "employer-size wage effect," or the "establishment-size wage effect" [4]. It is one of the most robust empirical labor market regularities, but conflicts with demand–supply theory, which assumes that the same kind of work should receive the same level of wages, regardless of the size of the employer. The efficiency wage channels are able to account for the size-wage effect:

- *Selection.* Large firms employ more capital per worker and operate with a deeper internal division of labor than comparable small firms. This permits them to achieve higher labor productivity, but means that they also incur higher overhead costs. The performance of the workforce is therefore more important for large firms than for small firms. Large firms make higher wage offers than small ones in order to attract more applicants so that they can select better performing employees, leaving the others to the smaller firms.
- *Turnover.* As large firms face higher overhead costs than small firms, the losses induced by higher rates of labor turnover are more pronounced for them than for small firms. Hence, large firms have a stronger incentive to reduce labor turnover by offering higher wages.
- *Discipline.* As labor productivity and overhead costs are higher in large firms, the losses due to sloth (i.e. lazy or unproductive workers) are more pronounced for them than for small firms. Hence, they are prepared to pay more than small firms to reduce sloth.
- *Morale.* If higher wages help to induce better morale, this channel becomes more important for large firms than for small firms, because they operate with higher labor productivity while facing larger overhead costs. In other words, if morale suffers, then large firms stand to lose more, financially, because each unit of lost production costs them more than small firms due to the larger overhead costs.

The congestion-wage effect

Wages for the same kind of work tend to be higher in congested metropolitan areas, as compared to sparsely populated regions [5]. The different channels through which efficiency wage theory works can account for this regularity and offer a plausible explanation for this effect:

- *Selection.* Because commuting costs are on average lower in congested areas, an increase in the wage offer reaches more potential applicants, some possibly working for other firms in the region. Hence, the effect obtainable from raising the wage offer will be more pronounced in a congested metropolitan region than in a sparsely populated rural region.
- *Turnover.* Labor supply reacts more strongly to wage changes in congested areas because more similar opportunities are available to prospective workers; the turnover effect will thus be stronger.
- *Discipline.* Because workers have more employment alternatives in metropolitan areas, the threat of dismissal is less effective at maintaining discipline; wages must therefore be higher in order to prevent sloth and shirking.
- *Morale.* This channel seems to be of less relevance regarding the congestion-wage effect.

Wage discrimination

Discrimination against certain employees occurs if workers with similar productivity characteristics are paid wages that differ according to group affiliation. Such discrimination is mostly discussed with respect to gender or race/ethnicity. In the following, only discrimination against women is considered, but the argument also applies to differential payment for similar workers across different occupations [6].

The mechanisms are similar to those discussed in the context of the size-wage effect and the congestion-wage effect, and hinge on the observation that women are traditionally more concerned with domestic duties, including childcare, than men. This makes non-financial job attributes more important to women than to men, and thus renders pay relatively less important.

- *Selection.* If women react less to higher wages than men do, the selection effect of higher wages will be less significant for women than for men. Firms will therefore rely less on the use of higher wage offers to broaden the pool of applicants for female-dominated jobs than for male-dominated jobs.
- *Turnover.* Because women care more for non-financial job attributes than men, firms will not use pay as a means of reducing turnover to the same extent for women as they would for men.
- *Discipline.* If women value certain attributes more than men, such as proximity to home or more family-friendly working hours, then they will likely have fewer alternatives than men to find employment elsewhere. The discipline effect from potential dismissal is thus not as reliant on higher wage premiums for women as it is for men, as firms can use these non-wage characteristics as leverage in lieu of wages.

- *Morale.* This channel seems to be of lesser relevance in supporting wage discrimination of women.

Hiring standards over the business cycle

Hiring standards vary over the business cycle. In periods of high unemployment, firms have ample choice among applicants, so hiring standards are rather stringent. Firms are willing to accept less qualified applicants in periods of high employment, due to greater hiring competition and a smaller applicant pool. The latter is extensively documented for multiple labor markets, even those that differ considerably in their institutional features, like the US and Germany [7].

At the same time, with high unemployment, wage offers tend to be lower relative to the market rate, as the selection problem is not as severe as in periods of high employment. As a consequence, wages vary pro-cyclically (i.e. wages go up when overall employment levels improve/during economic upturns, and go down during periods of high unemployment/economic downturns), hiring standards and over-qualification vary counter-cyclically (i.e. they are higher in periods of high unemployment and lower during low unemployment).

Inefficiency features of efficiency wages

Efficiency wages often generate different pay for the same type of work. Discrimination implies this directly, because—in the case of gender discrimination—women are paid less than men for the same work. In a similar fashion, the size-wage effect implies that two similar workers working in establishments of different size will be paid differently, and the congestion-wage effect leads to different pay for the same type of work in rural versus metropolitan areas. Efficiency wages that optimize the turnover cost/wage cost trade-off imply that the same work is rewarded differently according to the responsiveness of wages to turnover costs, and the selection mechanism entails that similar work may be rewarded differently according to the responsiveness of the pool of applicants to changes in the wage offered. Whenever the same type of work is paid differently, this implies economic inefficiency.

