Efficiency Wages and Inter-Industry Wage Differentials

Michelle Alexopoulos[*] [†]

July 2003

Abstract

Many empirical papers argue that a significant portion of inter-industry wage differentials is attributable to efficiency wage considerations. This paper presents a new multi-industry shirking efficiency wage model, in which detected shirkers face a monetary punishment, and examines whether it is consistent with the size and behavior of inter-industry wage differentials over time. The model suggests that: 1) small differences in detection rates can account for significant and persistent differentials, 2) high wage industries tend to have higher profits per worker and higher capital to labor ratios, and 3) inter-industry wage differentials are consistent with the empirical findings that they are generally acyclical.

Keywords: Inter-industry wage differentials, efficiency wages, business cycles
JEL Codes: E3, J3, J4

[*]Department of Economics, University of Toronto, Toronto, Canada M5S 3G7 e-mail: malex@chass.utoronto.ca

[†]The author would like to acknowledge helpful comments from Jon Cohen, Marty Eichenbaum, Angelo Melino, Steve Jones, Trisha Gladden, Michael McCraken, Marco Bassetto, Gadi Barlevy, participants of the Canadian Macroeconomic Study Group, the SED Summer Meetings, and seminar participants at the Minneapolis Federal Reserve, University of Western Ontario and SUNY. All error and omissions are the sole responsibility of the author.

"An understanding of the nature of inter-industry wage differentials could prove quite useful in determining the relevance of alternative models of wage determination."

-Dickens and Katz (1987)

I. INTRODUCTION

The presence of persistent inter-industry wage differentials in developed countries has been well documented by Krueger and Summers (1987, 1988), Dickens and Katz (1987), Murphy and Topel (1987), Gera and Grenier (1995), and Blackburn and Neumark (1992). Standard competitive labour market models offer two common explanations for the presence of these differentials. First, firms in different industries may offer identical workers different wages to compensate for different non-pecuniary job attributes. These compensating differentials ensure that identical individuals enjoy equal utility levels. Second, workers with the same measurable attributes in different industries may receive different wages if there are systematic differences in unmeasured worker ability that are correlated with industry affiliation.

Although the competitive explanations appear promising, Blackburn and Neumark (1992), Krueger and Summers (1988), Dickens and Katz (1987), and Borjas and Ramey (2000) have cast some doubt on the ability of these competitive explanations to account for much of the observed inter-industry wage differentials.[1] Instead, many of these authors argue that efficiency wages can account for a large component of the wage differentials and that efficiency wage models may also be consistent with the empirical regularities that suggest high wage industries are associated with higher capital to labour ratios and higher profits per worker

[1] Murphy and Topel's (1987) results are more supportive of the competitive labour model's unobserved ability explanation. However, even they found that on average approximately 30% of the estimated inter-industry wage differentials remained unaccounted for.

than low wage industries.[2] Despite their findings and the microeconomic evidence that suggests there is a link between the level of monitoring, worker productivity and wages, with the exception of Walsh (1999), no attempt has been made to directly evaluate the ability of a multi-sector shirking efficiency wage model to explain the presence of persistent inter-industry wage differentials.[3] Such an attempt is made in this paper using a general equilibrium framework.

In the first part of the paper, a formal model is developed, then analyzed, to determine if it is qualitatively consistent with: *(i)* the behavior of the inter-industry wage differentials observed in the data, and *(ii)* the high wage-low wage industry differences mentioned above.[4] The new model has three main features. First, firms imperfectly observe the effort levels of their workers during the production process. Second, the firms' ability to detect shirking workers varies across sectors. Third, detected shirkers are assumed to forgo some compensation. Absent bonding, these features can generate equilibrium unemployment, and cause firms to choose different levels of effort, wages and employment in equilibrium. An examination of the model's properties indicate that the model performs well along both dimensions of interest. I find that the model predicts: (i) the presence of persistent inter-industry wage differentials, and (ii) that high wage industries have higher capital per worker and higher profits per worker than low wage industries. Furthermore, versions of the model that allow for differences in worker ability suggest that differences in ability may not be able to explain sizable inter-industry wage differentials in economies where firms generally have problems detecting

[2] See Katz (1986) for a good overview of the different types of efficiency wage models.

[3] The links between wages, monitoring and productivity at the firm level have been explored by Krueger (1991), Groshen and Krueger (1990), Levine (1992), and Rebitzer (1995) among others.

[4] For a good discussion of the evidence that high wage industries also tend to have high profits per worker and high capital per worker, see Dickens and Katz (1987).

shirking workers.

In the second part of the paper, inter-industry wage differentials are estimated using data from the Current Population Survey (CPS) for the years 1973-1999. These estimates are used to determine the size and behavior of the inter-industry wage differentials in the U.S., and help calibrate the model. Three striking results emerge from this exercise. First, according to the model, only small differences in the detection rates are required to produce significant and persistent inter-industry wage differentials. Second, the inter-industry wage differentials predicted by the model are consistent with my empirical finding that the inter-industry wage differentials estimated from the CPS data are generally acyclical.[5] Third the movements of wages and employment, at the aggregate and industry levels, in the model are consistent with the low variation of real wages and high variation of employment observed over the business cycle.

All shirking efficiency wage models are based on Shapiro and Stiglitz (1984) in which a positive link exists between a worker's effort and his wage because firms imperfectly observe the effort levels of their workers.[6] In this type of model, individuals provide their contractual level of effort at the offered wage only if the expected utility of providing the effort is greater than or equal to the expected utility associated with shirking, (i.e., they will abide by the terms of their contracts if their incentive compatibility constraints are satisfied). Therefore, profit maximizing firms take the incentive compatibility constraints of their workers into account when choosing the terms of their employees' contracts, as well as the amounts of capital and

[5] The empirical result that the differentials are acyclical is consistent with Keane's (1993) findings using a different approach and data the National Longitudinal Survey of Young Men for the years 1966-1981. However, my finding that an efficiency wage model is consistent with this fact is contrary to Keane's conclusion that efficiency wage theories predict that the differentials should widen in recessions and shrink in expansions.

[6] Examples include Albrecht and Vroman (1992,1999), Gomme (1999), and Walsh (1999).

labour they will employ during the period. As a result, wages and employment in these models are determined by the interaction of the individuals' incentive compatibility constraints and labour demand, not the interaction between labour demand and labour supply as in the standard labour market clearing models. It follows that, if firms in different industries face different monitoring problems, the incentive compatibility constraints of identical workers hired by these firms will differ. This, in turn, causes the firms to offer identical workers different contracts. Consequently, a multi-sector shirking efficiency wage model may offer a compelling explanation as to why the wages of identical workers differ across industries.

In a novel attempt to determine if inter-industry wage differentials can be attributed to shirking efficiency wage considerations, Walsh (1999) analyses a multi-sector shirking efficiency wage model based on the original Shapiro-Stiglitz model. In his model, firms in different industries detect shirkers with different probabilities, and detected shirkers are fired. Using this model, he concludes that small differences in detection rates do not produce the large wage differentials attributed to the efficiency wage models in the empirical literature.

Although the methodology in Walsh's paper is appealing, the Shapiro-Stiglitz model used in his investigation is known to have problems reproducing the behavior of wages and employment over the business cycle.[7] These problems are attributable to the assumption that detected shirkers are fired in the model. Consequently, this paper develops an alternative framework in which detected shirkers are not dismissed but instead forgo a portion of their compensation.[8] This alternative "monetary punishment" is adopted for two reasons. First,

[7] See e.g., Gomme (1999).

[8] The model presented in this paper is similar in spirit to the type of shirking model investigated in Fehr, Kirchsteiger, and Riedl's (1996) experimental study of wage differentials and the shirking efficiency wage theory.

past work has demonstrated that a general equilibrium shirking efficiency wage model with this type of monetary punishment is better able to account for the low wage variation and high variation in employment seen in the U.S. data than standard business cycle models.[9] Second, survey evidence suggests firms more commonly rely on this type of "monetary punishment" to discipline workers than on outright dismissal.[10] The stark difference between Walsh's finding (i.e., that small differences in monitoring cannot explain sizable inter-industry wage differentials) and my finding that suggest the contrary are attributable to the differences in models' the punishments associated with shirking.

The remainder of the paper is organized as follows. Section II develops a multi-industry shirking efficiency wage model where detected shirkers forgo a bonus. Section III investigates the properties of this model to determine if they are consistent with: *(i)* persistent inter-industry wage differentials, and *(ii)* the empirical evidence about the links between wages and industry characteristics. In Section IV, an empirical analysis of the inter-industry wage differentials is presented, and Section V concludes the paper.

II. THE BASIC MODEL

For simplicity, the description of the model in this section focuses on the case where there are only 2 industries in the economy. However, it is simple to extend the model to allow for a finite number of different industries in the economy, and demonstrate that all of the results described in section III can be generalized.

[9] Papers such as Alexopoulos (2000), Burnside, Eichenbaum and Fisher (2000), and Felices (2001) examine the responses of this type of shirking efficiency wage model to shocks.

[10] Evidence in Agell and Lungborg (1995), Hall (1993), and Malcomson (1998), suggests that firms do not immediately fire detected shirkers. Instead, it is more common to reprimand detected shirkers, and remove opportunities from them that will result in monetary costs (e.g., lower bonuses or raises). In addition Bewley (1999) provides evidence that a number of firms tend to use bonuses and wage increases to motivate their workers.

The model is composed of four sectors. The first sector contains families and their members; the second sector is comprised of perfectly competitive final good firms; and the last two are made up of two monopolistically competitive intermediate good industries. Each sector's problem is discussed in detail below.

II.1. The Representative Family and its Members' Problem

Whenever there is a positive level of unemployment in equilibrium, if workers' income levels are not perfectly insured and individuals can transfer wealth between periods, the workers' problems become heterogeneous. In this model, the workers' problems are kept homogeneous by introducing a family construct.[11] Here it is assumed that workers cannot transfer wealth across periods, but their families can. Since some of the returns on assets are distributed to the family members, individuals' consumption levels are partially insured by their family. However, any one individual's employment status does not affect the savings decisions in the family, and when individuals make their decisions, the workers take as given the amount of consumption they expect to receive from their family.

