BANK OF CANADA
BANQUE DU CANADA

Working Paper/Document de travail
2015-35

The Carrot and the Stick: The Business Cycle Implications of Incentive Pay in the Labor Search Model

by Julien Champagne

Bank of Canada Working Paper 2015-35

September 2015

The Carrot and the Stick: The Business Cycle Implications of Incentive Pay in the Labor Search Model

by

Julien Champagne

Canadian Economic Analysis Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
julienchampagne@bankofcanada.ca

Bank of Canada working papers are theoretical or empirical works-in-progress on subjects in economics and finance. The views expressed in this paper are those of the author. No responsibility for them should be attributed to the Bank of Canada.

ISSN 1701-9397 © 2015 Bank of Canada

Acknowledgements

I am grateful to André Kurmann, Alain Delacroix, and Steve Ambler for their encouragement and many helpful comments. I also thank seminar participants at UQAM, the 2014 Canadian Economic Association annual meetings (Vancouver), the 2014 CEF annual conference (Oslo), and the 2014 European Economic Association meetings (Toulouse).

Abstract

This paper considers a real business cycle model with labor search frictions where two types of incentive pay are explicitly introduced following the insights from the micro literature on performance pay (e.g. Lazear, 1986). While in both schemes workers and firms negotiate ahead of time-*t* information, the object of the negotiation is different. The first scheme is called an "efficiency wage," since it follows closely the intuition of the shirking model by Shapiro and Stiglitz (1984), while the second is called a "performance-pay" wage, since the negotiation occurs over a wage schedule that links the worker's wage to the worker's output. The key feature here is that the worker can then adjust the level of effort (i.e. performance) provided in any period. I simulate a shift toward performance-pay contracts as experienced by the U.S. labor market to assess whether it can account simultaneously for two documented business cycle phenomena: the increase in relative wage volatility and the Great Moderation. While the model yields higher wage volatility when performance pay is more pervasive in the economy, it produces higher volatility of output and higher procyclicality of wages, two results counterfactual to what the U.S. economy has experienced during the Great Moderation. These results pose a challenge to the idea that higher wage flexibility through an increase in performance-pay schemes can account for business cycle statistics observed over the past 30 years.

JEL classification: E24, J33, J41
Bank classification: Business fluctuations and cycles; Labour markets

Résumé

L'auteur se penche sur un modèle de cycle réel avec frictions de recherche sur le marché du travail dans lequel il introduit explicitement deux modes de rémunération incitative, s'inspirant des constats d'un courant de littérature microéconomique sur la rémunération au rendement (Lazear, 1986, entre autres). Bien que, dans les deux cas, la négociation entre les travailleurs et les entreprises soit fondée sur l'information disponible avant la période *t*, l'objet de la négociation n'est pas le même. Le premier mode de rémunération est appelé « salaire d'efficience », car il est étroitement lié à l'intuition derrière le modèle dit du « tire-au-flanc » (*shirking*) de Shapiro et Stiglitz (1984), tandis que l'autre est qualifié de salaire « au rendement », puisque l'enjeu de la négociation consiste en une formule salariale dans laquelle le salaire du travailleur est fonction de sa production. L'élément-clé du salaire « au rendement » est que le travailleur peut moduler le niveau d'effort déployé (c.-à-d. le rendement) à chaque période. À l'aide de ce modèle, l'auteur simule une transition vers une économie où les contrats de rémunération au rendement sont plus fréquents, à l'instar de celle qui s'est opérée aux États-Unis, afin d'établir si cette transition permet d'expliquer simultanément deux phénomènes observés : l'accroissement de la volatilité relative des salaires et la Grande Modération. Le modèle génère une volatilité des salaires plus importante lorsque la rémunération au rendement devient plus répandue dans l'économie, mais il engendre aussi une volatilité accrue de la production et une procyclicité amplifiée des salaires. Ces deux résultats diffèrent de ce qui a été observé aux États-Unis durant la Grande Modération et remettent en question l'idée selon laquelle une plus grande flexibilité des salaires résultant de l'adoption d'un mode de rémunération au rendement peut expliquer les données statistiques sur le cycle économique des trente dernières années.

Classification JEL : E24, J33, J41
Classification de la Banque : Cycles et fluctuations économiques; Marchés du travail

Non-Technical Summary

The nature of business cycle fluctuations evolves and changes over time. A classic example of changing business cycle dynamics is the 25 years prior to the Great Recession, a period referred to as the Great Moderation, where the business cycle volatility of output and other macroeconomic aggregates fell by more than 50% relative to previous decades in the United States. However, this historically low macroeconomic volatility did not apply to one prominent labor market variable: real average hourly wages. Some authors have cited changes in labor market dynamics as a common explanation for both the decline in macroeconomic volatility and the increase in real wage volatility.

The increased incidence of performance-pay compensation schemes has been advocated as an explanation for the increase in wage volatility. For example, the incidence of performance-pay schemes has increased significantly during the past 30 years, and the wages of non-union workers with performance-pay contracts are more responsive to labor market shocks than are the wages of union workers without performance-pay contracts, implying that performance pay increases flexibility in wage setting.

This paper introduces two types of incentive pay schemes in a macroeconomic theoretical model of the business cycle with matching frictions in the labor market, and unobservable effort in production and wage bargaining between workers and firms. The two pay schemes differ in the way they incite the worker to supply the unobservable effort; in other words, the essence of incentive pay is different. The first one is an "efficiency-wage" type of pay scheme, where the worker is offered a fixed-wage amount in advance and the worker's effort can be monitored only with a given probability (think of a unionized employee who is paid a fixed salary independent of output). The second scheme is a "performance pay" one, where the contract simply specifies the wage as a function of output (think of a salesman whose salary is based mostly on commissions). The paper compares and analyzes the business cycle implications of the model for each compensation scheme separately. Finally, it evaluates how a structural change from one compensation scheme to the other, in the light of the above evidence, can account for the observed increase in the volatility of average real wages and the reduced volatility of other macroeconomic variables.

I find that while the model yields higher wage volatility when performance pay is more pervasive in the economy, it produces higher volatility of output and higher procyclicality of wages than that experienced during the Great Moderation. These results pose a challenge to the idea that higher wage flexibility through an increase in performance-pay schemes can account for business cycle statistics observed over the past 30 years.

1 Introduction

It has been well documented that the nature of business cycle fluctuations evolves over time. Many studies present evidence for changes in the dynamics of U.S. macroeconomic time series, such as McConnell and Perez-Quiros (2000), Stock and Watson (2002), Galí and Gambetti (2009), Galí and van Rens (2014). A classic example of changing dynamics is the 25 years prior to the Great Recession, a period referred to as the Great Moderation, where the business cycle volatility of output and other macro aggregates fell by more than 50%. However, this historically low macroeconomic volatility did not apply to one prominent labor market variable: real average wages. For instance, Champagne and Kurmann (2013) document that, from 1953-1983 to 1984-2006, the business cycle volatility of real average hourly wages relative to the volatility of aggregate output became 2.5 to 3.5 times larger over the two sample periods. As in Galí and van Rens (2014), they point toward changes in labor market dynamics as a common explanation for the decline in macro volatility and the increase in real wage volatility.

Among the documented changes in labor market dynamics, the increased incidence of performance-pay compensation schemes has been advocated as an explanation for the increase in wage volatility. For example, Lemieux et al. (2009a) show, using Panel Study of Income Dynamics (PSID) data, that the incidence of performance-pay schemes has increased significantly during the past 30 years in the United States. Moreover, Lemieux et al. (2009b) find that wages of non-union workers with performance-pay contracts are most responsive to local labor market shocks and least responsive for union workers without performance-pay contracts, implying that performance pay increases flexibility in wage setting. Finally, Champagne and Kurmann (2013) suggest that structural changes in the labor market, in the form of more flexible wage setting, are promising candidates to account for the increase in relative wage volatility.

Motivated by these observations, this paper first introduces two types of incentive pay schemes into a business cycle model with matching frictions and Nash bargaining. Second, it compares the business cycle implications of each compensation scheme, along with the basic the labor search model where the intensive margin is constant (e.g. Shimer (2005)). Finally, it evaluates how a structural change from one compensation scheme to the other, in the light of Lemieux et al.'s (2009a) evidence, can account for the observed increase in the relative volatility of average real wages and the dynamics of other labor market variables.

Specifically, I use a dynamic stochastic general-equilibrium (DSGE) real business cycle model with labor search frictions (e.g. Andolfatto, 1996; Trigari, 2009) and variable effort that is costly for the worker to supply. Then I use Lazear's (1986) insights on "input-based" and "output-based" compensation schemes to formulate two different wage-determination mechanisms.[1] Under each scenario, firms and workers negotiate *à la* Nash, but the wage outcome differs because the essence of incentive pay is different. Under the "input-based" scenario, workers and firms negotiate pay in advance subject to an incentive compatibility constraint that guarantees a minimum effort level (i.e. an efficiency-wage/shirking type of model). On the other hand, under the "ouput-based" wage contract (labelled "performance-pay" wage throughout the paper), the object of the negotiation is a wage schedule that links pay to effort (i.e. performance), which the worker supplies in order to maximize utility given the ex-ante negotiated wage schedule.[2] The first wage contract can be caricatured as "the stick," and the second as "the carrot."

Simulations of the model yield interesting results. First, the performance-pay scheme implies greater wage volatility than under the efficiency-wage scenario (and vs. the benchmark labor search model), a finding robust across different calibration strategies (e.g. Shimer, 2005; Hagedorn and Manovskii, 2008). This might suggest that changes in the way firms compensated workers over the past decades, i.e. from an efficiency-wage type of compensation to pay schemes that are linked to effort (and output), are at least partially responsible for the observed increase in relative wage volatility. Second, while the model is not able to replicate fluctuations in unemployment and vacancies as observed in the data (consistent with Shimer, 2005) under the preferred calibration strategy, it does fairly better under a more extreme calibration, as in Hagedorn and Manovskii (2008). When the economy is calibrated to match the average incidence of performance-pay wage contracts in the U.S. economy before and then after 1984, simulations show that an increase in the incidence of performance pay leads to an increase in relative wage volatility of about 10%. But it also leads to counterfactual results, such as an increase in output volatility and an increase in the correlation between wages and output. The reason is that effort is procyclical, amplifying the response of wages to technology shocks, but at the same time raising output volatility and the correlation between wages and output.

This paper introduces into a business cycle DSGE model ideas from the microeconomic literature on incentive pay. A large body of studies has researched into many forms of compensation schemes and pointed to different ways they can be used to incite effort from workers (e.g. Lazear, 1986;

[1] Lazear (1986) offers two simple examples to illustrate the difference between "input-based" and "output-based" wage contracts: "Two extreme examples are illustrative. Unskilled farm labor often is paid in the classic piece-rate fashion: an amount of payment per pound or piece harvested is specified in advance. Near the other extreme are middle managers of major corporations whose annual salaries are specified in advance, and who are then paid exactly that amount, independent of output. The qualifier is that, if effort falls below some specified level (e.g., he does not come to work regularly), the manager may be terminated."

[2] As put forward by the micro literature (e.g. Prendergast, 1999), this wage-determination mechanism provides a natural alternative incentive device for subtle effort supplies that are very hard to monitor.