The basic reason for this assertion can be illustrated as follows. Consider a baker who produces two kinds of bread. One requires much more expensive ingredients than the other, but otherwise the resources used to produce one loaf (i.e. time and energy) are the same in both cases. Here, the more expensive type of bread ought to be more expensive than the other type due exclusively to the price difference in ingredients, because all other resources going into production are identical. If the price difference were larger than the extra costs, this would direct consumption in the wrong way. Some consumers would buy the more expensive loaf at a price difference equal to the extra costs, but may refrain from doing so at a larger price mark-up. This would cause inefficiency, as the consumer would end up with the cheaper loaf but would have preferred the more expensive loaf if they were only required to pay the real extra costs. If there are two otherwise identical bakers, each specializing in one type of bread, and the one who bakes the more expensive bread is paid better, the price difference between the two types of bread would exceed the extra costs by the wage difference between the bakers, and inefficiency would arise.

This example can be generalized: Whenever price differences do not reflect differences in social costs, prices induce wrong allocation. Wage differences that are unrelated to differences in the kind of work generate such unwarranted price differences. These types of differences are generated when firms utilize efficiency wages. Based on the argument above, which suggests that progressive taxation reduces the incentive for firms to use the efficiency wage mechanism, progressive taxation can be a means of reducing such inefficiencies by curbing excessive wage inequality.

Empirical evidence

Empirical and laboratory evidence about the channels described throughout this paper are usually discussed separately. A detailed discussion would surpass the scope of the present article in space and technicality, but it can be said that evidence does exist in support of each channel. There is empirical evidence that large firms hire more productive applicants [4] and that labor turnover is reduced if pay is better [8], while laboratory evidence exists that suggests better pay induces better performance [9]. Further, there is massive empirical evidence for the size-wage effect, the congestion-wage effect, and wage discrimination [10], and some empirical evidence suggests that more progressive wage or income taxation reduces pre-tax wage inequality, as is predicted by efficiency wage theory [1].

Relationship to other theories

Phenomena like the size-wage effect, the congestion-wage effect, or wage discrimination are often held to be incompatible with standard demand–supply theory, but this is not entirely true. These phenomena may actually also be accounted for by standard demand–supply theory as well as by other theories. By making suitable assumptions about preferences and technology, nearly everything can be “explained” in the standard demand–supply framework: the size-wage effect can be described by an aversion to work in large firms, discrimination by discriminatory preferences, or the congestion-wage effect by a combination of the advantages of agglomeration with a preference for living in the countryside, etc.

Yet, some phenomena accounted for by efficiency wage theory elude demand–supply theory, such as labor markets characterized by a permanent excess of supply. (Excess supply in this sense means that there are, in the labor market under consideration, permanently more sufficiently qualified workers available than needed, without inducing a lowering of the wage level. Note that the workers in excess supply are not necessarily unemployed but may hold lower-paying jobs elsewhere.)

As such, efficiency wage considerations do not exclude other theoretical approaches, but rather complement them. Take dynamic monopsony theory [10]. A monopsonist (e.g. a market structure in which only one buyer interacts with many would-be sellers of a particular product) may care about selection, turnover, discipline, or morale just like anyone else, and take into account that its labor demand will affect the optimal efficiency wage that is to be offered [11]. Similarly, search theory can consider turnover costs, discipline, and morale effects just as it often includes selection effects, right from the beginning [12]. As a consequence, it does not appear very useful to test a monopsonistic

theory of wage setting or some search theory against some version of an efficiency wage theory. The important question really is to find out what can be gained by including efficiency wage considerations.

LIMITATIONS AND GAPS

Even though empirical evidence mostly supports the ideas and channels presented by efficiency wage theory, it remains controversial whether the size of a single effect is truly strong enough to lift wages above the level that would suffice to attract just the desired number of applicants. Yet, even if each single channel is rather weak, the joint effect of the channels may be quite considerable. For this reason, it is perhaps preferable to concentrate on the joint effects of the examined channels, rather than looking at them individually, as they add up and reinforce each other. However, the strength of the joint effect is even more difficult to gauge than those of each single channel individually. For lack of an independent measurement of the strength and interaction of the joint effects working through the different channels, only the overall predictions can be checked empirically, such as the effects of tax progressivity on pre-tax wage differentials. Large-scale systematic studies may be in order here.

Moreover, there is not much empirical and theoretical work available that explores how efficiency wage concerns on the side of the firms work out in collective bargaining, and how labor legislation may strengthen or weaken the desire by the firms to make use of efficiency wage setting.