The descriptions of the family and the family's problem are similar to the ones outlined in Alexopoulos (2000), Burnside, Eichenbaum and Fisher (2000), and Felices (2001). There are two main assumptions. First, there are a large number of identical families, each of which contain a [0,1]-continuum of identical members. Second, the families own all of the assets in the economy. Each family owns an equal share of the economy wide capital stock and an equal portion of firms' stocks in each industry. As a result, each family is entitled to a fraction of

[11] The properties of the model below are not significantly affected if, instead of having workers belonging to a family, it is assumed that there are two types of agents in the economy: entrepreneurs and workers. Similar to the environment seen in Gomme (1999), each worker would be unable to accumulate capital but would be endowed with some of the firms' shares, while each entrepreneur would be able to accumulate capital and would own the majority of shares in a firm. This case is discussed in Alexopoulos (2001).

each firm's profits.

In this model, the representative family must choose how much of its period t income (i.e., return on capital, $r_t^1 K_t^1 + r_t^2 K_t^2$, and profits from the different sectors, $\pi_t^1 + \pi_t^2$) to spend on investment in capital goods for each industry, I_t^1 and I_t^2, and how much to spend on final goods for their members to consume, c_t^f. This implies the following budget constraint for the family:

$$r_t^1 K_t^1 + r_t^2 K_t^2 + \pi_t^1 + \pi_t^2 \geq c_t^f + (K_{t+1}^1 - (1-\delta)K_t^1) + (K_{t+1}^2 - (1-\delta)K_t^2) \tag{1}$$

where δ is the depreciation rate of capital, K_t^j is the capital rented to industry $j \in \{1, 2\}$ and r_t^j is the rental rate of capital in industry $j \in \{1, 2\}$. The family chooses the levels of I_t^1, I_t^2 and c_t^f to maximize the expected discounted value of its lifetime utility subject to its budget constraint.[12]

II.1.1. Family Members

Individuals are assumed to have log-separable utility functions. The utility levels of individual family members employed in industry j and : *(i)* not shirking, *(ii)* shirking and detected, and *(iii)* shirking and not detected, are respectively:

$$U(c_t^j, e_t^j) = \ln(c_t^j) + \gamma \ln(T - \vartheta(e_t^j > 0)(e_t^j + \xi)) \tag{2}$$

$$U(c_t^{sj}, 0) = \ln(c_t^{sj}) + \gamma \ln(T) \tag{3}$$

$$U(c_t^j, 0) = \ln(c_t^j) + \gamma \ln(T) \tag{4}$$

where $\gamma > 0$, T is the individual's time endowment, e_t^j is the amount of effort the individual exerts on the job in industry $j \in \{1, 2\}$, and ξ is the fixed cost associated with providing any

[12] To keep the environment as simple as possible it is assumed that families do not believe that their choices can affect the employment probability of their members. This is done for simplicity. Alexopoulos (2001) provides assumptions that rationalize this assumption and leads to precisely the same allocations as in this model.

effort on the job.[13 ,14] In addition, $\vartheta(\cdot)$ is an indicator function that takes on the value 1 if the individual provides any effort and 0 otherwise. Further, c_t^j, and c_t^{sj} are the consumption levels of workers in industry j not detected shirking, and the consumption levels of workers in industry j who are detected shirking respectively, where $j \in \{1, 2\}$. An unemployed family member has a consumption level of c_t^u and a utility level of

$$U(c_t^u, 0) = \ln(c_t^u) + \gamma \ln(T) \tag{5}$$

The consumption of each of the family members during the period depends on their job market outcomes. Each member will either be: *(i)* employed in industry 1, *(ii)* employed in industry 2, or *(iii)* unemployed. In addition to the consumption differences caused by different employment statuses, workers' consumption levels can differ depending on whether or not they are detected shirking on the job.

In this simple version of the model, it is assumed that firms hire workers using a one period contract where a fraction, s, of a worker's possible wage during the period is paid to all workers while the remaining fraction, $(1 - s)$, of the total possible wage is only paid to workers not detected shirking during the period.[15,16] Consistent with the empirical evidence in surveys such as the Sage System Administrator Salary Profile (1999) and the International Consumer Service Association Incentive and Bonus Survey (2001), I assume that s is the same

[13] For simplicity, the notation that separates the actual effort provided and the level of effort specified in the firms' contracts is suppressed in the discussion since they are equal in equilibrium.

[14] To sharply focus on the effect of different detection rates, I will assume that these rates are the only difference between jobs, (e.g., ξ is the same across jobs).

[15] In this case, we can interpret sw_t^i as a base wage or salary, that is guaranteed to the worker, and $[1-s]w_t^i$ as a bonus payment.

[16] In this model s is an exogenous parameter and its value is chosen to be consistent with empirical evidence. However, the model delivers the same results as a model where there is a restriction on the minimum value of s, (e.g. a legal restriction or an industry norm), and s is chosen endogenously by firms. It is also possible to endogenously generate a constant positive value of s in this type of shirking model when firms can choose s, but know that they make mistakes when detecting shirking behavior (See Alexopoulos (2001)).

across industries. In addition, firms are assumed not to punish workers not detected shirking. These assumptions are made solely for simplicity. The results in this paper are unaffected if the model instead assumes: *(i)* there are continuing matches between workers and firms that breakup with an exogenous probability, *(ii)* a firm gets a reputation as a bad employer if it does not pay the bonus to non-detected shirkers, *(iii)* workers will not provide effort to bad employers because they believe that bad employers will fail to provide them with their bonus, and *(iv)* there are reasonable levels of markups in the economy.[17,18]

Letting N_t^j denote the number of workers employed in industry j, for $j \in \{1,2\}$, the consumption levels for non disciplined workers in industry j, detected shirkers in industry j and unemployed family members are:

$$c_t^j = w_t^j - Tr_t + c_t^f \tag{6}$$

$$c_t^{sj} = sw_t^j - Tr_t + c_t^f \tag{7}$$

$$c_t^u = \frac{N_t^1 + N_t^2}{(1 - N_t^1 - N_t^2)} Tr_t + c_t^f \tag{8}$$

Here, w_t^j is the real wage paid to individuals employed in industry j, Tr_t is the amount of the transfer that is paid to unemployed members in the economy. Employed workers pay Tr_t as an intra-family transfer to increase the consumption levels of the unemployed in their family.[19]

[17] This richer environment is closer to that seen in MacLeod and Malcomson (1998). An examination of this alternate environment demonstrates that when the household's discount rate, β, equals $\left(\frac{1}{1.03}\right)^{0.25}$, and when markups are greater than 1%, the firm will have no incentive to withhold bonuses from non-detected shirkers since the value of withholding the bonus in any period is less that the discounted value of the future profits they will lose because of the bad reputation they get if they withhold the bonuses, (i.e., the firm's incentive compatibility constraint will not be binding). As a result, the firm's problem will reduce back to a one period problem. Further, since it is assumed that the probability that the match breaks up is exogenous, in order to focus on the effects of having a monetary punishment, the individual's incentive compatibility constraints are the same as those in the simple version of the model.

[18] See Basu and Fernald (1994,1997) for some evidence on mark-ups for the U.S. economy.

In this simple case, all working members transfer the same amount to the unemployment insurance fund. Further, it is assumed that the transfers ensure that unemployed members enjoy the same utility as detected shirkers in the low wage industry. Therefore,

$$Tr_t = \min(sw_t^1, sw_t^2)(1 - N_t^1 - N_t^2) \tag{9}$$

In equilibrium, the model predicts that wages will be lower in the high detection industry and effort will be higher (See Propositions 2 and 3 below). Therefore, this transfer will imply that unemployed workers are involuntarily unemployed in equilibrium.[20] Although, assumptions about the intra-family transfer will have an affect on the dynamics of the model and will affect whether or not the unemployed family members are voluntarily or involuntarily unemployed, this assumption will not effect the theoretical properties of the model developed in the next section.[21]

II.2. The Family's Problem

Letting N_t^{sj} denote the number of shirking family members hired in industry j, and letting d_j denote the probability a shirker is detected by a firm in industry j, the family's problem

[19] This unemployment insurance could easily be modeled as being handled by the government, instead of the family.

[20] When the detection rate in industry 1 is less than the detection rate in industry 2 (i.e., $d_1 < d_2$), $e^1 < e^2$ and $w^1 > w^2$. Given these inequalities and the fact that the transfer is $Tr_t = sw_t^2(1 - N_t^1 - N_t^2)$, $c^{s1} > c^{s2} = c^u$. Since firms choose the terms for their contracts so that no one shirks in equilibrium, it follows from the individuals' incentive compatibility constraint that:

$$U(c_t^1, e_t^1) > U(c_t^2, e_t^2) = d_2 U(c_t^{s2}, 0) + (1 - d_2) U(c_t^2, 0) > U(c_t^u, 0)$$

because the individuals' utility function, $U(c, e)$, is increasing in c and decreasing in e. Therefore, unemployed individuals have a lower utility level than their working counterparts and these individuals are involuntarily unemployed in equilibrium.

[21] For a discussion about how changes in the levels of the intra-family transfers affects the quantitative properties in a similar model with only one sector of firms, see Alexopoulos (2000). In addition, Felices (2001) examines the effect of changing the form of unemployment insurance in a sticky price model.

can be expressed as:

$$\max_{\left\{K_{t+1}^1, K_{t+1}^2, c_t^f\right\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \begin{array}{c} [(N_t^1 - N_t^{s1})\, U(c_t^1, e_t^1) + (1-d_1) N_t^{s1} U(c_t^1, 0) + d_1 N_t^{s1} U(c_t^{s1}, 0)] \\ + [(N_t^2 - N_t^{s2})\, U(c_t^2, e_t^2) + (1-d_2) N_t^{s2} U(c_t^2, 0) + d_2 N_t^{s2} U(c_t^{s2}, 0)] \\ +(1 - N_t^1 - N_t^2) U(c_t^u, 0) \end{array} \right\}$$

subject to equation (1), where $c_t^j, c_t^{sj}, c_t^u, U(c_t^j, e_t^j), U(c_t^{sj}, 0), U(c_t^j, 0)$ and $U(c_t^u, 0)$ are described by equations (2)-(9) respectively.