Prendergast, 1999). On the macroeconomic side, some papers introduced variable effort in different contexts and studied its impact on different key macroeconomic variables over the business cycle (e.g. Burnside, Eichenbaum, and Rebelo (BER), 1993; in efficiency-wage frameworks: Alexopoulos, 2004; and Danthine and Kurmann, 2004; and in efficiency-wage and labor search frameworks: Costain and Jansen, 2010; Riggi, 2012). However, no studies have either tried to model the idea of performance pay in a DSGE framework, or looked at the consequences of incorporating different incentive pay schemes into a single DSGE framework. And while Costain and Jansen (2010) and Riggi (2012) study the implications of efficiency wages in a labor search framework, there is exogenous productivity in Costain and Jansen (2010), and no wage bargaining in Riggi (2012), two elements key to understanding the effects of incentive pay on wages and the business cycle in general.[3]

The rest of the paper proceeds as follows. Section 2 summarizes recent empirical evidence on the increased incidence of performance-pay contracts, deunionization and the increase in relative wage volatility in the United States, which all point toward more flexible wage setting in the past three decades. Section 3 presents the model with search frictions, variable effort and two different forms of incentive pay. Section 4 presents the calibration of the model, and section 5 the simulation results. Section 6 offers some conclusions.

2 Empirical evidence

This section summarizes the empirical evidence on the increased incidence of performance-pay wage contracts and relative wage volatility that serves as a motivation for the model and theoretical exercise developed in the next section.

The first piece of evidence comes from Lemieux, Macleod and Parent (2009a) who, using PSID data, document that the incidence of "output-based" compensation schemes (i.e. "performance-pay" contracts) has increased significantly during the 1980s and continued to rise (at a slower pace) in the 1990s, suggesting that it acted as an important driver behind the increase in wage inequality. At the same time, the United States experienced a sharp decline in unionization, which has been largely documented (e.g. Farber and Western, 2001; Hirsch and Macpherson (2010); Champagne

[3] Lastly, note that the concept of incentive pay in the current paper is very different from the "performance-pay" wage in Champagne and Kurmann (2013). Apart from the fact that the wage schedule here is determined ahead of time-t information, in the Champagne and Kurmann (2013) model the performance-pay wage is equal to the marginal rate of substitution between consumption and leisure hours times an optimal markup the worker commands, because of the imperfect substitutability of its labor service. The current paper thus presents a more serious, microfounded form of performance pay.

and Kurmann, 2013). Both phenomena lead to more flexibility in wage contracts between firms and workers. This evidence is shown in Figure 1; the left panel plots the non-union density for 1964 to 2006, while the right panel plots Lemieux, MacLeod and Parent's (2009a) measure of incidence of performance pay between 1976 and 1998, both for the non-farm business sector.

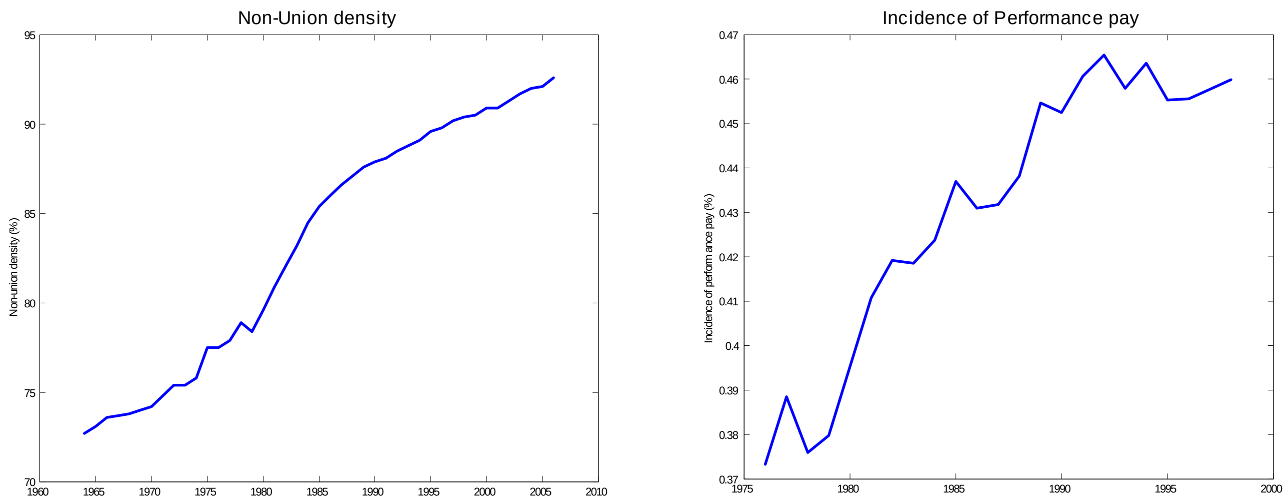

Figure 1: Evolution of non-union density (left panel) from non-farm business workers and incidence of performance pay (right panel) in the United States.

Second, Champagne and Kurmann (2013) document that, from 1953-1983 to 1984-2006, the business cycle volatility of average hourly wages increased by 15% to 60%, depending on the data set and filtering method used. As a result, the business cycle volatility of average hourly wages relative to the volatility of aggregate output became 2.5 to 3.5 times larger over the two sample periods. Champagne and Kurmann (2013) further document that this increase in relative wage volatility is pervasive across the labor market, albeit the magnitude of the increase varies for different groups of workers.[4] Table 1 presents a brief overview of these findings, updated to 2012, by showing volatilities and relative volatilities for real chained GDP, real average hourly compensation, and real

[4]On the individual level, it has been documented that earnings have also become more volatile in the past three decades. Starting with Gottschalk and Moffitt (1994), a number of papers using panel data show that labor income has, on average, become considerably more volatile across individual workers. Recent evidence based on PSID data by Dynan et al. (2008) and Jensen and Shore (2008) indicate that this increase in labor income volatility has remained approximately constant for most individuals but has increased greatly for individuals who already had volatile earnings in the past. Taken together, these panel studies imply that wages have become more volatile, on average, and much more volatile relative to output.

average weekly compensation for the non-farm business sector.[5,6] It shows that while the volatility of output decreased remarkedly after 1984, the volatility of both average hourly and weekly earnings has increased, so that the relative volatility (to output) of hourly and weekly earnings increased by a factor of 2.36 and 1.69, respectively, between the 1953Q2 to 1984Q1 and 1984Q2 to 2012Q4 periods.

	Standard Deviation			**Relative Standard Deviation**		
	Pre-84	*Post-84*	*Post/Pre-84*	*Pre-84*	*Post-84*	*Post/Pre-84*
Output	2.57	1.58	0.61	1.00	1.00	1.00
	(0.24)	(0.20)				
Avg. hourly comp.	0.65	1.02	1.56	0.26	0.65	2.50
	(0.06)	(0.10)		(0.03)	(0.11)	
Avg. weekly comp.	0.87	0.96	1.10	0.34	0.61	1.80
	(0.10)	(0.10)		(0.04)	(0.10)	

Notes : Standard and relative standard deviations for output, average hourly compensation and average w eekly compensation computed using quarterly, HP-filtered data. Total sample extends from 1953Q1 to 2012Q4. Non-farm business sector. PCE-deflated w ages. Standard errors computed using the generalized method of moments and the Delta method appear in parentheses below estimates.

Table 1: Business cycle volatilities.

While the timing of the increases in performance-pay contracts and in relative wage volatility might not be causal but coincidental, Lemieux et al. (2009a) document that performance pay is more frequent for skilled individuals that are employed in industries such as wholesale trade and finance, insurance, and real estate, and also more concentrated into the upper end of the wage distribution, which is precisely where wage volatility is highest and increased the most in the past three decades.[7] Finally, based on the same PSID data set as in Lemieux et al. (2009a), Lemieux et al. (2009b) find that wages of non-union workers with performance-pay contracts are most responsive to local labor market shocks, and least responsive for union workers without performance-pay contracts. Together, these observations suggest that the increased incidence of performance-pay contracts results in greater wage flexibility, making wages more responsive to business cycle shocks.

[5] I show volatilities for average weekly compensation because it is the appropriate measure of wages in the model presented below (i.e. there is no "hours margin" in the model).

[6] See the appendix for a detailed description of the data.

[7] See Champagne and Kurmann (2013) for a detailed account of the behavior of the relative volatility of wages across different segments of the workforce. Moreover, Champagne and Kurmann (2015) document that the increase in relative wage volatility was most pronounced for workers in the upper end of the wage distribution.

3 A DSGE model with incentive pay and matching frictions

The model I present in this section is a real business cycle DSGE model with a representative household, a continuum of firms offering a homogeneous good in a competitive market and labor search frictions. The model has two notable features. First, effort is a production input that firms cannot observe and therefore cannot directly contract upon. Second, in the spirit of Lazear (1986), firms incite effort from their workers according to one of two compensation schemes. The first scheme is one where a firm and a worker negotiate ahead of time-t information over a wage and a minimum required amount of effort.[8] With a given probability $0 < d < 1$, the firm can monitor whether the worker actually supplies this required amount of effort. A worker who is found to supply less than this level of effort is fired. I call this an "efficiency-wage" type of compensation, because it follows closely the intuition of the shirking model by Shapiro and Stiglitz (1984).[9] The second compensation scheme is one where the negotiation occurs again ahead of time-t information, but where the object of the negotiation is a wage schedule that links the worker's wage to the worker's output. The key feature here is that the level of effort can be adjusted by the worker in any period given the state of the economy. Consequently, even though the wage schedule is predetermined, the resulting wage is not. I call this compensation scheme "performance pay."

I assume a labor market where a fraction $1-p$ of firms negotiate with workers over an efficiency-wage type of compensation, while the remaining firms (fraction p) negotiate over a performance-pay wage. Firms cannot switch from one compensation scheme to the other.

Timing. After random matching occurs, the firm negotiates with the worker over one of the two types of compensation schemes, depending on its type. Then, shocks are observed, and firms take their optimal decisions over vacancies (for next period's hiring), while households choose their optimal consumption level. If an individual is matched to a performance-pay firm, the individual chooses the optimal level of effort given the bargained wage schedule.

Below I provide the details of the model, starting with a description of the labor market, the households' and firms' optimization problems, and finally the bargaining process. At this last stage I will describe separately the mechanisms that determine the efficiency wage and performance-pay wage, since this is where the differences arise.

[8] Consequently, the wage is predetermined in period t.

[9] Think of this efficiency-wage contract as one where the worker is offered a predetermined wage and, in return, has to show up to work and supply a fixed amount of effort. The firm can observe with probability d whether the worker shows up to work, and fire the worker for failing to do so.

3.1 Labor market

The labor market is characterized by matching frictions (e.g. Shimer, 2005). Search is not directed: unemployed workers automatically search at no cost and firms pay to post vacancies. Matching between unemployed individuals and vacancies occurs randomly according to an aggregate matching function:

$$m(v_t, u_t) = (v_t)^{\sigma} (u_t)^{1-\sigma}, \tag{1}$$

where u_t is the measure of workers searching for a job and v_t is the aggregate number of vacancies during period t. The parameter σ denotes the elasticity of job matches with respect to the vacancy input. Finally, I define the labor market tightness, θ_t, as the vacancy-unemployment ratio, $\frac{v_t}{u_t}$; the probability that an unemployed individual is matched to an open vacancy at date t is denoted $f_t = \frac{m_t}{u_t}$; similarly, the probability that any open vacancy is matched with a searching worker at date t is $q_t = \frac{m_t}{v_t}$. Households and firms take these probabilities as given.