SUMMARY AND POLICY ADVICE

Efficiency wage theory, which explains the implications of wage setting through different channels, mostly through firms' selection, labor turnover, workers' discipline, and morale, contributes to a unified understanding of a range of diverse and empirically important labor market phenomena.

Each single phenomenon—such as discrimination, the size-wage effect, and the congestion-wage effect—has been rationalized by many theoretical considerations. The strength of the efficiency wage approach is that it suggests a comprehensive explanation for all these diverse and empirically well-established phenomena. Therefore, these phenomena offer collaborative empirical support for the efficiency wage perspective.

For policy purposes, it is important to empirically evaluate such joint effects of the different channels, which is unfortunately currently lacking. Nevertheless, separate evaluation of each phenomenon supports the efficiency wage hypothesis, which could have important policy implications, particularly by improving the efficiency of income taxation.

A potentially important policy implication of efficiency wage theory relates to the increase in wage inequality, a trend that is currently causing many political and social concerns. Efficiency wage theory accounts for this trend by pointing out that technical progress has entailed, in most occupations, a sharp increase in non-routine cognitive tasks, such as doing research, planning, or selling, and a pronounced decline in algorithmic or routine tasks [13]. It can be expected that this will further sharpen the problems of selection, turnover, and morale in the course of technological progress and thereby render efficiency wage effects more pronounced over time, leading to increasing (and inefficient) wage

inequality and over-education. To counter such socially and economically problematic tendencies, the incentives for firms to use the efficiency wage mechanism ought to be curtailed by means of taxation, labor legislation, and collective bargaining mechanisms.

Acknowledgments

The author thanks two anonymous referees and the IZA World of Labor editors for many helpful suggestions on earlier drafts. The author also thanks Johannes Ludsteck and Joachim Möller.

Competing interests

The IZA World of Labor project is committed to the *IZA Guiding Principles of Research Integrity*. The author declares to have observed these principles.

© Ekkehart Schlicht

REFERENCES

Further reading

Akerlof, G. A., and J. Yellen (eds). *Efficiency Wage Models of the Labor Market*. Cambridge: Cambridge University Press, 1986.

Katz, L. F. "Efficiency wage theories: A partial evaluation." In: Fischer, S. (ed.). *NBER Macroeconomics Annual 1986*. Cambridge, MA: MIT Press, 1986.

Key references

- [1] OECD. *Divided We Stand. Why Inequality Keeps Rising*. Paris: OECD Publishing, 2011; Table 2.4.
- [2] Dolton, P., C. R. Bondibene, and J. Wadsworth. "Employment, inequality and the UK national minimum wage over the medium term." *Oxford Bulletin of Economics and Statistics* 74:1 (2012): 78–106.
- [3] Schlicht, E. "Wage generosity." *Journal of Institutional and Theoretical Economics* 148:3 (1992): 437–451.
- [4] Oi, W., and T. Idson. "Firm size and wages." *Handbook of Labor Economics* 3 (1999): 2165–2214.
- [5] Glaeser, E. L., and D. C. Maré. "Cities and skills." *Journal of Labor Economics* 19 (2001): 316–342.
- [6] Schlicht, E. "Selection wages and discrimination." *Economics* 4 (2010): 1–30.
- [7] Büttner, T., P. Jacobebbinghaus, and J. Ludsteck. "Occupational upgrading and the business cycle in West Germany." *Economics* 4 (2010): 1–37.
- [8] Campbell, C. M. "Do firms pay efficiency wages? Evidence with data at the firm level." *Journal of Labor Economics* 11:3 (1993): 442–470.
- [9] Fehr, E., G. Kirchsteiger, and A. Riedl. "Does fairness prevent market clearing? An experimental investigation." *The Quarterly Journal of Economics* 108:2 (1993): 437–459.
- [10] Manning, A. *Monopsony in Motion: Imperfect Competition in the Labor Market*. Princeton, NJ: Princeton University Press, 2013.
- [11] Dube, A. "The envelope (theorem) please: Profits, efficiency wages, and monopsony." 2015. Online at: <http://arindube.com/2015/06/10/the-envelope-theorem-please-profits-efficiency-wages-and-monopsony/>
- [12] Mortensen, D. "Job search, the duration of unemployment and the Phillips curve." *American Economic Review* 60 (1970): 847–862.
- [13] Spitz-Oener, A. "Technical change, job tasks, and rising educational demands: Looking outside the wage structure." *Journal of Labor Economics* 24:2 (2006): 235–270.

Online extras

The **full reference list** for this article is available from:

<http://wol.iza.org/articles/efficiency-wages-variants-and-implications>

View the **evidence map** for this article:

<http://wol.iza.org/articles/efficiency-wages-variants-and-implications/map>