Profit maximizing firms ensure that no one shirks in equilibrium. Furthermore, $r_t^1 = r_t^2$ for capital to be employed in both industries. It follows that, in equilibrium, the marginal utility of the family at time t is a weighted average of the marginal utility of the different types of family members, where the weights are given by the percent of family members of each type, $\left(\text{i.e., } \frac{N_t^1}{c_t^1} + \frac{N_t^2}{c_t^2} + \frac{1-N_t^1-N_t^2}{c_t^u}\right)$.

II.3. Final Good Firms

Firms in the final goods sector are perfectly competitive and use output from the intermediate goods firms for production according to the following production function:

$$Y_t = \left(a_1 \left(\int_0^1 Y_t^1(i)^{\frac{1}{\phi}} di \right)^{\frac{\phi}{\mu}} + a_2 \left(\int_0^1 Y_t^2(i)^{\frac{1}{\phi}} di \right)^{\frac{\phi}{\mu}} \right)^{\mu}$$

where $\mu \geq 1, \phi \geq 1$, and $Y_t^1(i)$ and $Y_t^2(i)$ are the goods from intermediate goods firm i in industry 1 and 2 respectively. The representative final good firm's problem can be written as:

$$\max_{Y_t, \left\{Y_t^1(i)\right\}_0^1, \left\{Y_t^2(i)\right\}_0^1} Y_t - \int_0^1 P_t^1(i) Y_t^1(i) di - \int_0^1 P_t^2(i) Y_t^2(i) di$$

Here $P_t^1(i)$ and $P_t^2(i)$ are the prices for intermediate goods from firm i in industry 1 and 2 respectively normalized by the price of the final good. Combining the first order necessary conditions, and the zero profit conditions the following equations emerge:

$$Y_t^j(i) = a_j^{\frac{\mu}{\mu-1}} Y_t(\overline{P}_t^j)^{\frac{\mu-\phi}{(\mu-1)(\phi-1)}} P_t^j(i)^{\frac{-\phi}{\phi-1}} \text{ and } \left\{ a_1^{\frac{\mu}{\mu-1}} \left(\overline{P}_t^1\right)^{\frac{-1}{\mu-1}} + a_2^{\frac{\mu}{\mu-1}} \left(\overline{P}_t^2\right)^{\frac{-1}{\mu-1}} \right\}^{1-\mu} = 1$$

where $\overline{P}_t^j = \left[\int_0^1 P_t^j(i)^{\frac{-1}{\phi-1}} di\right]^{1-\phi}$ for $j = 1, 2$.

For the remainder of the paper the following notation is used: $Y_t^1 = \int_0^1 Y_t^1(i)di$, $Y_t^2 = \int_0^1 Y_t^2(i)di$ and $y_t = \frac{Y_t^2}{Y_t^1}$. It is also assumed that all firms in an industry are symmetric and $a_1 = a_2 = 1$. It follows that:

$$\frac{P_t^1}{P_t^2} = \frac{P_t^1(i)}{P_t^2(i)} = y_t^{\frac{\mu-1}{\mu}}$$

Therefore, as the size of industry 1 increases relative to industry 2, y_t decreases and the price of goods in industry 1 falls relative to the price of goods in industry 2.

II.4. The Intermediate Goods Firms

In this economy, there are two different industries (sectors) with intermediate goods firms. The intermediate goods firms in both industries are monopolistic competitors. Entry of new firms over time is ruled out and no firms are able to switch industries. Intermediate goods firms are assumed to require effort from their employees in order to produce goods during the period. Since effort is only imperfectly observed, firms will choose to offer workers a contract that will satisfy their incentive compatibility constraint.

Given the level of the family's transfers to their unemployed members, the individuals' rationality constraint never binds, so it will always be in the best interest of workers to accept employment from firms. As a result, the relationship between a worker's wage and his effort level will be determined by the fact that firms will make workers just indifferent between shirking on the job and providing the profit maximizing effort level specified in the contract.

Given the demand function from the final good firms, the exogenous detection rate, d_j, and the fact that the workers' incentive compatibility constraints hold with equality in equilibrium, the monopolistically competitive firm's problem can be expressed as:

$$\max_{\{K_t^j, N_t^j w_t^j, P_t^j\}} P_t^j Y_t^j - r_t^j K_t^j - w_t^j N_t^j \, subject \ to$$

$$N_t^j \leq x_t^j$$

$$Y_t^j = \theta_t^j (K_t^j)^\alpha (e_t^j N_t^j)^{1-\alpha}$$

$$e_t^j = T\left(1 - \left(\frac{c_t^j}{c_t^{sj}}\right)^{-\frac{d_j}{\gamma}}\right) - \xi$$

$$Y_t^j = Y_t (\overline{P}_t^j)^{\frac{\mu-\phi}{(\mu-1)(\phi-1)}} P_t^{j\,\frac{-\phi}{\phi-1}}$$

Here N_t^j and K_t^j are the number of people hired and the amount of capital rented respectively, x_t^j is the number of people who are looking for work at the representative firm in industry j and θ_t^j is the level of technology in industry j.[22] Since there is unemployment in the U.S. economy, the remainder of the paper will focus on the model's results for the case when firms in neither sector are constrained by the number of employees looking for work, (i.e., the case where $x_t^1 > N_t^1$ and $x_t^2 > N_t^2$).

When the firm's level of employment is not constrained by the number of applicants, the

[22] Here I assume there is only one type of labour employed by firms. However, Alexopoulos (2002) demonstrates that a variant of this model where there are multiple types of jobs within a firm has the same reduced form production function as used in this paper. Moreover, the model with multiple occupations is consistent with the empirical observation that the wages of all workers employed in the high wage industry receive high wages, even if the model assumes that the probability of detection only differs for one of the occupations across industries.

first order necessary conditions for industry j, can be written as:

$$\frac{e_{w_t^j}(w_t^j)}{e(w_t^j)} w_t^j = 1 \text{ (Solow Condition)}$$

$$\frac{(1-\alpha)}{\phi} \frac{P_t^j Y_t^j}{N_t^j} - w_t^j = 0$$

$$\frac{\alpha}{\phi} \frac{P_t^j Y_t^j}{K_t^j} - r_t = 0$$

$$\text{where } e_t^j = e(w_t^j) = T\left(1 - \left(\frac{c_t^j}{c_t^{sj}}\right)^{-\frac{d_j}{\gamma}}\right) - \xi \text{ and } e_{w_t^j}(w_t^j) = \frac{\partial e(w_t^j)}{\partial w_t^j}$$

assuming firms are symmetric within industries. Using the fact that $w_t^j, e_{w_t^j}(w_t^j)$ and $e(w_t^j)$ can be expressed as functions of c_t^j and c_t^{sj}, the Solow condition implies that wages are chosen such that the ratio $\frac{c_t^j}{c_t^{sj}}$ is a constant greater than one, (i.e., the consumption of non-shirkers is greater than the consumption of detected shirkers), and effort is constant across all states of the world and across time, (i.e., $e_t^j = e^j$).

III. PROPERTIES OF THE MODEL WITH UNEMPLOYMENT

III.1. Wages and Effort in the Different Industries

The Solow conditions for the firms in different industries imply that wages will be chosen so that consumption of the non-shirkers is directly proportional to the consumption of the detected shirkers since:

$$\frac{c_t^j}{c_t^{sj}} = \Pi_j \text{ for } j \in \{1, 2\}$$

This equation allows us to derive the following expressions for wages:

$$w_t^j = \frac{(\Pi_j - 1)(c_t^f - Tr_t)}{(1 - s\Pi_j)} \text{ for } j \in \{1, 2\}$$

which implies:

$$\frac{w_t^1}{w_t^2} = \frac{(1 - s\Pi_2)(\Pi_1 - 1)}{(1 - s\Pi_1)(\Pi_2 - 1)} = \mathcal{D}$$

where $\mathcal{D}$ is the inter-industry wage differential. Clearly this differential does not depend on the time period, and is therefore very persistent over time and across different states of the world. The value of the differential, $\mathcal{D}$, is only affected by the parameters that affect the individuals' incentive compatibility constraint, (i.e., s, d_j, γ, T, and ξ).

In addition to proving that the model predicts a persistent differential, it is useful to determine if lower detection rates coincide with higher wages in the model, (i.e., if $d_1 < d_2$ implies that $w_t^1 > w_t^2$). To demonstrate that this is the case, the first proposition proves that $\Pi_1 > \Pi_2$ when $d_1 < d_2$.[23] Intuitively, this finding implies that as the detection rate drops, firms must increase the punishment associated with detection, (i.e., increase the wage of non-shirkers relative to the wage of detected shirkers), in order to prevent workers from shirking on the job. This result would be trivial to show if it was assumed that all firms in the economy require the same level of effort regardless of their detection rate.[24] However, in this model, firms in industries with low detection rates may choose a lower contractual effort level than firms with higher detection rates. If firms with the low probability of detecting shirkers choose a much lower contractual effort level than the firms in the industry with the high detection rate, they may offer their workers contracts that pay a lower wage than firms in the industry with the high detection rate. It is therefore necessary to examine how changes in the detection rates effects both the effort and wage levels. These results are provided in

[23] The proof for this proposition, and all others, are provided in Appendix A.

[24] If all firms required the same effort level, $w_t^1 > w_t^2$ since the individuals IC constraints will imply that $\left(\frac{c_t^1}{c_t^{s1}}\right)^{d_1} = \left(\frac{c_t^2}{c_t^{s2}}\right)^{d_2} \rightarrow \left(\frac{c_t^1}{c_t^{s1}}\right) > \left(\frac{c_t^2}{c_t^{s2}}\right)$ because $d_1 < d_2$ which implies $w_t^1 > w_t^2$.

Propositions 2 and 3.

Proposition 1: Firms will offer workers a contract that ensures that the consumption of non-shirking workers relative to the consumption of detected shirkers in the industry increases as the probability of detecting a shirker decreases, (i.e., $\Pi_1 > \Pi_2$ when $d_1 < d_2$).

This can be formally shown by using the assumption that the value of the constants, Π_j are not much larger that 1, and the Solow condition. From this Proposition it follows that wages will be higher in industries that cannot detect shirkers easily.

Proposition 2: In this economy, workers in the low detection industry will be paid a higher wage than workers in a high detection industry, (i.e., $\frac{w_t^1}{w_t^2} = \mathcal{D} > 1$).