Employment evolves according to the following dynamic equation:

$$n_{t+1} = (1 - s)\, n_t + f_t u_t. \tag{2}$$

At the beginning of period $t + 1$, employment is equal to the number of surviving matches from period t, plus the new ones ($m(v_t, u_t) = f_t u_t$). Matches are separated each period with exogenous probability s ($0 < s < 1$). The number of unemployed individuals at the beginning of any period t (when production occurs) is $1 - n_t$.[10] However, this is different from the number of individuals searching for a job during period t, which is given by

$$u_t = 1 - (1 - s) n_t. \tag{3}$$

The measures of unemployment ($1 - n_t$.) and job seekers (u_t) differ, since some workers who produced in period t can then be exogenously separated and search for next-period employment.[11]

3.2 Households

The households are thought of as very large "families" or "units" comprising a continuum of members along the unit interval. I label variables "*ew*" for those pertaining to the efficiency-wage

[10] Because the labor force is normalized to one, $1 - n_t$ also corresponds to the unemployment rate.

[11] I based this sequencing of events on the insights of Ravenna and Walsh (2012) to allow some workers to work and search in the same period. As they state in their paper: "In search models based on a monthly period of observation, it is more common to assume workers hired in period t do not produce until period $t + 1$. In this case, the number of job seekers in period t plus the number of employed workers adds to the total work force. Because we base our model on a quarterly frequency, we allow for some workers seeking jobs to find jobs and produce within the same period."

segment, and "pp" for the performance-pay wage segment. The household has period utility

$$u\left(c_t, e_t\right) = c_t - \left[(1-p)n_{ew,t}\frac{\left(e_{ew,t}\right)^{1+\eta}}{1+\eta} + pn_{pp,t}\frac{\left(e_{pp,t}\right)^{1+\eta}}{1+\eta}\right],$$

where c_t denotes consumption and $e_{ew,t}$ denotes the level of effort supplied by a fraction $1-p$ of household members employed by efficiency-wage firms; $e_{ew,t} \in [0, \overline{e}]$, depending on whether the employed member supplies the required amount of effort ($e_{ew,t} = \overline{e}$) or shirks ($e_{ew,t} < \overline{e}$). $e_{pp,t}$ denotes the level of effort supplied by household members employed by performance-pay firms, and η is a parameter governing the effort supply elasticity. The household's period utility thus includes the gain in utility of consuming c_t, minus the disutility of supplying effort sending $n_t = (1-p)n_{ew,t} + pn_{pp,t}$ members in the labor market.[12]

Households in each period face the following budget constraint:

$$\begin{aligned} c_t &= (1-p)n_{ew,t}w_{ew,t} + pn_{pp,t}w_{pp,t} \\ &\quad +(1-(1-p)n_{ew,t} - pn_{pp,t})b + \Pi_t - T_t, \end{aligned} \tag{4}$$

where b represents unemployment benefits (financed by lump-sum taxes on households, T_t); $w_{ew,t}$, $w_{pp,t}$ denote the efficiency- and performance-pay wages, respectively; and $\Pi_t = p\Pi_{pp,t} + (1-p)\Pi_{ew,t}$ denotes the household's profits share from the firms. Note that $w_{ew,t}$, $w_{pp,t}$, $e_{ew,t}$ and $e_{pp,t}$ will be determined during the bargaining process.

The household's value function can therefore be written as

$$W(\Omega_t) = \max_{c_t} \left\{ \begin{array}{c} c_t - \left[(1-p)n_{ew,t}\frac{\left(e_{ew,t}\right)^{1+\eta}}{1+\eta} + pn_{pp,t}\frac{\left(e_{pp,t}\right)^{1+\eta}}{1+\eta}\right] \\ +\beta E_t\left[W\left(\Omega_{t+1}\right)\right] \end{array} \right\},$$

subject to the budget constraint (4) and employment evolution (2). $\Omega_t = (n_t; z_t)$ represents the state vector of the economy.

Because I follow the labor search literature (e.g. Shimer, 2005; Hagedorn and Manovskii, 2008) and assume linear utility of consumption, the marginal utility of consumption of the household is constant; however, I provide an explicit form for the disutility of supplying effort instead of assuming that the outside option of the worker is constant and equal to b.[13]

[12] As is standard in the unemployment literature, I assume that households provide perfect consumption insurance to their members. As a result, the consumption and investment decision rules are the same for every household member. See Andolfatto (1996) for a detailed structure that implements this full-insurance assumption in a search and matching framework, or Alexopoulos (2004) for a detailed structure in an efficiency-wage context.

[13] I assume that b represents unemployment benefits, the "constant" portion of the worker's outside option. See the bargaining section for more details on the workers' outside option.

3.3 Firms

There is a continuum of identical firms on the unit interval. As stated above, a fraction $1-p$ of firms bargain with workers over an efficiency-wage type of compensation, while the remaining firms (fraction p) bargain with workers over a performance-pay wage schedule that links the worker's wage to output. Firms are owned by the households, and thus they discount expected future values according to

$$\Delta_{t,t+1} = \beta E_t \frac{u_1(c_{t+1})}{u_1(c_t)},$$

which is constant and equal to β because of linear utility. When a firm is matched with a suitable worker, it bargains over the wage and then observes time-t information. Thereafter, it chooses the number of vacancies to post, v_t (for next period's hiring) at fixed cost per vacancy κ, and finally produces according to the following linear production function:

$$F(n_t e_t; z_t) = y_t = z_t n_t e_t, \tag{5}$$

where z_t is a technology shock. As a result, in each period, the firm chooses the number of vacancies v_t to post such as to maximize the present discounted value of their future profits stream. Since this decision problem is similar for all firms (for a given bargained wage), the firm's value function can be written as

$$\begin{aligned} V(\Omega_t) &= \max_{v_t} \left\{ \begin{array}{c} F(n_t e_t; z_t) - n_t w_t - \kappa v_t \\ +\beta E_t \left\{ V\left(\Omega_{t+1}\right) \right\} \end{array} \right\} \qquad (6) \\ \text{s.to} &: \quad n_{t+1} = (1-s)\, n_t + q_t v_t. \end{aligned}$$

The first-order condition is

$$\kappa = \beta E_t \left\{ \frac{\partial V(\Omega_{t+1})}{\partial n_{t+1}} \frac{\partial n_{t+1}}{\partial v_t} \right\}, \tag{7}$$

where $\frac{\partial n_{t+1}}{\partial v_t} = q_t$. The value of an additional worker for the firm, i.e. $\frac{\partial V(\Omega_t)}{\partial n_t}$, is

$$\frac{\partial V(\Omega_t)}{\partial n_t} = V_n(\Omega_t) = \frac{\partial F(n_t e_t; z_t)}{\partial n_t} - w_t + \beta E_t \left\{ \frac{\partial V(\Omega_{t+1})}{\partial n_{t+1}} \frac{\partial n_{t+1}}{\partial n_t} \right\}.$$

Updating $V_n(\Omega_t)$ by one period, using equation (5) and substituting back into (7) yields the vacancy-creation condition:

$$\frac{\kappa}{q_t} = \beta E_t \left\{ \left[\frac{y_{t+1}}{n_{t+1}} - w_{t+1} + (1-s) \frac{k}{q_{t+1}} \right] \right\}. \tag{8}$$

The vacancy-creation condition states that, in equilibrium, the expected cost of hiring a worker is equal to the expected value of a match. Equation (8) shows that an increase in expected future profits will decrease q_t, implying that the number of posted vacancies must rise. This increase in vacancies will then increase employment next period.

3.4 Bargaining

As mentioned above, I assume that bargaining occurs before time-t shocks are realized. Since the bargaining problems for each compensation scheme differ substantially, I describe them separately below.

3.4.1 Efficiency-wage bargaining

Under this bargaining scenario, firms ask workers to supply a minimum amount of effort in return for a predetermined wage. They incite effort using a punishment scheme: with a given detection probability d, they can catch shirkers (if caught shirking, workers are fired). The important thing to note here is that this required level of effort is not an equilibrium outcome, as in Alexopoulos (2004), but an implicit assumption that firms can only monitor basic effort such as showing up to work: I assume that the constant detection probability d is an outcome of a contract enforcement device that can monitor only some such basic type of effort. However, it cannot help in enforcing more subtle effort supplies that are likely to be variable.[14]

The wage paid to workers is determined via bargaining over the match surplus before time-t is revealed. This surplus-sharing rule can be formulated as

$$w_{ew,t} = \underset{w_{ew,t}}{\arg\max} E_{t-1} \left\{ \left[W^{ns}_{n,ew}(\Omega_t) - W^{s}_{n,ew}(\Omega_t) \right]^{\xi} V_{n,ew}(\Omega_t)^{1-\xi} \right\}, \tag{9}$$

where $W^{ns}_{n,ew}(\Omega_t)$ and $W^{s}_{n,ew}(\Omega_t)$ are the values of being employed supplying effort level $\overline{e}$ and being employed shirking, respectively; $V_n(\Omega_t)$ is the firm's value of hiring an additional worker, ξ is the worker's bargaining power, and expectations are in $t-1$, since bargaining occurs before time-t shocks are realized. Even though the problem is standard, the household's surplus in the match is not. Why such a formulation of the household's surplus? Because under this efficiency-wage scenario, the "threat point" of the worker is not the value of being unemployed, but the value of shirking at work. A worker who does not get the minimum wage at which the no-shirking condition binds will shirk instead of going into the unemployment pool, because that worker is strictly better off shirking than being unemployed.

Before solving the bargaining problem, it is convenient to define the relevant surplus from employment for the firms and for the worker. As laid out above, the firm's surplus from employment

[14] Riggi (2012) assumes that the level of effort is not fixed and thus can vary with the state of the economy. For instance, after a negative shock to the level of capital, firms fire workers and those who keep their jobs increase their level of effort due to the "unemployment threat," thereby increasing productivity and having prolonged (negative) effects on employment and job creation. Here, I assume that the firm's monitoring technology does not permit verification of more subtle effort supplies.

is

$$V_{n,ew}(\Omega_t) = \frac{y_{ew,t}}{n_{ew,t}} - w_{ew,t} + \frac{\kappa}{q_t}(1-s). \tag{10}$$

For the household, the value of having an additional member employed is different whether the employed member supplies effort or not. Since every worker who supplies $e_{ew,t} < \bar{e}$ will be considered to be shirking, the household maximizes utility by choosing $e_{ew,t} = \bar{e}$ if the household wants its members to exert any effort, or $e_{ew,t} = 0$ otherwise.[15]

Consequently, we write the values (in terms of current consumption) of being employed supplying effort level $\bar{e}$, $W^{ns}_{n,ew}(\Omega_t)$, and of being employed shirking, $W^{s}_{n,ew}(\Omega_t)$, as[16]

$$\begin{aligned} W^{ns}_{n,ew}(\Omega_t) &= w_{ew,t} - b - \frac{\bar{e}^{1+\eta}}{(1+\eta)} + \beta\left[(1-s)(1-f_t)\right] E_t\left\{W_{t,ew}(\Omega_{t+1})\right\} \qquad (11) \\ W^{s}_{n,ew}(\Omega_t) &= (1-d)w_{ew,t} - b + \beta\left[(1-s)(1-f_t)(1-d)\right] E_t\left\{W_{t,ew}(\Omega_{t+1})\right\}. \end{aligned}$$

The first expression in (11) is standard: it states that the surplus from employment (in terms of current consumption) for a worker exerting the desired effort level $\bar{e}$, $W^{ns}_{n,ew}(\Omega_t)$, is equal to the worker's wage minus the forgone unemployment benefits and the cost of supplying effort, plus the discounted expected future value of being employed in the next period, i.e. $W_{n,ew}(\Omega_{t+1})$. The second expression, $W^{s}_{n,ew}(\Omega_t)$, states that the value of being employed shirking, in terms of current consumption, is the wage (discounted by the probability d of being caught shirking), less the forgone unemployment benefits, plus the discounted expected future value of being employed in the next period.