Given that $d_1 < d_2$ implies $w_t^1 = w_t^2 \mathcal{D} > w_t^2$, proposition 2 can be used to generate the model's prediction with respect to the variance of wages in the different industries. Specifically, the model predicts that $Var(w_t^1) = \mathcal{D}^2 Var(w_t^2) > Var(w_t^2)$, (i.e., wages in the low detection industry will be more volatile than wages in the high detection industry).

The first two propositions illustrate that in the model $d_1 < d_2$ implies $\Pi_1 > \Pi_2$, and $w_t^1 > w_t^2$. However, since firms are free to alter the contractual level of effort, the finding that wages are higher in the low detection industry does not ensure that effort in the low detection industry is larger than the contractual effort in the high detection industry. Proposition 3 describes how the differences in the detection rates affect the contractual effort levels in the different industries.

Proposition 3: The contractual effort rate decreases as the probability of detection decreases, (i.e., $e^1 < e^2$ when $d_1 < d_2$).

This proposition indicates that firms in the low detection industry will require less effort than firms in the high detection industry. This occurs because as the detection rate decreases, the expected utility of shirking increases for a given wage and effort pair. To prevent shirking at the new detection rate, firms can satisfy the individual's incentive compatibility constraint by: *(i)* leaving the wage unchanged and reducing the required effort level, *(ii)* raising the wage and leaving the effort level unchanged, or *(iii)* raising wages and lowering the required effort. The firm's first order conditions in this model imply that profit will be maximized by firms adjusting both wages and effort levels in response to a change in the detection rate.

III.2. The Relative Size of the Industries

From the firms' first order necessary conditions, we can determine the relative size of the industries. Given that there are positive levels of capital rented by firms in both industries, the relationship between the industries' ratios of capital to labour can be expressed as:

$$\frac{\alpha}{\phi}\frac{P_t^1 Y_t^1}{K_t^1} = \frac{\alpha}{\phi}\frac{P_t^2 Y_t^2}{K_t^2} \rightarrow \frac{\frac{K_t^1}{N_t^1}}{\frac{K_t^2}{N_t^2}} = \left(\frac{P_t^1}{P_t^2}\frac{\theta_t^1}{\theta_t^2}\right)^{\frac{1}{1-\alpha}}\frac{e^1}{e^2}$$

Then, using the fact that $\frac{w_t^1}{w_t^2} = \mathcal{D}$ it follows that:

$$\begin{aligned}
\mathcal{D} &= \frac{P_t^1}{P_t^2}\frac{\theta_t^1}{\theta_t^2}\left(\frac{\frac{K_t^1}{N_t^1}}{\frac{K_t^2}{N_t^2}}\right)^{\alpha}\left(\frac{e^1}{e^2}\right)^{1-\alpha} \rightarrow \mathcal{D} = \left(\frac{P_t^1}{P_t^2}\frac{\theta_t^1}{\theta_t^2}\right)^{\frac{1}{1-\alpha}}\left(\frac{e^1}{e^2}\right) \\
\rightarrow & \frac{P_t^1}{P_t^2} = \left(\frac{e^2}{e^1}\mathcal{D}\right)^{1-\alpha}\frac{\theta_t^2}{\theta_t^1}
\end{aligned}$$

Since $\frac{P_t^1}{P_t^2} = y_t^{\frac{\mu-1}{\mu}}$, it is clear that:

$$y_t = \left[\left(\frac{e^2}{e^1}\mathcal{D}\right)^{1-\alpha}\frac{\theta_t^2}{\theta_t^1}\right]^{\frac{\mu}{\mu-1}}$$

As a result, the relative size of the industries is:

$$\frac{P_t^1 Y_t^1}{P_t^2 Y_t^2} = y_t^{\frac{-1}{\mu}} = \left[\left(\frac{e^2}{e^1}\mathcal{D}\right)^{1-\alpha}\frac{\theta_t^2}{\theta_t^1}\right]^{\frac{-1}{\mu-1}}$$

Therefore, as $\frac{\theta_t^2}{\theta_t^1}$ increases, industry one decreases in size relative to industry 2. Also, if $\frac{\theta_t^2}{\theta_t^1}$ is constant, y_t is constant and the relative size of the industry remains unchanged over time.

III.3. The Capital to labour Ratio, and the Relative Size of Capital and Employment in the Industries

The Euler equations also imply that:

$$\mathcal{D} = \frac{P_t^1}{P_t^2}\frac{Y_t^1}{Y_t^2}\frac{N_t^2}{N_t^1} \rightarrow \frac{N_t^2}{N_t^1} = \mathcal{D}y_t^{\frac{1}{\mu}} \text{ and } \frac{K_t^1}{N_t^1} = \mathcal{D}\frac{K_t^2}{N_t^2} \text{ implying that } \frac{K_t^1}{K_t^2} = \mathcal{D}\frac{N_t^1}{N_t^2} = y_t^{\frac{-1}{\mu}}$$

From these equations it is clear that there is a higher capital to labour ratio in the low detection industry. In addition, these relationships imply that if $\theta_t^1 = \omega\theta_t^2$, for all t where ω is a constant, then $y_t = y$ and the employment in the different industries and the capital in the different industries are always proportional. However, if $\frac{\theta_t^2}{\theta_t^1}$ increases, $\frac{N_t^2}{N_t^1}$ and $\frac{K_t^2}{K_t^1}$ increase.

Proposition 4, describes the model's predictions about how a change in the detection rate in industry 1 will affect both the level of employment in industry 1, and the level of employment in industry 2 in steady state.

Proposition 4: A decrease in the detection rate in industry 1 relative to the detection rate in industry 2 causes employment in industry 1 to fall, and causes employment in industry 2 to rise in steady state.

III.4. Profit Rates

It also follows from the Euler equations that higher wages are paid by firms with higher profit rates where the profit rate is defined by profit per worker. Specifically,

$$\frac{\pi_t^1}{\pi_t^2}\frac{N_t^2}{N_t^1} = \frac{P_t^1Y_t^1 - r_tK_t^1 - w_t^1N_t^1}{P_t^2Y_t^2 - r_tK_t^2 - w_t^2N_t^2}\frac{N_t^2}{N_t^1} = \frac{P_t^1Y_t^1\left(1 - \frac{\alpha}{\phi} - \frac{1-\alpha}{\phi}\right)}{P_t^2Y_t^2\left(1 - \frac{\alpha}{\phi} - \frac{1-\alpha}{\phi}\right)}\frac{N_t^2}{N_t^1} = y_t^{-\frac{1}{\mu}}\mathcal{D}y_t^{\frac{1}{\mu}} = \mathcal{D} > 1$$

Therefore, the model predicts that firms in the low detection industry (high wage industry) will have higher profits per worker than firms in the high detection industry (low wage industry).

However, this does not imply that total profits are higher in the low detection industry. For example, if $\frac{\theta_t^2}{\theta_t^1} = 1$, in equilibrium:

$$y_t = y = \left[\frac{e^2}{e^1}\mathcal{D}\right]^{\frac{\mu(1-\alpha)}{\mu-1}} > 1$$

In this case, $\frac{\pi_t^1}{\pi_t^2} = y_t^{-\frac{1}{\mu}} < 1$ and profits would be higher in the industry with lower wages, (i.e., industry 2). However, if $\frac{\theta_t^2}{\theta_t^1} \neq 1$, it is possible that $\frac{\pi_t^1}{\pi_t^2} = y_t^{-\frac{1}{\mu}} > 1$ causing profits to be higher in the low detection industry. Consequently, which industry earns the higher profits will depend on the relative levels of technology in the industries.

III.5. Relative Prices

In the model, the relative prices of goods in the two industries are:

$$\frac{P_t^1}{P_t^2} = y_t^{\frac{\mu-1}{\mu}}$$

As y_t increases (i.e., $\frac{\theta_t^2}{\theta_t^1}$ increases), then the price of industry 1's output increases relative to the price of the output in industry 2.

III.6. The Industries' labour Forces

The previous results are derived under the condition that $x_t^j > N_t^j$. However, there has been no discussion about how x_t^i is determined. Here, the sizes of the industries' labour forces and the size of their job queues are examined under two different assumptions. In the first case, I assume that workers are free to look for work in both sectors simultaneously while in the second case I assume that workers can search in only one sector at a time.

Case 1: If individuals are free to look for work in both sectors simultaneously, all workers will look for jobs in both sectors since there is no cost to searching. However, workers will prefer to be employed in industry 1 rather than in industry 2, since wages are higher and effort is lower in industry 1. As a result, all workers would accept a job from a firm in industry 1, while only workers without a job offer from industry 1 would accept employment offers from industry 2. Therefore, when workers are free to look in both sectors simultaneously, industry 1 will have a larger pool of potential employees than firms in industry 2.

Case 2: If workers can only look for employment in one industry during the period, then $x_t^1 + x_t^2 = 1$. Further, an individual going to industry 1 must have the same expected utility as a worker seeking employment from industry 2 in equilibrium, since workers get to choose what industry they will search in. Consequently, the economy's equilibrium market clearing condition is:

$$\frac{N_t^1}{x_t^1}U(c_t^1, e_t^1) + \left(1 - \frac{N_t^1}{x_t^1}\right)U(c_t^u, 0) = \frac{N_t^2}{x_t^2}U(c_t^2, e_t^2) + \left(1 - \frac{N_t^2}{x_t^2}\right)U(c_t^u, 0)$$

$$\text{which implies } x_t^1 = \frac{f_0}{f_1\frac{N_t^2}{N_t^1} + f_0} \text{ and } x_t^2 = 1 - x_t^1$$

$$\text{where } f_0 = \ln\left(\frac{c_t^1}{c_t^u}\right) + \gamma\ln\left(\frac{T - e^1 - \xi}{T}\right) \text{ and } f_1 = \ln\left(\frac{c_t^2}{c_t^u}\right) + \gamma\ln\left(\frac{T - e^2 - \xi}{T}\right)$$

Changes in x_t^1 depend solely on changes in $\frac{N_t^2}{N_t^1}$ since effort levels in the different industries are constant across time and across states of the world, and:[25]

$$\ln\left(\frac{c_t^2}{c_t^u}\right) = \ln\left(\frac{c_t^2}{c_t^{s2}}\right) = \ln(\Pi_2) \text{ and } \ln\left(\frac{c_t^1}{c_t^u}\right) = \ln\left(\mathcal{D}\frac{\Pi_2 - 1}{1 - \frac{1}{\Pi_1}}\right)$$

As a result, if $\theta_t^1 = \omega\theta_t^2$ for all t, $x_t^1 = x^1$ and $x_t^2 = 1 - x_t^1 = 1 - x^1 = x^2$. Further, if $\frac{\theta_t^2}{\theta_t^1}$ increases, $\frac{N_t^2}{N_t^1}$ increases, which causes x_t^1 to decrease and x_t^2 to increase. Therefore, as jobs increase in one sector relative to another, more people look for work in the growing sector.

III.7. Queues

Past work on efficiency wages has suggested that there should be large queues in industries that pay higher wages. If we measure the queue size by the ratio of the number of workers

[25] $\frac{c_t^1}{c_t^u}$ is constant since $c_t^u = c_t^{s2}$, $\frac{c_t^1 - c_t^{s1}}{1-s} = w_t^1$, $\frac{c_t^2 - c_t^{s2}}{1-s} = w_t^2$ and $\frac{w_t^1}{w_t^2} = \mathcal{D}$ implies $\frac{c_t^1}{c_t^{s2}} = \mathcal{D}\frac{\Pi_2 - 1}{1 - \frac{1}{\Pi_1}}$.

willing to accept employment in that industry to the number of jobs available in the industry, it is clear that the assumptions concerning where individuals are able to search for work during the period will affect the findings.

In the first case, individuals could search for work in both sectors simultaneously. Since all workers would accept employment in the high wage industry, while only workers without job offers from the high wage industry would accept employment in the low wage industry, it follows that $x_t^1 > x_t^2$. As a result, whether the queue is larger in the high wage industry than in the low wage industry, $\left(\text{i.e., whether } \frac{x_t^1}{N_t^1} > \frac{x_t^2}{N_t^2}\right)$, will depend on the relative size of the number of jobs available in each industry. For example, if $N_t^1 < N_t^2$, as it is in the data, then clearly $\frac{x_t^1}{N_t^1} > \frac{x_t^2}{N_t^2}$. However, if the number of jobs in the high wage industry was much larger than the number of jobs in the low wage industry, then $\frac{x_t^1}{N_t^1} < \frac{x_t^2}{N_t^2}$. In general, whether or not $N_t^1 < N_t^2$ depends on the parameters of the model.

Although it is not possible to prove that the size of the total number of workers queuing for jobs is always greater in the high wage industry when workers are able to search for jobs in both industries simultaneously, it is possible to prove that $\frac{x_t^1}{N_t^1} > \frac{x_t^2}{N_t^2}$ when workers are only able to search in one industry at a time. In particular,

$$\frac{x_t^1}{N_t^1} - \frac{x_t^2}{N_t^2} > 0 \iff \frac{x_t^1}{N_t^1}\frac{N_t^2}{x_t^2} > 1 \iff \frac{N_t^2}{N_t^1}\left(\frac{f_0}{f_1 \frac{N_t^2}{N_t^1}}\right) > 1 \iff f_0 > f_1$$

However, $f_0 > f_1$ since wages are higher and effort is lower in the low detection industry. Therefore, if workers only search in one sector of the economy at a time, the model predicts that queues will always be larger in the high wage industry.

III.8. Differences in Unmeasured Ability

An alternate theory of inter-industry wage differentials proposes that workers with the same measurable attributes receive different wages because there are systematic differences in unmeasured worker ability that are correlated with industry affiliation. Papers, such as Murphy and Topel (1987), have argued that unobserved worker ability, not efficiency wage considerations, are responsible for the majority of the wage differentials seen in the data. Therefore, it is important to determine how the addition of ability affects the results of the efficiency wage model.

To explore the consequences of adding workers with different ability to the model, I consider a simple case where there are only two types of workers in the economy, and explore the properties of the model for the equilibrium where the high ability individuals are employed in the high paying industry. Type 1 agents are assumed to have higher ability than type 2 agents in Industry 1 (the high wage industry), while both agents have the same ability in Industry 2. Type 2 agents are assumed to have the same ability in industry 1 and 2. Furthermore, I assume that there are a large number of agents of both types so there is no shortage of either type of worker. I consider two ways of incorporating unobserved ability into the model.

In the first case, I assume that the worker's ability affects the amount of effective labour that is provided to firms. In particular, if a_i^j is the level of individual $i's$ ability in industry j, then for a given level of effort in industry j, e_i^j, this individual contributes $a_i^j e_i^j$ units of effective labour to his firm. Since type 1 individuals have higher ability in industry 1, for a given level of effort, $a_1^1 e > a_2^1 e$. However, in industry 2, for a given level of effort and hours, $a_1^2 e = a_2^2 e$. All workers, in this case, will have the same incentive compatibility constraints as before because a worker's ability does not directly enter into the worker's utility function.

Firms in industry j now have the following production function:

$$Y_t^j = \theta_t^j \left(K_t^j\right)^\alpha (N_{1t}^j a_1^j e_{1t}^j + N_{2t}^j a_2^j e_{2t}^j)^{1-\alpha}$$

where N_{it}^j is the number of workers of type i hired by the firm in industry j.

It is straightforward to demonstrate that when there is no shortage of workers, firms in industry 1 will choose to hire only one type of worker in equilibrium (the high ability type) while firms in industry 2 will hire either type of worker. In the sorting equilibrium, where type 1 workers are hired in industry 1, the high ability of the type 1 individuals acts as an increase in the level of technology used in industry 1 and causes more type 1 workers to be hired into that industry. Therefore, the model is consistent with the empirical observation that high ability individuals are generally found in the high wage sector. However, the model also demonstrates that wages and effort levels will still be determined by the Solow condition, and the inter-industry wage differential remains as it was before, $\left(\text{i.e., } \frac{w_t^1}{w_t^2} = \frac{(1-s\Pi_2)(\Pi_1-1)}{(1-s\Pi_1)(\Pi_2-1)} = \mathcal{D}\right)$. This implies that high ability workers are hired into the high wage firms because their ability increases the profitability of the firm, but unlike labour-market clearing models the value of the wage differential and the wages paid to individuals in the shirking model are not dependant on the ability of workers. The inter-industry wage differential remains solely determined by the parameter values $\xi, s, \frac{d_1}{\gamma}, \frac{d_2}{\gamma}$, and T.

In the second case, I assume that an individual's ability affects his utility from working by altering the fixed cost associated with working. Specifically, individuals with higher ability will have lower costs associated with providing effort on the job:

$$\begin{aligned} \xi_i^1 &= \xi - h(a_i^1) \\ \xi_i^2 &= \xi \end{aligned}$$

where $h(\cdot) > 0$, and $h'(\cdot) > 0$ for all i. Therefore, type 1 workers incur lower fixed costs in industry 1 than type 2 workers, while the fixed cost is independent of type in industry 2.

Firms in industry j face the following problem when there is no shortage of workers:

$$\max_{\left\{K_t^j, N_{1t}^j, w_{1t}^j,, N_{2t}^j, w_{2t}^j, P_t^j\right\}} P_t^j Y_t^j - r_t^j K_t^j - w_{1t}^j N_{1t}^j - w_{2t}^j N_{2t}^j$$

subject to:

$$Y_t^j = \theta_t^j (K_t^j)^\alpha (N_{1t}^j e_{1t}^j + N_{2t}^j e_{2t}^j)^{1-\alpha}$$

$$e_{it}^j = T\left(1 - \left(\frac{c_{it}^j}{c_{it}^{sj}}\right)^{-\frac{d_j}{\gamma}}\right) - \xi_i^j \text{ for } i = 1, 2$$

$$Y_t^j = Y_t(\overline{P}_t^j)^{\frac{\mu-\phi}{(\mu-1)(\phi-1)}} P_t^{j\,\frac{-\phi}{\phi-1}}$$

Here N_{it}^j is the number of workers of type i hired by the firm in industry j, e_{it}^j is the effort required from a worker of type i working for a firm in industry j, ξ_i^j is type $i's$ fixed cost of providing effort in industry j, c_{it}^j is the consumption enjoyed by a type i worker who is not detected shirking, and c_{it}^{sj} is the consumption enjoyed by a type i worker who is detected shirking. Again it is straightforward to show that the firms in industry 1 will only hire one type of worker, while firms in industry 2 are again indifferent between the different types. Therefore, the model will again be consistent with the empirical observation that high ability workers are more frequently found in high paying industries. However, the model again suggests that differences in workers' ability may not be a major determinant of large inter-industry wage differentials.

In the sorting equilibrium, where type 1 workers are hired by firms in industry 1, type 1 workers (i.e., the high ability workers) employed in industry 1 receive lower wages than type

2 workers would receive *if* they firms in industry 1 had to hire them. This result follows from the assumption that high ability workers have a lower fixed cost of providing effort in this industry. Firms know that it is not as costly to induce high ability types to provide effort as low ability types because of the difference in their fixed costs. Therefore, the high ability workers will not require as high a wage as low ability workers for a given level of effort in industry 1. Since ability puts downward pressure on wages in the high paying industry in this model, the addition of ability it makes it harder, not easier, to explain sizable inter-industry wage differentials. In particular, as the fixed cost of providing effort decreases for the high ability workers in industry 1 (i.e., as ability in industry 1 rises relative to ability in industry 2), Π_1 decreases, which causes the wage differential, $\frac{w_t^1}{w_t^2} = \frac{(1-s\Pi_2)(\Pi_1-1)}{(1-s\Pi_1)(\Pi_2-1)}$, to fall.

The two methods of modelling unobserved ability examined suggest that, even though differences in ability has been proposed as a major determinant of inter-industry wage differentials, the presence of workers with different abilities may not help explain high inter-industry wage differentials when there is imperfectly observable effort. However, the model with workers of different abilities across industries does provide an explanation for the observation that high ability workers are more often employed in high wage industries. The model's results are consistent with Gibbons and Katz's (1992) empirical findings that: (1) if workers' traits had any direct effect on wages, the effect is small, and (2) workers' traits affect workers' mobility.

IV. AN EMPIRICAL ANALYSIS

IV.1. Estimating the Inter-Industry Wage Differentials

Although previous studies, such as Krueger and Summers (1988), report that the inter-industry wage differentials are highly persistent over time, little information exists regarding how inter-industry wage differentials behave over the business cycle.[26] This additional information can help test whether a multi-sector model's predictions are consistent with the empirical evidence.