Incentive compatibility constraint. For workers to exert any effort, firms must offer workers a wage that satisfies their incentive compatibility constraint. Define this constraint as the "no shirking condition," expressed as $W^{ns}_{n,ew}(\Omega_t) \geq W^{s}_{n,ew}(\Omega_t)$. Using (11) above, we get:

$$\frac{\bar{e}^{1+\eta}}{(1+\eta)} \leq d\left[w_{ew,t} + \beta(1-s)(1-f_t)E_t\left\{W_{n,ew}(\Omega_{t+1})\right\}\right], \tag{12}$$

or, alternatively:

$$w_{ew,t} \geq \frac{\bar{e}^{1+\eta}}{(1+\eta)}\frac{1}{d} - \beta(1-s)(1-f_t)E_t\left\{W_{n,ew}(\Omega_{t+1})\right\}. \tag{13}$$

Workers will exert the desired amount of effort $\bar{e}$ only if the loss they would incur if detected shirking, weighted by the probability of being detected (d), is greater or equal to their disutility (in terms of current consumption) of supplying $\bar{e}$. This loss is the sum of two components: the forgone real wage value if detected shirking, plus the expected discounted value of a match in the next period. Consistent with the efficiency-wage literature (e.g. Shapiro and Stiglitz, 1984), the

[15] Since all workers are similar, there will be only one equilibrium efficiency wage and performance-pay wage.

[16] The detailed derivations of the surplus from employment are provided in the appendix.

"no shirking" wage is higher when: (i) the level of effort to be supplied is higher; (ii) the detection probability (d) is lower; (iii) the exogenous separation rate is higher (i.e. the fact that matches have a high probability of being terminated in the near future increases the incentive to shirk); (iv) the discount factor β is lower (since low value on employment next period implies lower loss if the worker is detected shirking).[17]

With the surpluses from the match defined, it is straightforward to solve the bargaining problem. The first-order condition yields the optimality condition:

$$(1-\xi)E_{t-1}\left\{W_n^{ns} - W_n^s\right\} = d\xi E_{t-1}\left\{V_n\right\}. \tag{14}$$

Expanding (14) using (10) and (11) and simplifying, we get the wage equation:

$$w_{ew,t} = \xi\left[E_{t-1}\left\{\frac{y_{ewt}}{n_{ew,t}}\right\} + (1-s)E_{t-1}\left\{\frac{\kappa}{q_t}\right\}\right] \tag{15}$$

$$+(1-\xi)\left[\frac{\overline{e}^{1+\eta}}{(1+\eta)}\frac{1}{d} - \beta(1-s)(1-E_{t-1}\left\{f_t\right\})E_{t-1}\left\{[W_{n,ew}(\Omega_{t+1})]\right\}\right], \tag{16}$$

where

$$W_{n,ew}(\Omega_t) = w_{ew,t} - b - \frac{\overline{e}^{1+\eta}}{(1+\eta)} + \beta(1-s)(1-f_{ew,t})E_t\left\{W_{n,ew}(\Omega_{t+1})\right\}.$$

The resulting efficiency wage is thus a predetermined variable; it is an expected sum (weighted by the worker's bargaining power) of the marginal product of a worker plus the expected cost of a vacancy and the discounted disutility of supplying effort, minus the (discounted) value of being employed next period.[18]

3.4.2 Performance-pay bargaining

Here, firms and workers negotiate prior to observing time-t shocks; after time-t information is revealed, workers can adjust their level of effort. The outcome of this negotiation is a predetermined wage schedule that links the worker's wage to output. As a result, the wage schedule $w_{pp,t}$ will satisfy

[17] Another interesting comparative static is the higher the job-finding rate f_t, the higher the no-shirking wage must be. Using equation (11), rewrite (13) as

$$w_{ew,t} \geq \frac{\overline{e}^{1+\eta}}{(1+\eta)}\frac{1}{d}$$
$$-\beta(1-s)E_t\left\{f_{t+1}w_{t+1}^{eff} - b + \frac{\overline{e}^{1+\eta}}{(1+\eta)}\left(\frac{1-d-f_{t+1}}{d}\right)\right\}.$$

The shorter the time it takes to get a job back after being fired, the higher the incentive to shirk.

[18] If the value of being employed next period is expected to be high, then the bargained efficiency wage will be lower today, since the worker has an incentive to stay on the job for the next period.

the optimality condition:

$$w_{pp,t} = \underset{\{w_{pp,t}\}}{\arg\max} E_{t-1}\left\{[W_{t,pp}(\Omega_t)]^{\xi} V_{t,pp}(\Omega_t)^{1-\xi}\right\}, \tag{17}$$

where, as above, $V_{n,pp} = \frac{y_{pp,t}}{n_{pp,t}} - w_{pp,t} + (1-s)\frac{\kappa}{q_t}$ is the firm's value of an additional worker, and $W_{n,pp}(\Omega_t)$ denotes the worker's surplus from employment (in terms of current consumption):

$$W_{n,pp}(\Omega_t) = w_{pp,t} - b - \frac{e_{pp,t}^{1+\eta}}{(1+\eta)} + \beta(1-s)(1-f_t)E_t\{W_{n,pp}(\Omega_{t+1})\}. \tag{18}$$

The first-order condition yields the optimality condition:

$$(1-\xi)E_{t-1}\{W_{2,pp}(\Omega_t)\} = \xi E_{t-1}\{V_{2,pp}(\Omega_t)\}. \tag{19}$$

Using the expression for $W_{2,pp}(\Omega_t)$ and $V_{2,pp}(\Omega_t)$, we get the wage schedule:

$$w_{pp,t} = \xi\left[z_t e_{pp,t} + (1-s)\kappa E_{t-1}\{\theta_{pp,t}\}\right] + (1-\xi)\left[\frac{e_{pp,t}^{1+\eta}}{(1+\eta)} + b\right], \tag{20}$$

where $e_{pp,t}$ is the optimal level of effort determined after observing time-t shocks (see optimal condition below). This wage schedule is similar to the basic labor search model with Nash bargaining, since it depends on both the marginal product of the worker and the worker's marginal rate of substitution. However, it also depends on $t-1$ expectations of labor market outcomes ($E_{t-1}\{\theta_{pp,t}\}$). The key feature in the above wage equation is that while the wage schedule is predetermined, $w_{pp,t}$ is not. This performance-pay scheme resembles a right-to-manage assumption, where workers have the right to manage their effort as a function of the bargained wage.[19]

Effort determination. After the wage schedule (20) is determined, workers observe shocks in t and choose the amount of effort to supply to maximize the value of being employed:

$$\max_{e_{pp,t}} \{W_{n,pp}(\Omega_t)\}$$

$$\text{s.to} \; : \; w_{pp,t} = \xi\left[z_t e_{pp,t} + (1-s)\kappa E_{t-1}\{\theta_{pp,t}\}\right] + (1-\xi)\left[\frac{e_{pp,t}^{1+\eta}}{(1+\eta)} + b\right].$$

The optimal effort condition is thus

$$e_{pp,t}^{\eta} = z_t. \tag{21}$$

The choice of effort equalizes the marginal product of effort and the marginal rate of substitution, and is privately efficient.[20]

[19] This right-to-manage analogy describes nicely the idea that performance pay creates an incentive mechanism, inciting workers to supply more effort.

[20] Note that private efficiency occurs only because of linear utility.

3.5 Aggregation and model dynamics

To close the model, I need to derive the aggregate identities for the variables that differ across the two firm types. First, the aggregate matching function is

$$\begin{aligned} m(u_t, v_t) &= v_t^{\sigma} u_t^{1-\sigma} \\ m(u_t, v_t) &= \left[p v_{pp,t} + (1-p) v_{ew,t}\right]^{\sigma} \left[p u_{pp,t} + (1-p) u_{ew,t}\right]^{1-\sigma}, \end{aligned}$$

since aggregate variables, such as output (y_t), vacancies (v_t), job searchers (u_t), and so forth, can be expressed as

$$\begin{aligned} y_t &= (1-p) y_{ew,t} + p y_{pp,t} \\ v_t &= (1-p) v_{ew,t} + p v_{pp,t} \\ u_t &= (1-p) u_{ew,t} + p u_{pp,t} \\ &\ldots \end{aligned}$$

For variables where we use the "per worker value," such as the average wage per worker or effort per worker, we aggregate these variables as

$$\begin{aligned} e_t &= (1-p) \frac{n_{ew,t}}{n_t} \bar{e} + p \frac{n_{pp,t}}{n_t} e_{pp,t} \\ w_t &= (1-p) \frac{n_{ew,t}}{n_t} w_{ew,t} + p \frac{n_{pp,t}}{n_t} w_{pp,t}. \end{aligned}$$

Moreover, we get the aggregate resources constraint from the household's budget constraint (4), substituting in the definition of profits and using Euler's theorem. This yields

$$y_t = c_t + \kappa v_t + T_t. \tag{22}$$

The model dynamics are obtained by taking a loglinear approximation around the steady state of the model. The appendix provides a complete set of the equations of the model.

4 Calibration and steady states

The model is calibrated to quarterly data for the U.S. economy. I lay out the calibration strategy in four steps. First, some parameters of the model are standard and thus are calibrated according to the related literature. For example, the quarterly discount factor β is set to 0.99; the elasticity of effort supply $(1/\eta)$ to 1; the elasticity of matches to vacancies, σ, to 0.6, which is about the midpoint of what is typically used in the literature;[21] ξ is set to 0.4 such that the Hosios condition

[21] See for example Andolfatto (1996), Petrongolo and Pissarides (2001), or Trigari (2009).

is satisfied; and s, the separation rate, is set to 0.10 as in den Haan, Ramey, and Watson (2000; DRW hereafter) and Shimer (2005). These standard parameters appear in the upper portion of Table 2.

Calibrated parameter values

Parameter	Definition	Value
β	Discount factor	0.99
$1/\eta$	Elasticity of effort supply	1
ξ	Worker's bargaining power	0.40
s	Separation rate	0.10
σ	Elasticity parameter, matching function	0.60
n	Employed / (Employed + unemployed)	0.89
1-n	Unemployment rate	0.11
f	Average job-finding rate	0.45
outsideOpt	Value of non-market activity	0.70
b/w	Unemployment benefits as a fraction of the wage	0.15
$\kappa v/y$	Vacancy-posting costs as a fraction of output	0.08
d	Implied detection probability, Efficiency wage	0.78

Table 2: Calibrated parameter and steady-state values.