For the empirical analysis of inter-industry wage differentials in the paper, data from the yearly May Current Population Surveys from 1973 to 1999 are used.[27] The May supplement is chosen for two reasons. First, it contains pertinent information about labour force participants who are 14 years of age or older. Second, only the May survey contains information on weekly earnings in the earlier years of the sample.

This section adopts the same general strategy as Krueger and Summers (1988). Specifically, the sample that is analyzed contains full and part-time non-agricultural workers, who are at least 16 years of age and who report usually earning between \$1.00 and \$250.00 an hour.[28]

To determine the importance of industry affiliation in determining usual hourly earnings, the log hourly wage is regressed on the one-digit census dummy variables, and human capital and demographic controls. The human capital and demographic controls consist of 9 occupation variables, a measure of education, an age variable, sex, race, union status, veteran status, a central city dummy variable, marital status, and a number of interaction terms.[29] The

[26] Keane (1993) also examines the cyclicality of inter-industry wage differentials using a different approach and data from the National Longitudinal Study of Young Men from 1966-1981.

[27] Although it would be useful to have data over a longer period of time, there have been significant changes in the occupation and industry codes and significant changes in the way that usual earnings and hours are reported. As a result, it is difficult to expand the analysis to include data from the years prior to 1973 since the changes make it very difficult to make meaningful comparisons between pre-1973 estimates and the estimates from 1973 on.

[28] Usual hourly earnings are defined as usual weekly earnings divided by usual weekly hours. The upper bound reflects topcoding in the data as well as an assumption made about feasible earnings for the categories of workers we examine. For example, a full time employee at 40 hours a week who earned \$250 an hour could earn as much as \$520,000. Therefore, eliminating people above this level will simply remove some of the outliers in the sample.

[29] The interaction terms include: (age)X(sex), (martial status)X(sex), (education)X(sex) and (education^2)X(sex). The union variable was unavailable in 1982. However, sensitivity analysis suggests that

yearly estimates of the inter-industry wage differentials are obtained by estimating the wage equation:

$$\ln(w) = \alpha + \beta X + D\gamma + \Omega O + \varepsilon$$

where w is the hourly wage of an individual, γ is a vector of mutually exclusive industry dummy variables, X is a vector of individual characteristics and locational variables, and O is a vector of mutually exclusive occupation dummy variables.[30] To facilitate a comparison between yearly estimates, an attempt is made to keep the definitions of the variables used in the regressions as consistent as possible over the years since there are a number of changes in the available data.[31]

The employment weighted average of the wage differentials, D, is computed. These inter-industry differentials, $\mathfrak{D}$, are differences between the industry differentials and the weighted average and are reported in Table 1.[32] Here, the resulting statistics correspond to the percentage increase (or decrease) in wages between an average worker in the specific industry and the average wage of a worker in all industries after controlling for demographic and human

the results do not critically depend on this variable. A more detailed description of the variables is available from the author upon request.

[30] Since the regression includes a constant, the omitted industry variable, (i.e., mining), is treated as having a zero effect on wages, following Krueger and Summers' (1988) methodology.

[31] One example of such as change is found in the way the worker's education is reported. Prior to 1991, the CPS reports a measure of education that can determine the number of years of schooling that the individual has completed. In 1991, the measure of education was changed. The new measure's categories report whether an individual has completed a high-school degree, college degree, etc. As a result, these two measures of education are not directly comparable. The solution adopted in the paper is to define a new measure of education as in Jaeger (1997) which is more consistent across the time periods in question.

[32] Following Krueger and Summers (1987), the differentials are computed as follows for industries j=0,1,...6 where s_j is the share of employment in industry j

$$\mathfrak{D}_j = D_j - \sum_{j=0}^{6} s_j D_j = D_j - \sum_{j=1}^{6} s_j D_j$$

using the assumption that $D_0 = 0$ for mining.

capital variables. Figure I graphs the inter-industry wage differential and the unemployment rate to illustrate the behavior of the differentials over time and over the business cycle. The findings suggest that the mining industry generally pays workers the highest wages, the service industry always has the second lowest wages and the wholesale and retail trade industry consistently has the lowest wages. However, the rankings of the manufacturing industry (MANUF), the construction industry (CONSTR), the transportation industry (TRANSP) and the finance, insurance and real estate industry (FINANCE), are not as clear. This is apparent in Table 2, where the average value of the differentials, as well as the average rank of the differential for each industry, are reported.[33] These statistics suggest that workers in the construction and transportation industries generally have higher wages than workers in either the finance, insurance, and real estate industry or the manufacturing industry. However, no conclusive statement can be made concerning ranking wages in the transportation industry versus wages in the construction industry, or wages in the finance, insurance and real estate industry versus the wages in the manufacturing industry.

Although it appears from Figure I that the differentials are not cyclical over the time period, it is useful to check this formally. This is done by running the following regression:

$$\Delta\widehat{\mathfrak{D}}_t^i = \alpha + \beta\Delta u_t + \gamma t + v_t$$

where v_t is white noise, t is a time trend, and $\Delta\widehat{\mathfrak{D}}_t^i$ and Δu_t are the first differences of the estimated interindustry wage differentials and the unemployment rate respectively.[34] The

[33] To calculate the average rank of the differentials, the industries in each year were ranked according to their wages. Each industry was assigned a value that corresponded to its ranking, (i.e., the industry with the highest wages was assigned a value of 1, the industry with the second highest wages received the value 2, etc.). Next, the average of the ranking for each industry over the years was computed.

[34] The unemployment rate that is used is the yearly average unemployment rate from May to May. This data was chosen to coincide with the dates of the CPS data.

results of this regression are found in Table 3. The results support the hypothesis that there is no significant cyclicality in the inter-industry wage differentials.[35]

The model in Section II predicts that wages in high wage industries have a higher variance than hourly wages in low wage industries. To examine the validity of this prediction, the variance of HP filtered hourly wages for the one-digit industries are compared.[36] The results are also reported in Table 2. These results suggest that hourly wages in high wage industries are likely to be more variable than the hourly wages in low wage industries.

IV.2. A Calibrated Model with 7 Industries

To analyze whether the model is consistent with the empirical findings, and to determine if small differences in the rates of detection across industries can produce sizable inter-industry wage differentials, a version of the model with 7 industries is calibrated. The calibration exercise is discussed below.

In the model, the percentage difference between the wage in industry i, i=1,..7, and the average wage when there are 7 industries in the economy is:

$$\begin{aligned}\frac{w^i}{\overline{w}} &= \frac{w^i}{\left(\sum_{i=1}^{7} w^i N^i\right) / \left(\sum_{i=1}^{7} N^i\right)} = \frac{w^i\left(\sum_{i=1}^{7} N^i\right)}{\left(\sum_{i=1}^{7} w^i N^i\right)} \\ &= \frac{w^7 \mathcal{D}^i N^7 \left(\sum_{i=1}^{7} \frac{N^i}{N^7}\right)}{w^7 N^7 \left(\sum_{i=1}^{7} \left(\mathcal{D}^i \frac{N^i}{N^7}\right)\right)} = \frac{\mathcal{D}^i \left(\sum_{i=1}^{7} \frac{N^i}{N^7}\right)}{\left(\sum_{i=1}^{7} \left(\mathcal{D}^i \frac{N^i}{N^7}\right)\right)}\end{aligned}$$

where $\mathcal{D}^i$ is the constant differential, $\frac{w^i}{w^7}$, implied by the model when w^7 is the wage in the industry with the highest rate of detection, (i.e., the lowest wage). Since the differentials $\{\mathcal{D}^1, \mathcal{D}^2, ..., \mathcal{D}^7\}$ are independent of the time period and state of the world, the behavior of $\frac{w^i}{\overline{w}}$

[35] This finding is consistent with Keane's (1993) results.

[36] Here, the hourly wage data from the citibase dataset is used. This data is created by Citibase on a monthly basis using the data from the BLS's CPS dataset. Since the data is monthly, the HP filter's smoothing parameter is set equal to 129600 as suggested by Ravn and Uhlig (1997).

depends on the behavior of $\frac{N^i}{N^7}$. From the results in Section III, it is clear that $\frac{w^i}{w}$ is acyclical whenever θ_t^1 is proportional to θ_t^2. Therefore, the 7 industry model is consistent with acyclical differentials if the technology levels in the different industries are directly proportional to one another.

Given these results, and acyclicality of the estimated inter-industry wage differentials, the model is calibrated under the assumption that $\theta^i = \omega_i \theta^7$. Furthermore, it is assumed that the technology level in industry i is determined by the equation:

$$\left(\theta_{t+1}^i - \theta^{iss}\right) = \rho_\theta \left(\theta_t^i - \theta^{iss}\right) + \varepsilon_t$$

where θ^{iss} is the steady state level of the technology in industry i, $0 < \rho_\theta < 1$ and ε is a mean zero i.i.d. shock. For this case the parameter values chosen are:

ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ρ_θ	β	α	$\mu = \phi$	δ
0.85	1.13	1.10	1.06	1.37	1.14	1.00	0.9	$\left(\frac{1}{1.03}\right)^{.25}$	0.34	1.2	0.025

$\frac{d_1}{\gamma}$	$\frac{d_2}{\gamma}$	$\frac{d_3}{\gamma}$	$\frac{d_4}{\gamma}$	$\frac{d_5}{\gamma}$	$\frac{d_6}{\gamma}$	$\frac{d_7}{\gamma}$	T	ξ	s
0.4795	0.4950	0.4950	0.5052	0.5045	0.5382	0.5625	1369	0.95	0.95

These values are chosen so that: *(i)* the aggregate unemployment rate is approximately 6.0%, *(ii)* the percentage of individuals employed in each industry is in the range of the employment shares seen in the CPS data set, *(iii)* the average wage differentials in the data are equal to the differentials predicted by the model, and *(iv)* the values for α, β, δ, T, and the markups are consistent with values seen in the real business cycle literature.[37] Here the notation implies that industry 1 has the highest wage and industry 7 has the lowest wage. For industries 2 to 6, the employment shares in the model are matched to the industry data for the

[37] See e.g., Burnside and Eichenbaum (1996), and Alexopoulos (2000).

industry with the same rank based on the rankings suggested by the average rank of the industry's differential. Table IV and Table V report the model's steady state values of employment shares and the differentials for these parameter values. These results indicate that small differences between the industries' probability of detecting shirkers, $\left(\text{i.e., } \left|\frac{d_1 - d_7}{d_7}\right| \times 100\% \simeq 17\%\right)$, can reproduce the substantial differences in wages across industries.[38] ,[39]

In general, there is debate in the literature about how much of the differentials are due to efficiency wage considerations. Simulations of the model suggest that a difference in detection rates of less than 5% between the highest and lowest paid industries is required if it is only necessary to explain one third of the differentials, which is the estimate of the unexplained portion of inter-industry wage differentials in Murphy and Topel (1987). Given that our measures of monitoring, such as the ratio of supervisors to employees, are generally very noisy, these results suggest that the small differences seen in these measures are not necessarily evidence that efficiency wage considerations are unimportant for explaining inter-industry wage differentials.

IV.3. Response to an Aggregate Technology Shock

In addition to evaluating the model's ability to predict the value of the predicted inter-industry wage differentials, it is interesting to determine if the model is able to generate small wage movements alongside large employment changes in response to technology shocks, since the aggregate evidence on wages and employment suggests that the real wage is much less variable over the business cycle than employment. Only the aggregate responses are reported

[38] To change this computation into levels, a value of γ must be choosen. For example, when $\gamma = 1.75$, the difference between the detection rates in the highest and lowest paid industries would be approximately 14%.

[39] Using a model where firms employ workers in different occupations, Alexopoulos (2002) demonstrates that large inter-industry wage differentials can emerge even when there are small differences in detection rates for only occupational group across industries.

in Figure II since the responses for the industry level variables are the same as the aggregate level responses in this model. It is clear from the results that the multi-sector model is able to account for small wage responses and large employment responses following a positive technology shock. For example, the model suggests that, for the parameter values used, the percent deviation of employment from its steady state value in response to a technology shock is at least twice the percent deviation of wages from its steady state value.

V. CONCLUSIONS

The paper presents a new multi-industry shirking efficiency wage model and investigates if this type of model is qualitatively consistent with *(i)* the behavior of the inter-industry wage differentials observed over time in the data, and *(ii)* the observations that high wage industries also tend to have higher profits per worker and higher levels of capital per worker. All firms, in the model, imperfectly observe their workers' effort levels in production. However, the firms' ability to detect shirking workers varies across sectors. In this environment, detected shirkers are assumed to forgo a portion of their compensation, such as a bonus. Absent bonding, these assumptions can generate equilibrium unemployment, and the different detection rates cause firms to choose different optimal levels of effort, wages and employment.

Three main results emerge from analyzing the model. First, the model with monetary punishments is able to reproduce the stability of the inter-industry wage differentials observed in the economy. Second, small differences in the probability of detection across industries can generate substantial inter-industry wage differentials in the model. Third, the model is consistent with the empirical evidence that implies that high wage industries have higher capital to labour ratios and higher profits per worker.

In addition to these results, the model predicts that workers in industries with higher detection rates will have lower wages and provide higher effort levels, than their counterparts in low detection industries. The model's predictions about the size of the queues in different industries is also examined. The results suggest that the model can predict larger queues for jobs in high paying industries, however these results depend on the parameter values and the assumptions concerning where individuals can look for work during the period.

This paper also provides new evidence about how inter-industry wage differentials behave over the business cycle using Krueger and Summers' (1988) methodology and the May Current Population Survey from 1973-1999. The findings confirm the results seen in previous studies which suggest that inter-industry wage differentials are very persistent over time, and are acyclical.

Finally, the paper examines the model's responses to technology shocks to determine whether the responses are consistent with large employment responses and small wage responses. The results demonstrate that the calibrated version of the model predicts that real wages in all industries are weakly procyclical, employment in all is strongly procyclical, and the inter-industry wage differentials are acyclical.

The work in this paper provides additional support for the theory that shirking efficiency wage frictions can help explain a significant part of inter-industry wage differentials, as well as other labour market phenomena, such as the presence of unemployment and the observed behavior of wages and employment over the business cycle. Although the paper does not rule out other potential explanations of inter-industry wage differentials, it does illustrate that differences in unmeasured worker ability may not help explain sizable inter-industry wage differentials in economies where firms face problems detecting shirking workers. Given that a number of empirical studies have documented relationships between monitoring, effort and wages in many industries, this theoretical finding may help explain the results in Blackburn and Neumark (1992) and Krueger and Summers (1988) which suggest that correcting for individuals' ability levels do not significantly reduce the size of the estimated inter-industry wage differentials. Future work should concentrate on designing alternative tests to clarify the importance of these efficiency wage considerations in determining wages and inter-industry wage differentials.

APPENDIX A: Proofs of Propositions

Proof of Proposition 1

Let $\Pi_j = \frac{c_t^j}{c_t^{sj}}$. Then the value of Π_j is determined by the Solow condition

$$0 = T\frac{d_j}{\gamma}(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{(1-s\Pi_j)}{1-s}(\Pi_j - 1) - T\left(1-(\Pi_j)^{-\frac{d_j}{\gamma}}\right) + \xi \equiv H$$

Notice that this equation, and the condition that effort is positive, implies that $\Pi_j \in (1, \frac{1}{s})$.

Next use the implicit function theorem to get $\frac{\partial \Pi_j}{\partial \frac{d_j}{\gamma}}$.

$$\begin{aligned}
H_{\Pi_j} &= T\frac{d_j}{\gamma}\left[-\left(\frac{d_j}{\gamma}+1\right)\right](\Pi_j)^{-\frac{d_j}{\gamma}-2}\frac{(1-s\Pi_j)}{1-s}(\Pi_j-1) - T\frac{d_j}{\gamma}(\Pi_j)^{-\frac{d_j}{\gamma}-1} \\
&\quad -T\frac{d_j}{\gamma}(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{s}{1-s}(\Pi_j-1) + T\frac{d_j}{\gamma}(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{(1-s\Pi_j)}{1-s} \\
&= T\frac{d_j}{\gamma}\left[-\left(\frac{d_j}{\gamma}+1\right)\right](\Pi_j)^{-\frac{d_j}{\gamma}-2}\frac{(1-s\Pi_j)}{1-s}(\Pi_j-1) \\
&\quad +T\frac{d_j}{\gamma}(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{1}{1-s}\{-s(\Pi_j-1)+1-s\Pi_j-(1-s)\} \\
&= T\frac{d_j}{\gamma}\left[-\left(\frac{d_j}{\gamma}+1\right)\right](\Pi_j)^{-\frac{d_j}{\gamma}-2}\frac{(1-s\Pi_j)}{1-s}(\Pi_j-1) \\
&\quad +T\frac{d_j}{\gamma}(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{1}{1-s}\{-s\Pi_j+s+1-s\Pi_j-1+s)\} \\
&= T\frac{d_j}{\gamma}\left[-\left(\frac{d_j}{\gamma}+1\right)\right](\Pi_j)^{-\frac{d_j}{\gamma}-2}\frac{(1-s\Pi_j)}{1-s}(\Pi_j-1) + T\frac{d_j}{\gamma}(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{1}{1-s}\{2s(1-\Pi_j)\}
\end{aligned}$$

Therefore, $H_{\Pi_j} < 0$ since $\Pi_j > 1$. Next, the sign of $H_{\frac{d_j}{\gamma}}$ needs to be determined.

$$H_{\frac{d_j}{\gamma}} = T(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{(1-s\Pi_j)}{1-s}(\Pi_j-1) - T\frac{d_j}{\gamma}(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{(1-s\Pi_j)}{1-s}(\Pi_j-1)\ln(\Pi_j) - \ln(\Pi_j)T(\Pi_j)^{-\frac{d_j}{\gamma}}$$

Remembering that the definition of H implies that

$$-\ln(\Pi_j)T(\Pi_j)^{-\frac{d_j}{\gamma}} = \ln(\Pi_j)\left[T\frac{d_j}{\gamma}(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{(1-s\Pi_j)}{1-s}(\Pi_j-1) - T + \xi\right]$$

and substituting this expression into the previous equation produces the expression:

$$
\begin{aligned}
H_{\frac{d_j}{\gamma}} &= T(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{(1-s\Pi_j)}{1-s}(\Pi_j-1)\left[1-\ln(\Pi_j)\frac{d_j}{\gamma}+\ln(\Pi_j)\frac{d_j}{\gamma}\right]-\ln(\Pi_j)\left[T-\xi\right] \\
&= T(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{(1-s\Pi_j)}{1-s}(\Pi_j-1)-\ln(\Pi_j)\left[T-\xi\right] \\
&= \frac{T\left(1-(\Pi_j)^{-\frac{d_j}{\gamma}}\right)-\xi}{\frac{d_j}{\gamma}}-\ln(\Pi_j)\left[T-\xi\right] \text{ (also from the definition of H)}
\end{aligned}
$$

Clearly if Π_j is close to 1, $H_{\frac{d_j}{\gamma}} < 0$. However, this is more clearly shown by defining a function $J(\Pi_j)$, where

$$
J(\Pi_j) = \frac{T\left(1-(\Pi_j)^{-\frac{d_j}{\gamma}}\right)-\xi}{\frac{d_j}{\gamma}}-\ln(\Pi_j)\left[T-\xi\right]
$$

$J(\Pi_j)$ is a strictly decreasing function when $\Pi_j \in (1, \frac{1}{s})$. i.e.

$$
\begin{aligned}
J'(\Pi_j) &= T(\Pi_j)^{-\frac{d_j}{\gamma}-1}-\frac{T-\xi}{\Pi_j}=\frac{1}{\Pi_j}\left[-(T-\xi)+T\Pi_j^{-\frac{d_j}{\gamma}}\right] \\
&= \frac{1}{\Pi_j}\left[-(T-\xi)-T\frac{d_j}{\gamma}(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{(1-s\Pi_j)}{1-s}(\Pi_j-1)+(T-\xi)\right] \text{ (From H=0)} \\
&= \frac{1}{\Pi_j}\left[-T\frac{d_j}{\gamma}(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{(1-s\Pi_j)}{1-s}(\Pi_j-1)\right] < 0
\end{aligned}
$$