The second step consists of finding steady-state values for n, f, u, and the unemployment rate $(1-n)$ using the steady-state equivalents of equations (2) and (3):

$$sn = fu$$
$$u = 1-(1-s)n.$$

To do this, I follow DRW's (2000) strategy, abstracting from labor force participation decisions and interpreting unmatched workers as including "both unemployed individuals and those not in the labor force but stating that they want a job." According to DRW (2000), the steady-state ratio of unmatched to matched workers (i.e. $(1-n)/n$), using the above definition of unmatched workers, is around 12%, which yields a value of $n = 0.89$. Consequently, the steady-state unemployment rate $1-n$ is equal to 0.11, and the job-finding rate f is equal to 0.45.[22]

The third step of the calibration strategy is less trivial because the wage equations differ substantially under each compensation scheme, while key labor market parameters, such as the

[22] Because I follow DRW (2000) to find target values for n, u, and f and cannot find different values for the efficiency-wage and performance-pay scenarios, I have to set $n_{ew} = n_{pp} = n$, and thus $u_{ew} = u_{pp} = u$. Consequently, I assume one steady state and look at the dynamics around it for both pay schemes.

unemployment-benefits-to-wage ratio (b/w), or the vacancy-creation cost κ, need to be the same under each bargaining scenario. I thus start with the remaining steady-state equations defining the performance-pay segment of the model, find what are the implied values for b/w and κ/q, and then solve the rest of the steady-state system.

The remaining equations defining the steady state of the performance-pay segment of the model are the steady-state equivalents of the production function (5); the vacancy-creation condition (8); the performance-pay wage equation (20); the effort condition (21); and the aggregate resources constraint (22). Since there is one more variable and free parameter than there are equations left, I need one other assumption to solve the system. I thus assume that the worker's outside option is equal to 70% of the wage, a value consistent with Hall and Milgrom (2008) and in between Shimer (2005) and Hagedorn and Manovskii (2008). With these assumptions in hand, solving the performance-pay segment is straightforward; first, note that the steady-state vacancy-creation condition in the performance-pay segment can be written as

$$\frac{\kappa v}{y_{pp}} = \frac{\kappa}{q} f \frac{u}{y_{pp}} = \frac{s}{(1/\beta - 1 + s)} \left[1 - \frac{w_{pp} n}{y_{pp}} \right], \tag{23}$$

while the performance-pay condition can be rewritten as

$$\frac{w_{pp} n}{y_{pp}} = \frac{\xi}{[1 - (1 - \xi) 0.7]} \left[1 + (1 - s) \frac{\kappa}{q} f \frac{n}{y_{pp}} \right]. \tag{24}$$

Substituting (24) into the vacancy-creation condition (23) defines aggregate vacancy-posting costs as a fraction of output (i.e. $\frac{\kappa}{q} f \frac{u}{y_{pp}}$) at a value of about 8%. Using the national accounting equation, the effort conditon and the production function, I can find values for $\frac{c_{pp}}{y_{pp}}$, e_{pp} and the steady-state output level y_{pp}. I finally find the wage w_{pp} from (24). With these steady-state values I can find (b/w) such that the outside option is equal to 70% of the steady-state wage, and $\frac{\kappa}{q}$ such that $\frac{\kappa}{q} f \frac{u}{y_{pp}} = 0.08$. These values are reported in the bottom half of Table 2.

The fourth and final step is to close the steady-state system by calibrating the remaining values in the efficiency-wage bargaining scenario. The additional parameter specific to this segment is d, the detection probability. For simplicity, I will assume that the labor share implied in my performance-pay calibration above is the same under the efficiency-wage scenario, and let d be determined by the steady-state system. This implies a detection probability of $d = 0.78$, similar to the preferred value of Riggi (2012).

One could argue that the steady-state values found above, especially from the third step onwards, are debatable. For example, aggregate vacancy-posting costs, at 8% of output, are arguably high.[23]

[23] Andolfatto (1996) sets vacancy-posting costs as a share of output at 1%, while Ravenna-Walsh (2012) set them at 5%.

In the next section, I will show that a very high value for the worker's outside option à la Hagedorn and Manovskii (2008) yields a higher labor share and, most importantly, lower vacancy-posting costs as a share of output. I will also provide another calibration strategy (details in the appendix) and show how simulation results differ under the alternative strategy.

Finally, for the technology shock, I assume that its logarithm follows an independent AR(1) process:

$$z_t = \rho_a z_{t-1} + \varepsilon_{zt} \text{ with } \varepsilon_{at} \; iid \; (0, \sigma^2_{\varepsilon_z}),$$

where $z_t = \log Z_t$. For the simulations below, I set $\rho_a = 0.978$ and $\sigma^2_{\varepsilon_z} = 0.80$, values found in Champagne and Kurmann (2013) for the 1953-2006 period.

5 Simulations

In this section, I first simulate the model either assuming all firms offer efficiency-wage compensation schemes (i.e. setting $p = 0$) and/or assuming performance-pay compensation schemes (i.e. setting $p = 1$), to picture how the model behaves under each bargaining scenario, and compare the results to U.S. data and to a benchmark labor search model without effort. If wages turn out to be more volatile in the performance-pay segment of the model, then a shift toward performance-pay schemes in the past 30 years, as documented in Section 2, could be a potential explanation for the observed increase in relative wage volatility during the same period. Second, I discuss the results along with the problems the search and matching model has in amplifying fluctuations in the labor market, and I propose an alternative calibration in the spirit of Hagedorn and Manovskii (2008) and show how it affects the simulation results. Finally, I provide a quantitative exercise where I vary exogenously the proportion of efficiency-wage and performance-pay firms in the economy (i.e. vary p), in the spirit of Champagne and Kurmann (2013), to show how a reasonable shift toward performance-pay schemes can account for the increase in relative wage volatility.

5.1 Impulse-response functions

To build intuition and to better understand the second moments below, it is useful to start by looking at impulse-response functions following a 1% technology shock, first when all firms offer efficiency-wage contracts, and then when all firms offer performance-pay contracts, separately. Under the efficiency-wage scenario, firm and worker negotiate prior to time-t information, and agree on a wage to be paid for a fixed amount of effort. Therefore, the wage is predetermined and the effort level cannot be adjusted by the worker, since it is constant. This is shown in Figure 2: when the technology shock hits the economy, the efficiency wage (dashed blue line) does not react on impact, since it is predetermined. In the next period, it jumps to adjust to the positive shock, and then decreases back at a sluggish pace toward its steady-state level.

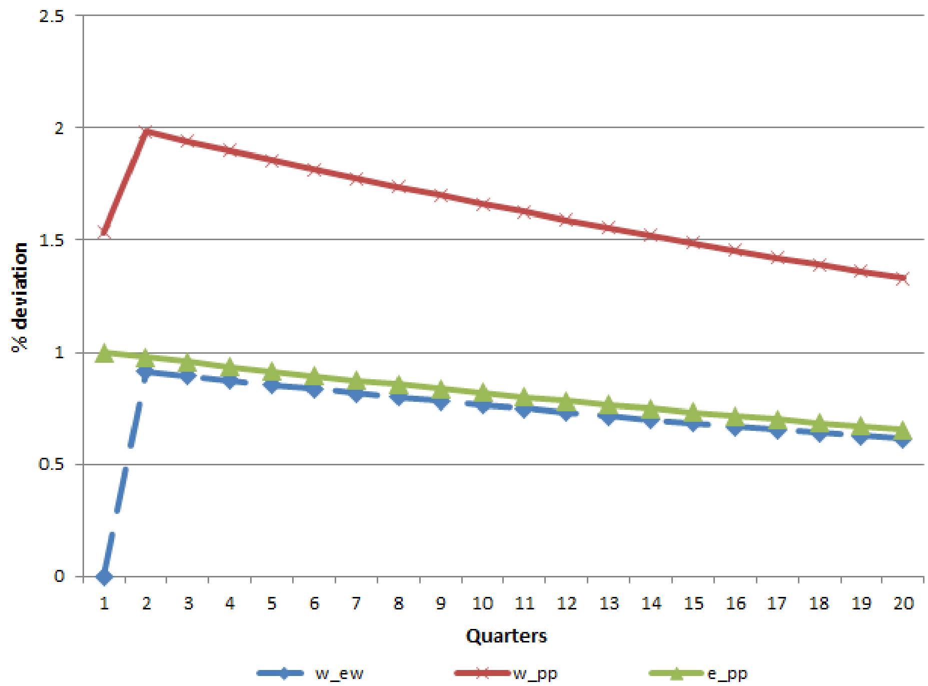

Figure 2. Impulse responses for different wage schemes following 1% technology shock.

On the other hand, the performance-pay wage is not predetermined, although the wage schedule is bargained before time-t technology shocks are realized. The reason is that workers adjust their effort supplies according to the state of the economy in period t. This is apparent from the red and green lines in Figure 2, which represent the impulse-response functions of the wage and effort following the same technology shock. The performance-pay wage reacts contemporaneously following a technology shock because of contemporaneous adjustments along the effort margin by workers. Since the performance-pay wage reacts each time a technology shock hits the model economy and the efficiency wage does not, we can safely assume that it is also more volatile than the efficiency wage. I show in the next subsection that it is indeed the case.

5.2 Second moments

Table 3 reports statistics summarizing the cyclical properties of the U.S. and model economies. The first three columns display second moments for the U.S. economy in three different subperiods;

the next three columns display moments from three model economies, i.e. a "benchmark model," the efficiency-wage segment, and the performance-pay segment of the model, respectively. The benchmark model is a standard labor search model without effort,[24] and wage bargaining occurs as in the standard search and matching literature (e.g. Pissarides, 2000; Shimer, 2005).

	US Data			Benchmark	Efficiency wage	Performance pay
	All	*Pre-84*	*Post-84*	*no effort*	*e={0,constant}*	*(σ(e)/ σ(y)=0.45)*
σ(y)	2.14	2.57	1.58	1.30	1.39	2.32
σ(n)/ σ(y)	0.75	0.66	0.95	0.27	0.45	0.15
σ(w)/ σ(y)	0.42	0.34	0.61	0.75	0.68	0.86
σ(y/n)/ σ(y)	0.75	0.66	0.95	0.80	0.75	0.90
σ(urate)/ σ(y)	6.17	5.59	7.04	2.29	3.73	1.28
σ(v)/ σ(y)	6.68	6.28	7.50	2.27	2.52	1.27
ρ(y,w)	0.51	0.75	0.18	0.98	0.83	0.99
ρ(y,y/n)	0.79	0.80	0.79	0.98	0.91	0.99
ρ(v,y/n)	0.87	0.91	0.82	0.92	0.99	0.92
ρ(v,urate)	-0.92	-0.94	-0.90	-0.29	-0.23	-0.29

Note : The first row reports the standard deviation for output; the next five rows report the relative standard deviations relative to standard deviation of output for, respectively, employment per capita, real average weekly wage, output per worker, the unemployment rate and vacancies index. The next two rows report the correlation between output and, respectively, the real average weekly wage and output per worker. The last two rows report, respectively, the correlation between vacancies and output per worker and the unemployment rate. All series are HP-filtered and the real wage series is PCE-deflated. U.S. data: Total sample extends from 1953Q1 to 2012Q4 with split in 1984Q1 using quarterly data for the non-farm business sector.

Table 3: Business cycle volatilities and correlations.

Data

The data sample for the U.S. non-farm business economy spans from 1953 to 2012, in quarterly terms; all series are logged and HP-filtered.[25] The first column displays second moments for the whole sample, whereas the second and third show second moments for two subsamples, before and after 1984Q1. The sample split is motivated by the Great Moderation literature that estimates a break in output volatility in 1984 (e.g. McConnell and Perez-Quiros, 2000),[26] whereas many other papers document the changing business cycle behavior of other prominent macroeconomic variables (e.g. Barnichon, 2010; Champagne and Kurmann, 2013; Galí and Van Rens, 2010; Nucci and Riggi, 2013). As documented in Champagne and Kurmann (2015) and shown in Table 3, the relative volatility of real earnings per worker (measured as real average weekly earnings) to output has increased significantly after 1984, from a ratio of 0.34 between 1953Q1 and 1984Q1 to a ratio of 0.61 between 1984Q2 and 2012Q4. Employment and labor productivity (measured as real-chained GDP

[24] Since there is no effort in the benchmark model, there is no disutility of providing effort in the worker's outside option (which equals b). Consequently, b in the benchmark model is calibrated to equal the same outside option value as in the efficiency-wage/performance-pay model.

[25] See the appendix for a detailed description of the data.

[26] Even though the Great Recession period (2007-2009) has been a turbulent period of economic activity, we see from Table 3 that output volatility has decreased by 40% after 1984, a period, up to the financial crisis, known as the Great Moderation.

divided by employment) also experienced increases in relative volatility after 1984, but to a lesser degree than earnings per worker. In terms of correlations, we see that real average weekly earnings were strongly procyclical before 1984, and mildly procyclical since 1984.[27] Labor productivity, as measured by output per worker, remained strongly procyclical throughout the sample, which stands in stark contrast to the vanishing procyclicality of output per hour documented in Galí and Van Rens (2010) and Champagne and Kurmann (2013).[28] Finally, as documented in many papers (e.g. Shimer, 2005; Hagedorn-Manovskii, 2008), there is a strong positive (negative) correlation between vacancies and labor productivity (unemployment).[29]

Model economies

The last three columns of Table 3 present the results for a benchmark model with no effort à la Shimer (2005), along with the model simulated with all firms bargaining over efficiency-wage contracts ($p = 0$) and over performance-pay contracts ($p = 1$). The first striking observation is that neither model comes close to matching the relative (to output) volatility of unemployment and vacancies, especially the performance-pay model. This result is a well-established one and known as the "Shimer puzzle" (Shimer, 2005). Moreover, when output is endogenous, and technology shocks are the only exogenous shocks in the model, it turns out that the labor search model is not only unable to generate a lot of amplification in unemployment and vacancies, but in output as well. This second result is consistent with DRW (2000), who show that without endogenous separations (i.e. fluctuations in the job-destruction rate), the labor search model does not propagate technology shocks well. The second striking, and more interesting, observation concerns the relative volatility of earnings per worker; in the performance-pay segment of the labor market, earnings are more volatile than in the efficiency-wage segment (by about 27%), and also vs. the benchmark model ($\sim 15\%$). Furthermore, as anticipated, the efficiency wage is less volatile than the standard Nash-bargained wage, implying that the shirking model along with the $t-1$ bargaining assumption induces stickiness in the wage. Finally, it is interesting to note that while the performance-pay model worsens the Shimer (2005) puzzle (vs. the benchmark case), the efficiency wage generates more amplification in unemployment and vacancies; because the wage is predetermined and does not react contemporaneously, firms have a greater incentive to post vacancies following a positive

[27] Even though the drop in correlation of weekly earnings with output after 1984 is consistent with the drop in the correlation with output of real average hourly earnings after 1984 (as documented in Champagne and Kurmann (2013, 2015)), the magnitude of the correlations is different. Champagne and Kurmann (2013, 2015) report that hourly wages went from being mildly procyclical before 1984 to mildly countercyclical after 1984.

[28] Champagne and Kurmann (2013), Table 3, show that the correlation between labor productivity (as measured by output per hour) and real non-farm GDP went from 0.65 (1964-1984) to 0.01 (1984-2006).

[29] Again, this is in stark contrast with the correlation between vacancies and output per hour; Barnichon (2010) reports (Table 1) that this correlation is 0.34 for the 1948-1984 period, and -0.31 for the 1984-2008 period.

technology shock.[30] Also note that all three models are able to generate a negative Beveridge curve, even though the curve is not as steep as in the data.

Lastly and unsurprisingly, the performance-pay model generates a larger correlation between the wage and output than the efficiency wage (and also than the benchmark model). It is unsurprising in the sense that, by definition, a "performance-pay" wage should follow production more closely than a more sticky wage. At the same time, it is inconsistent with the observed decline of this correlation in the U.S. data.

5.3 The Shimer puzzle and an alternative calibration strategy

As discussed in the calibration section above, one could argue that some values implied by the steady-state system of equations are more or less realistic, such as vacancy-posting costs at 8% of output. High vacancy-creation costs reduce the incentive to post vacancies, and thus worsen the unemployment/vacancies volatility puzzle (e.g. Shimer, 2005). Essentially, Shimer (2005) states that the standard labor search model cannot reproduce the volatility observed in the unemployment rate and vacancy posting: following a positive productivity shock, the increase in the job-finding rate pulls down unemployment and thus the $\frac{v}{u}$ ratio increases, raising the workers' threat point and consequently raising wages, which then take the bulk of the productivity increase and eliminate the incentive to post vacancies. Here, I propose a new calibration strategy in the spirit of Hagedorn and Manovskii (2008), where I set the worker's outside option at a very high value (95% of the steady-state wage value). As a result, $\frac{\kappa}{q}$ and b will be different than in the above strategy.

[30] This is the argument put forward by Shimer (2005) and Hall (2005), i.e. sticky wages are a potential solution to the unemployment and vacancies volatility puzzle.

Table 4 presents the parameters and steady-state values for this second calibration strategy.

Calibrated parameter values

Parameter	Definition	Value
β	Discount factor	0.99
$1/\eta$	Elasticity of effort supply	1
ξ	Worker's bargaining power	0.40
s	Separation rate	0.10
σ	Elasticity parameter, matching function	0.60
n	Employed / (Employed + unemployed)	0.89
1-n	Unemployment rate	0.11
f	Average job-finding rate	0.45
outsideOpt	Value of non-market activity	95%
b/w	Unemployment benefits as a fraction of the wage	0.44
$\kappa v/y$	Vacancy-posting costs as a fraction of output	1%
d	Implied detection probability, Efficiency wage	54%

Table 4: Alternative calibration strategy.

We can see that this new calibration strategy using an outside option value of 95% of the steady-state wage level, a very high number in the spirit of Hagedorn and Manovksii (2008), yields a very low value for vacancy-posting costs (as a share of output). This can be easily seen using equation (24); for a given labor share, the higher the outside option of the worker (as a fraction of the wage), the lower the vacancy-posting costs (as a share of output).

Table 5 presents the results for this alternative calibration strategy. As in Table 3, the first three columns present the same key data moments for three different sample periods, and the last three columns present moments for the model economies (i.e. the benchmark, efficiency-wage and performance-pay models).

	US Data			Benchmark	Efficiency wage	Performance pay
	All	*Pre-84*	*Post-84*	*no effort*	*e={0,constant}*	*(σ(e)/ σ(y)=0.28)*
σ(y)	2.14	2.57	1.58	2.76	3.89	3.67
σ(n)/ σ(y)	0.75	0.66	0.95	0.72	0.87	0.54
σ(w)/ σ(y)	0.42	0.34	0.61	0.32	0.23	0.50
σ(y/n)/ σ(y)	0.75	0.66	0.95	0.38	0.27	0.57
σ(u)/ σ(y)	6.17	5.59	7.04	5.98	7.28	4.50
σ(v)/ σ(y)	6.68	6.28	7.50	5.94	4.93	4.47
ρ(y,w)	0.51	0.75	0.18	0.83	0.75	0.96
ρ(y,y/n)	0.79	0.80	0.79	0.83	0.58	0.91
ρ(v,y/n)	0.87	0.91	0.82	0.92	0.99	0.92
ρ(v,u)	-0.92	-0.94	-0.90	-0.29	-0.23	-0.29

Note : The first row reports the standard deviation for output; the next five rows report the relative standard deviations relative to standard deviation of output for, respectively, employment per capita, real average weekly wage, output per worker, the unemployment rate and vacancies index. The next two rows report the correlation between output and, respectively, the real average weekly wage and output per worker. The last two rows report, respectively, the correlation between vacancies and output per worker and the unemployment rate. All series are HP-filtered and the real wage series is PCE-deflated. U.S. data: Total sample extends from 1953Q1 to 2012Q4 with split in 1984Q1 using quarterly data for the non-farm business sector.

Table 5: Business cycle volatilities and correlations.

First note that this calibration strategy yields substantially more fluctuations in output than the previous strategy in all three models. Second, as put forward by Hagedorn and Manovskii (2008), calibrating the outside option of workers at a very high proportion of the wage helps to solve the Shimer (2005) puzzle. As we see in Table 5, the benchmark model comes close to replicating the relative volatility of unemployment and vacancies to output. To understand this result, we simply need to look at the volatility of average wages: in all three models, the relative volatilities of wages are much lower than in the previous calibration. This implies more stickiness in the wage, increasing the firms' incentive to post vacancies. However, and most importantly, even if wages are less volatile in all three models under this calibration strategy, it is consistent with the previous calibration in two respects: (1) the performance-pay model generates higher relative volatilities of earnings per worker (about 91% higher than the efficiency wage, and 31% higher than the benchmark model); (2) the performance-pay model worsens the Shimer (2005) puzzle, in the sense that it yields lower relative volatilities of vacancies and unemployment than the benchmark model.

5.4 Quantitative exercise

To further assess how the increased incidence of performance-pay contracts can account for the increase in average wage volatility, I use the first calibration of the previous section, set the proportion of performance-pay firms in the economy (i.e. p) as in Champagne and Kurmann (2013), in order to match the pre-1984 average incidence of performance-pay contracts in the U.S. economy, and simulate the model. Then I set p to match the post-1984 average incidence of performance-pay contracts and again simulate the model to see how the change in the pervasiveness of pay-for-performance can account for the increase in relative wage volatility in the model. Table 6 presents the results of the quantitative exercise.

	US Data		Model	Model
	Pre-84	*Post-84*	**p=0.30**	**p=0.60**
$\sigma(y)$	2.57	1.58	1.68	1.90
$\sigma(n)/\sigma(y)$	0.66	0.95	0.33	0.25
$\sigma(w)/\sigma(y)$	0.34	0.61	0.71	0.82
$\sigma(y/n)/\sigma(y)$	0.66	0.95	0.82	0.86
$\sigma(urate)/\sigma(y)$	5.59	7.04	3.08	2.62
$\sigma(v)/\sigma(y)$	6.28	7.50	2.10	1.77
$\rho(y,w)$	0.75	0.18	0.93	0.96
$\rho(y,y/n)$	0.80	0.79	0.96	0.96
$\rho(v,y/n)$	0.91	0.82	0.98	0.95
$\rho(v,urate)$	-0.94	-0.90	-0.24	-0.27

Note: See Tables 3 and 5 for data description. p is set at 0.30 to match pre-1984 average of performance-pay incidence, and at 0.6 to match post-1984 average of performance-pay incidence.

Table 6: Business cycle volatilities and correlations.

Consistent with the results found in the previous subsection, more performance-pay contracts in the economy increase average wage volatility, but only by about 10%. Labor productivity also becomes more volatile, but the increase is again very small. Counterfactually, the increase in performance-pay schemes increases the volatility of output, along with the correlation between average wages and output. As mentioned earlier, this is due to the fact that effort is strongly procyclical under the performance-pay scenario, linking more closely the wage to output.

This quantitative exercise shows the challenges faced by advocates (e.g. Lemieux et al., 2009a; Champagne and Kurmann, 2013) of the performance-pay story (i.e. changes in the structure of pay over the past three decades as a driver of the increase in relative wage volatility and, by the same token, of increased macroeconomic stability): when one takes a microfounded approach to model performance pay in a general-equilibrium framework with explicit effort determination, it is difficult to generate business cycle statistics, as observed before and during the Great Moderation.

6 Conclusion

Some researchers have argued that changes in the dynamics of labor markets can be a potential explanation for the changing nature of business cycle fluctuations. For instance, during the Great Moderation, a period of unprecedented macroeconomic stability, average real wages have become more volatile in the U.S. economy, putting the labor market on the front stage.

Among the documented changes in labor market dynamics, the increased incidence of performance-pay compensation schemes has been advocated as an explanation for the increase in wage volatility. For example, Lemieux et al. (2009a) show that the incidence of performance-pay schemes has increased significantly during the past 30 years in the United States. Moreover, Lemieux et al. (2009b) find that wages of non-union workers with performance-pay contracts are most responsive to local labor market shocks, and least responsive for union workers without performance-pay contracts, implying that performance pay increases flexibility in wage setting. Finally, Champagne and Kurmann (2013) suggest that structural changes in the labor market, in the form of more flexible wage setting, are promising candidates to account for the increase in relative wage volatility.

Motivated by these observations, this paper first introduces two types of incentive pay schemes into a business cycle model with matching frictions and wage bargaining. Second, it compares the business cycle implications of each compensation scheme, along with the basic labor search model where the intensive margin is constant (e.g. Shimer (2005)). Finally, it evaluates how a structural change toward more performance-pay contracts, in the light of Lemieux et al.'s (2009a) evidence, affects the relative volatility of average real wages and other labor market variables.

Specifically, I use a real business cycle DSGE model with labor search frictions, and two types of incentive pay with explicit effort determination. To model incentive pay, I use Lazear's (1986) insights on "input-based" and "output-based" compensation schemes to formulate two different wage-determination mechanisms: one where bargaining occurs over an efficiency-wage (i.e. no-shirking) type of compensation scheme, and the other where bargaining occurs over a wage schedule that links pay to effort (i.e. performance) that is costly for the workers to supply. The first wage contract can be caricatured as "the stick," and the second as "the carrot."

Simulations of the model yield interesting and counterfactual results. First, the performance-pay scheme implies greater wage volatility than under the efficiency-wage scenario (and vs. the benchmark labor search model), a finding robust across different calibration strategies (e.g. Shimer, 2005; Hagedorn and Manovskii, 2008). Second, while the model is not able to replicate fluctuations in unemployment and vacancies as observed in the data (consistent with Shimer, 2005) under the preferred calibration strategy, it does fairly better under a more extreme calibration, as in Hagedorn and Manovskii (2008). Third, when the economy is calibrated to match the average incidence of performance-pay wage contracts in the U.S. economy before and then after 1984, simulations show that an increase in the incidence of performance pay leads to an increase in relative wage volatility of about 10%. But it also leads to counterfactual results, such as an increase in output volatility and an increase in the correlation between wages and output. The reason is that effort is procyclical, amplifying the response of wages to technology shocks, but at the same time raising output volatility and the correlation between wages and output.

These results pose a challenge to the theory that an increase in performance-pay wage contracts yields more flexibility in wage setting, increasing average wage volatility and lowering fluctuations in output. When one tries to model the idea of performance pay seriously (i.e. more microfounded with an explicit effort determination), it is difficult to generate the business cycle fluctuations observed over the past three decades.

Nonetheless, the idea of changes in the pay structure as an explanation behind changes in labor market dynamics and business cycle fluctuations is intriguing. Having a good theory of incentive pay in a DSGE framework can give good insight not only into the business cycle, but it can also be used to assess the influence of incentive pay on wage inequality, on economic growth, etc. Having the right incentive-pay framework would allow us to pursue other research avenues: for example, heterogeneous workers with different skill sets who sort themselves out toward (or not) performance-pay jobs; endogenous separation rates, where matches between low-skill workers and performance-pay jobs are terminated at a high frequency; or endogenizing the firm's decision to offer some type of incentive contract.

A Appendix

A.1 Data

All the data used in Tables 1, 3, 5, and 6 are in quarterly terms. Data from Figure 1 are taken from Champagne and Kurmann (2013).

- **Output**: Gross Domestic Product, Non-farm business, Chained-$2005. From the NIPA tables of the Bureau of Economic Analysis (BEA). Series ID: A358RX1. I divide this series by the U.S. population (see below) to get a GDP per capita measure.

- **Price deflator**: The main series used is the Personal Consumption Expenditure (PCE) deflator, from the NIPA tables of the BEA; index, 2009=100. Series ID: A002RD3.

- **Population**: Non-civilian population, 16 years old and over; from the Bureau of Labor Statistics' (BLS) Labor Productivity and Costs (LPC) program. Series ID: LNU00000000Q.

The rest of the variables come from the Major Productivity and Costs program of the BLS, which produces labor productivity and costs (LPC) measures for the private-sector U.S. economy.

- **Compensation**: Total compensation from the LPC data set comprises a "wages and salaries" component, and a "supplements" component.[31] The "wages and salaries" component is based on earnings data from the Quarterly Census of Employment and Wages (QCEW), previously known as the BLS ES-202 program. The QCEW is "...a cooperative program involving the Bureau of Labor Statistics (BLS) of the U.S. Department of Labor and the State Employment Security Agencies (SESAs)...[and] produces a complete tabulation of employment and wage information for workers covered by State unemployment insurance (UI) laws and Federal workers covered by the Unemployment Compensation for Federal Employees (UCFE) program." This represents about 98 percent of all U.S. jobs. The definition of labor earnings in the QCEW is very comprehensive. Specifically: "Wage and salary disbursements consist of the monetary remuneration of employees (including the salaries of corporate officers, commissions, tips, bonuses, and severance pay); employee gains from exercising nonqualified stock options; distributions from nonqualified deferred compensation plans; and an imputation for pay-in-kind (such as the meals furnished to the employees of restaurants)." See http://www.bea.gov/regional/pdf/spi2005/Complete_Methodology.pdf for more

[31] The proportion of wages and salaries in total compensation has been trending downwards in a constant way through time, from around 91% of total compensation in the mid-1960s to 80% in 2010.

information.

The "supplements" component consists of employer contributions for employee pension and insurance funds and employer contributions for government social insurance.[32] To derive total compensation for the non-farm business sector, the LPC subtracts compensation of employees working in public administration offices, in the farm sector, and in non-profit institutions and private households.[33] Moreover, the LPC imputes earnings of self-employed individuals using comparable data from workers in the CPS.

The total compensation measure we use from LPC is series ID: PRS85006063 (in levels), continuously updated each quarter here: www.bls.gov/lpc/special_requests/msp_dataset.zip.

- **Hours**: The total hours measure we use is LPC series ID: PRS84006033 (in levels), continuously updated each quarter here: www.bls.gov/lpc/special_requests/msp_dataset.zip.
- **Total employment:** I use LPC's employment series PRS85006013, which, as for compensation and hours above, is in levels.
- **Average weekly earnings:** I divide by 52 the ratio of total compensation to total employment.
- **Average hourly wage**: I compute average hourly earnings by dividing average weekly earnings with average weekly hours.

A.2 Surplus from employment

Here I describe in detail how I derive the households' value of having an additional member working under each wage-bargaining scenario.

A.2.1 Efficiency-wage sector

For illustrative purposes, let us assume that all workers and firms bargain over efficiency-wage contracts (i.e. $p = 0$). In that case, the household's problem becomes

$$W(\Omega_t) = \max_{c_t} \left\{ c_t - \left[n_{ew,t} \frac{(e_{ew,t})^{1+\eta}}{1+\eta} \right] + \beta E_t \left[W\left(\Omega_{t+1}\right) \right] \right\}, \tag{25}$$

[32] The estimates for the "supplements" portion of total compensation come from various sources, such as the IRS, the Medical Expenditure Panel Survey, or the American Council on Life Insurance. The estimates are compiled by the Bureau of Economic Analysis (BEA).

[33] Note that workers employed in "general government" are not included in the non-farm business measure, while workers in "governement enterprises" are.

subject to :

$$\begin{aligned} c_t &= \left(\begin{array}{c} \left[n^{ns}_{ew,t} + (1-d)n^{s}_{ew,t}\right] w_{ew,t} + \\ +(1-\left[n^{s}_{ew,t} + n^{ns}_{ew,t}\right])b + \Pi_{ew,t} \end{array} \right) \\ n_{ew,t+1} &= \left\{(1-s)\left[(1-d)n^{s}_{ew,t} + n^{ns}_{ew,t}\right] + f_{ew,t}u_{ew,t}\right\}, \end{aligned} \tag{26}$$

where $u_{ew,t} = 1-(1-s)\left[(1-d)n^{s}_{ew,t} + n^{ns}_{ew,t}\right]$. As mentioned in the main text, the surpluses from employment are different whether an employed member is shirking or supplying effort level $\overline{e}$. These surpluses are found by taking the first-order conditions of $W(\Omega_t)$ with respect to $n^{s}_{ew,t}$ and $n^{ns}_{ew,t}$, respectively, subject to the budget constraint and employment evolution equation above (26). This yields

$$\begin{aligned} \frac{\partial W(\Omega_t)}{\partial n^{s}_{ew,t}} &= W^{s}_{2,ew}(\Omega_t) = (1-d)w_{ew,t} - b + \beta E_t\left[W_2\left(\Omega_{t+1}\right)\frac{\partial n_{ew,t+1}}{\partial n^{s}_{ew,t}}\right] \\ \frac{\partial W(\Omega_t)}{\partial n^{ns}_{ew,t}} &= W^{ns}_{2,ew}(\Omega_t) = w_{ew,t} - b - \frac{\overline{e}^{1+\eta}}{(1+\eta)} + \beta E_t\left[W_2\left(\Omega_{t+1}\right)\frac{\partial n_{ew,t+1}}{\partial n^{ns}_{ew,t}}\right], \end{aligned} \tag{27}$$

where

$$\frac{\partial n_{ew,t+1}}{\partial n^{s}_{ew,t}} = \left((1-s)(1-d)(1-f_t)\right),$$

and

$$\frac{\partial n_{ew,t+1}}{\partial n^{ns}_{ew,t}} = \left((1-s)(1-f_t)\right).$$

Note that the surpluses from employment above are already expressed in terms of current consumption, since utility is linear. Rearranging yields

$$\begin{aligned} W^{s}_{2,ew}(\Omega_t) &= (1-d)w_{ew,t} - b + \beta\left[(1-s)(1-d)(1-f_t)\right]E_t\left[W_2\left(\Omega_{t+1}\right)\right] \\ W^{ns}_{2,ew}(\Omega_t) &= w_{ew,t} - b - \frac{\overline{e}^{1+\eta}}{(1+\eta)} + \beta\left[(1-s)(1-f_t)\right]E_t\left[W_2\left(\Omega_{t+1}\right)\right], \end{aligned} \tag{28}$$

which are equivalent to the surpluses from employment (11) in the main text.

A.2.2 Performance-pay wage sector

Again for illustrative purposes, let us assume now that all workers and firms bargain over performance-pay wage contracts (i.e. $p = 1$). The household's problem becomes

$$W(\Omega_t) = \max_{c_t}\left\{c_t - \left[n_{pp,t}\frac{\left(e_{pp,t}\right)^{1+\eta}}{(1+\eta)}\right] + \beta E_t\left[W\left(\Omega_{t+1}\right)\right]\right\}, \tag{29}$$

subject to :

$$\begin{aligned} c_t &= \left(n_{pp,t}w_{pp,t} + (1-n_{pp,t})b + \Pi_{pp,t}\right) \\ n_{pp,t+1} &= \left\{(1-s)\,n_{pp,t} + f_t u_{pp,t}\right\}, \end{aligned} \tag{30}$$

where $u_{pp,t} = 1 - (1-s)n_{pp,t}$. The surplus from employment (or the household's value of having an additional member employed) can be derived from the first-order condition (with respect to $n_{pp,t}$) of the household's problem (29) subject to (30):

$$\frac{\partial W(\Omega_t)}{\partial n_{pp,t}} = W_{2,pp}(\Omega_t) = w_{pp,t} - b - \frac{e_{pp,t}^{1+\eta}}{(1+\eta)} + \beta E_t \left\{ W_2\left(\Omega_{t+1}\right) \frac{\partial n_{t+1}}{\partial n_{pp,t}} \right\}, \tag{31}$$

where

$$\frac{\partial n_{t+1}}{\partial n_{pp,t}} = \left(\left(1-s\right)\left(1-f_{pp,t}\right)\right).$$

Rearranging yields

$$W_{2,pp}(\Omega_t) = w_{pp,t} - b - \frac{e_{pp,t}^{1+\eta}}{(1+\eta)} + \beta\left[\left(1-s\right)\left(1-f_t\right)\right] E_t\left[W_2(\Omega_{t+1})\right], \tag{32}$$

which represents the household's value, in terms of current consumption, of having one additional member employed. Equation (32) is equivalent to (18) in the main text.

A.3 System of equations

Here I list the system of equations of the model, including the firm-specific equations and aggregate identities.

1. Matching function:

$$m(v_t, u_t) = v_t^{\sigma} u_t^{1-\sigma}$$

2. Performance-pay employment evolution:

$$n_{pp,t+1} = \left(1-s\right) n_{pp,t} + f_t u_{pp,t}$$

3. Efficiency wage:

$$\begin{aligned} w_{ew,t} \quad &= \quad \xi \left[E_{t-1}\left\{ \frac{y_{ewt}}{n_{ew,t}} \right\} + (1-s)E_{t-1}\left\{ \frac{\kappa}{q_t} \right\} \right] \\ &\quad + (1-\xi)\left[\frac{\bar{e}^{1+\eta}}{(1+\eta)}\frac{1}{d} - \beta(1-s)(1-E_{t-1}\left\{f_t\right\})E_{t-1}\left\{\left[W_{n,ew}(\Omega_{t+1})\right]\right\} \right] \end{aligned}$$

4. Surplus from employment, under efficiency-wage scenario:

$$W_{n,ew}(\Omega_t) = w_{ew,t} - b - \frac{\bar{e}^{1+\eta}}{(1+\eta)} + \beta(1-s)(1-f_{ew,t})E_t\left\{W_{n,ew}(\Omega_{t+1})\right\}$$

5. Performance-pay wage:

$$w_{pp,t} = \xi\left[z_t e_{pp,t} + (1-s)\kappa E_{t-1}\left\{\theta_{pp,t}\right\}\right] + (1-\xi)\left[\frac{e_{pp,t}^{1+\eta}}{(1+\eta)} + b\right]$$

6. Effort condition, performance pay:

$$e_{pp,t}^{\eta} = z_t$$

7. Production function, efficiency-wage firm:

$$y_{ew,t} = z_t n_{ew,t}\bar{e}$$

8. Production function, performance-pay firm:

$$y_{pp,t} = z_t n_{pp,t} e_{pp,t}$$

9. Vacancy-creation condition, efficiency-wage firm:

$$\frac{\kappa}{q_t} = \beta E_t\left\{\frac{y_{ew,t+1}}{n_{ew,t+1}} - w_{ew,t+1} + (1-s)\frac{\kappa}{q_{t+1}}\right\}$$

10. Vacancy-creation condition, performance-pay firm:

$$\frac{\kappa}{q_t} = \beta E_t\left\{\frac{y_{pp,t+1}}{n_{pp,t+1}} - w_{pp,t+1} + (1-s)\frac{\kappa}{q_{t+1}}\right\}$$

11. Job-finding rate:

$$f_t = \frac{m(v_t, u_t)}{u_t}$$

12. Job-filling rate:

$$q_t = \frac{m(v_t, u_t)}{v_t}$$

13. Market tightness:

$$\theta_t = \frac{v_t}{u_t}$$

14. Aggregate number of job searchers:

$$u_t = 1 - (1-s)n_t$$

15. Unemployment rate:

$$urate_t = 1 - n_t$$

16. Aggregate output:

$$y_t = py_{pp,t} + (1-p)y_{eff,t}$$

17. Aggregate employment:

$$n_t = pn_{pp,t} + (1-p)n_{eff,t}$$

18. Average effort (or effort per worker):

$$e_t = p\frac{n_{pp,t}}{n_t}e_{pp,t} + (1-p)\frac{n_{ew,t}}{n_t}\bar{e}$$

19. Aggregate vacancies:

$$v_t = pv_{pp,t} + (1-p)v_{ew,,t}$$

20. Aggregate job searchers:

$$u_t = pu_{pp,t} + (1-p)u_{ew,,t}$$

21. Average wage:

$$w_t == p\frac{n_{pp,t}}{n_t}w_{pp,t} + (1-p)\frac{n_{ew,t}}{n_t}w_{ew,t}$$

22. Aggregate resource constraint:

$$y_t = c_t + \kappa v_t$$

References

[1] Alexopoulos, M., 2004. Unemployment and the business cycle. Journal of Monetary Economics 51, pp. 277–298.

[2] Andolfatto, D. 1996. Business cycles and labor-market search. American Economic Review, Vol. 86, No. 1 (March), pp. 112-32.

[3] Barnichon, R., 2010. Productivity, aggregate demand, and unemployment fluctuations. Journal of Monetary Economics 57, pp. 1013-1025.

[4] Burnside, C., Eichenbaum, M., Rebelo, S., 1993. Labor hoarding and the business cycle. Journal of Political Economy, Vol. 101, No. 2, pp. 245-73.

[5] Champagne, J., Kurmann, A., 2013. The great increase in relative wage volatility in the U.S. Journal of Monetary Economics 60, pp. 166-183.

[6] Champagne, J., Kurmann, A., Stewart, J. 2015. Reconciling the divergence in aggregate U.S. wage series. Working paper.

[7] Costain, J., Jansen, M., 2010. Employment fluctuations with downward wage rigidity: The Role of Moral Hazard. Scandinavian Journal of Economics, Wiley Blackwell, Vol. 112, No. 4 (December), pp. 782-811.

[8] Danthine, J-P., Kurmann, A., 2004. Fair wages in a New Keynesian Model of the business cycle. Review of Economic Dynamics, Elsevier, Vol. 7, No. 1 (January), pp. 107-142.

[9] den Haan, W. J., Ramey, G., Watson, J., 2000. Job destruction and propagation of shocks. American Economic Review, Vol. 90, No. 3 (June), pp. 482-498.

[10] Dynan, K.E., Elmendorf, D.W., Sichel, D.E., 2008. The evolution of household income volatility. Federal Reserve Board Finance and Economics Discussion Series 1997-61.

[11] Farber, H.S., Western, B., 2001. Accounting for the decline of unions in the private sector, 1973-1998. Journal of Labor Research 22, pp. 385-401.

[12] Galí, J., Gambetti, L., 2009. On the sources of the Great Moderation. American Economic Journal: Macroeconomics 1, pp. 26-57.

[13] Galí, J., Van Rens, T., 2014. The vanishing procyclicality of labor productivity. Working paper.

[14] Gottschalk, P., Moffitt, R.A., 1994. The growth of earnings instability in the U.S. labor market. Brookings Papers on Economic Activity 2, pp. 217-272.

[15] Hagedorn, M., Manovskii, I. 2008. The cyclical behavior of equilibrium unemployment and vacancies. American Economic Review, Vol. 98, No. 4 (September), pp. 1692–1706.

[16] Hall, R., 2005. Employment fluctuations with equilibrium wage stickiness. American Economic Review, Vol. 95, No. 1 (March), pp. 50-65.

[17] Hall, R., Milgrom, P., 2008. The limited influence of unemployment on the wage bargain. American Economic Review, Vol. 98, No. 4 (September), pp. 1653-1674.

[18] Hirsch, B.T., Macpherson, D.A., 2010. Union membership and earnings data book: compilations from the Current Population Survey. Bureau of National Affairs, Washington; plus associated database at http://www.unionstats.com.

[19] Jensen, S.T., Shore, S.H., 2008. Changes in the distribution of income volatility. Working paper.

[20] Lazear, E., 1986. Salaries and piece rates. Journal of Business, 59:3 (July), pp. 405-31.

[21] Lemieux, T., MacLeod, W.B., Parent, D., 2009a. Performance pay and wage inequality. Quarterly Journal of Economics 124, pp. 1-49.

[22] Lemieux, T., MacLeod, W.B., Parent, D., 2009b. Performance pay, wage flexibility, and hours of work. Working paper.

[23] McConnell, M. M., Perez-Quiros, G., 2000. Output fluctuations in the United States: what has changed since the early 1980s?. American Economic Review 90, pp. 1464–1476.

[24] Nucci, F., Riggi, M., 2013. Performance pay and changes in U.S. labor dynamics. Journal of Economic Dynamics and Control, vol. 37, Issue 12, pp. 2796-2813.

[25] Petrongolo, B., Pissarides, C., 2001. Looking into the black box: A survey of the matching function. Journal of Economic Literature, Vol. 34 (June), pp. 390–431.

[26] Pissarides, C., 2000. The unemployment volatility puzzle: is wage stickiness the answer?. Econometrica, Vol. 77, No. 5 (September), pp. 1339–1369.

[27] Prendergast, C., 1999. The provision of incentives in firms. Journal of Economic Literature 37, pp. 7-63.

[28] Ravenna, F., Walsh, C., 2012. Labor market flows with skill heterogeneity in a monetary policy model. Journal of Money, Credit and Banking, Vol. 44:s2 (December), pp. 31-71.

[29] Riggi, M., 2012. Capital destruction, jobless recoveries and the discipline device role of unemployment. Bank of Italy Working Paper 871.

[30] Shapiro, C., Stiglitz, J., 1984. Equilibrium unemployment as a worker discipline device. American Economic Review, Vol. 74, No. 3 (June), pp. 433-444.

[31] Shimer, R. 2005. The cyclical behavior of equilibrium unemployment and vacancies. American Economic Review, Vol. 95, No. 1 (March), pp. 25-49.

[32] Stock, J.H., Watson, M., 2002. Has the business cycle changed and why?. In: Gertler, M., Rogoff, K. (Eds.), NBER Macroeconomics Annual. MIT Press, pp. 159-218.

[33] Trigari, A. 2009. Equilibrium unemployment, job flows, and inflation dynamics. Journal of Money, Credit and Banking, Vol. 41, No. 1 (February), pp. 1-33.