It then follows that $H_{\frac{d_j}{\gamma}} < \max J(\Pi_j) = J(1) < 0$. As a result, the Implicit Function Theorem implies that $\frac{\partial \Pi_j}{\partial \frac{d_j}{\gamma}} = -\frac{H_{\frac{d_j}{\gamma}}}{H_{\Pi_j}} < 0$, and since d_j only appears in the equations in the ratio $\frac{d_j}{\gamma}$, it follows that

$$
\frac{\partial \Pi_j}{\partial d_j} = \frac{\partial \Pi_j}{\partial \frac{d_j}{\gamma}}\frac{\partial \frac{d_j}{\gamma}}{\partial d_j} = \frac{\partial \Pi_j}{\partial \frac{d_j}{\gamma}}\frac{1}{\gamma} < 0
$$

Proof of Proposition 2

$\frac{w_t^1}{w_t^2} = \mathcal{D} = \frac{(1-s\Pi_2)(\Pi_1-1)}{(1-s\Pi_1)(\Pi_2-1)}$. Also we know that $d_1 < d_2$ implies that $\Pi_1 > \Pi_2$ according to the previous proposition. To prove that $\mathcal{D} > 1$, examine the function $M(z) = \frac{z-1}{1-sz}$. Notice $M'(z) = \frac{1-s}{(1-sz)^2} > 0$. Therefore $\Pi_1 > \Pi_2 \rightarrow H(\Pi_1) > H(\Pi_2) \rightarrow \mathcal{D} = \frac{H(\Pi_1)}{H(\Pi_2)} > 1$

Proof of Proposition 3

$$\frac{\partial e^i}{\partial \frac{d_j}{\gamma}} = T\ln(\Pi_j)(\Pi_j)^{-\frac{d_j}{\gamma}} + T\frac{d_j}{\gamma}(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{\partial \Pi_j}{\partial \frac{d_j}{\gamma}} = T(\Pi_j)^{-\frac{d_j}{\gamma}-1}\left[\Pi_j \ln(\Pi_j) + \frac{d_j}{\gamma}\frac{\partial \Pi_j}{\partial \frac{d_j}{\gamma}}\right]$$

To sign $\left[\Pi_j \ln(\Pi_j) + \frac{d_j}{\gamma}\frac{\partial \Pi_j}{\partial \frac{d_j}{\gamma}}\right]$ it is useful to break the expression into parts. First

$$\begin{aligned}
\frac{d_j}{\gamma}\frac{\partial \Pi_j}{\partial \frac{d_j}{\gamma}} &= -\frac{d_j}{\gamma}\frac{T(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{1}{1-s}\left\{(1-s\Pi_j)(\Pi_j-1)\left(1-\frac{d_j}{\gamma}\ln(\Pi_j)\right) - \Pi_j\ln(\Pi_j)(1-s)\right\}}{T(\Pi_j)^{-\frac{d_j}{\gamma}-1}\frac{1}{1-s}\frac{d_j}{\gamma}\left\{\frac{-\left(\frac{d_j}{\gamma}+1\right)}{\Pi_j}(1-s\Pi_j)(\Pi_j-1)+2s(1-\Pi_j)\right\}} \\
&= \frac{-\left\{(1-s\Pi_j)(\Pi_j-1)\left(1-\frac{d_j}{\gamma}\ln(\Pi_j)\right) - \Pi_j\ln(\Pi_j)(1-s)\right\}}{\left\{\frac{-\left(\frac{d_j}{\gamma}+1\right)}{\Pi_j}(1-s\Pi_j)(\Pi_j-1)+2s(1-\Pi_j)\right\}} \equiv \frac{A}{B}
\end{aligned}$$

Then let $C \equiv B\Pi_j \ln(\Pi_j)$. In this case, the sign of $\frac{\partial e^i}{\partial \frac{d_j}{\gamma}}$ is determined by the sign of $\frac{A+C}{B}$. Since $\Pi_j \in (1, \frac{1}{s})$, $B < 0$. Therefore, the sign of $\frac{\partial e^i}{\partial \frac{d_j}{\gamma}}$ is determined by the sign of $A + C$.

$$\begin{aligned}
A+C &= -(1-s\Pi_j)(\Pi_j-1)\left(1-\frac{d_j}{\gamma}\ln(\Pi_j)\right) + \Pi_j\ln(\Pi_j)(1-s) \\
&\quad +\Pi_j\ln(\Pi_j)\left\{\frac{-\left(\frac{d_j}{\gamma}+1\right)}{\Pi_j}(1-s\Pi_j)(\Pi_j-1)+2s(1-\Pi_j)\right\} \\
&= -(1-s\Pi_j)(\Pi_j-1)\left(1-\frac{d_j}{\gamma}\ln(\Pi_j)\right) + \Pi_j\ln(\Pi_j)(1-s) \\
&\quad +\ln(\Pi_j)\left\{-\left(\frac{d_j}{\gamma}+1\right)(1-s\Pi_j)(\Pi_j-1)+2s\Pi_j(1-\Pi_j)\right\} \\
&= -(1-s\Pi_j)(\Pi_j-1)(1-\frac{d_j}{\gamma}\ln(\Pi_j)) + \Pi_j\ln(\Pi_j)(1-s) \\
&\quad -\left(\frac{d_j}{\gamma}+1\right)(1-s\Pi_j)(\Pi_j-1)\ln(\Pi_j) - \ln(\Pi_j)2s\Pi_j(\Pi_j-1) \\
&= -(1-s\Pi_j)(\Pi_j-1)\left[1-\frac{d_j}{\gamma}\ln(\Pi_j)+\left(\frac{d_j}{\gamma}+1\right)\ln(\Pi_j)\right] + \Pi_j\ln(\Pi_j)\left[(1-s)-2s\Pi_j+2s\right] \\
&= -(1-s\Pi_j)(\Pi_j-1)\left[1+\ln(\Pi_j)\right] + \Pi_j\ln(\Pi_j)\left[(1+s-2s\Pi_j\right] \\
&= -\ln(\Pi_j)\left[(1-s\Pi_j)(\Pi_j-1) - \Pi_j(1-s\Pi_j) + \Pi_j s(\Pi_j-1)\right] - (1-s\Pi_j)(\Pi_j-1)
\end{aligned}$$

$$= \quad -\ln(\Pi_j)\left[(1-s\Pi_j)\Pi_j-(1-s\Pi_j)-\Pi_j(1-s\Pi_j)+\Pi_j s(\Pi_j-1)\right]-(1-s\Pi_j)(\Pi_j-1)$$

$$= \quad -\ln(\Pi_j)\left[-1+s\Pi_j-\Pi_j s+s(\Pi_j)^2\right]-(1-s\Pi_j)(\Pi_j-1)$$

$$= \quad -\ln(\Pi_j)\left[s(\Pi_j)^2-1\right]-(1-s\Pi_j)(\Pi_j-1)$$

If $[s(\Pi_j)^2-1]>0$, then clearly $A+C<0$. However, it is necessary to consider the case where $[s(\Pi_j)^2-1]<0$ and $\Pi_j\in(1,\frac{1}{s})$.Note that $\Pi_j>1\Leftrightarrow(\Pi_j)^2>\Pi_j\Leftrightarrow -s(\Pi_j)^2< -s\Pi_j\Leftrightarrow 1-s(\Pi_j)^2<1-s\Pi_j\Leftrightarrow \ln(\Pi_j)(1-s(\Pi_j)^2)<\ln(\Pi_j)(1-s\Pi_j)$. Therefore, under these conditions,

$$\begin{aligned} A+C &= \ln(\Pi_j)\left[1-s(\Pi_j)^2\right]-(1-s\Pi_j)(\Pi_j-1) \\ &< \ln(\Pi_j)\left[1-s\Pi_j\right]-(1-s\Pi_j)(\Pi_j-1)=-(1-s\Pi_j)\left[\Pi_j-1-\ln(\Pi_j)\right] \end{aligned}$$

Let $W(\Pi_j)\equiv\Pi_j-1-\ln(\Pi_j)$. Then $W'(\Pi_j)=1-\frac{1}{\Pi_j}>0$ for $\Pi_j\in(1,\frac{1}{s})$, and $W(1)=0$. As a result, $W(\Pi_j)>W(1)=0$ for $\Pi_j\in(1,\frac{1}{s})$. Therefore, $A+C<0$ if $[s(\Pi_j)^2-1]<0$. More importantly, since all cases imply that $A+C<0$, $\frac{A+C}{B}>0\rightarrow\frac{\partial e^i}{\partial\frac{d_j}{\gamma}}>0$. Since d_j only appears in the equations in the ratio $\frac{d_j}{\gamma}$, it follows that

$$\frac{\partial e^i}{\partial d_j}=\frac{\partial e^i}{\partial\frac{d_j}{\gamma}}\frac{\partial\frac{d_j}{\gamma}}{\partial d_j}=\frac{\partial e^i}{\partial\frac{d_j}{\gamma}}\frac{1}{\gamma}>0$$

Therefore, $d_1<d_2\rightarrow e^1<e^2$.

Proof of Proposition 4

Here the variables without time subscripts denote the steady state levels of the variables.

$$\frac{\partial N_1}{\partial\frac{d_1}{\gamma}}=\frac{\partial\left[\frac{(1-\alpha)(1-s)r}{\left[(1-\alpha)rs\left(\mathcal{D}y^{\frac{1}{\mu}}+1\right)+\alpha\left(\mathcal{D}y^{\frac{1}{\mu}}+\mathcal{D}\right)\left(r^{\frac{\mu-1+\alpha}{\alpha}}-\delta\right)\right](\Pi_2-1)}\right]}{\partial\frac{d_1}{\gamma}}$$

$$=\frac{-(1-\alpha)(1-s)r\left[\left((1-\alpha)rs+\alpha\left(r^{\frac{\mu-1+\alpha}{\alpha}}-\delta\right)\right)\frac{\partial\mathcal{D}y^{\frac{1}{\mu}}}{\partial\frac{d_1}{\gamma}}+\frac{\partial\mathcal{D}}{\partial\frac{d_1}{\gamma}}\alpha\left(r^{\frac{\mu-1+\alpha}{\alpha}}-\delta\right)\right]}{(\Pi_2-1)\left[(1-\alpha)rs\left(\mathcal{D}y^{\frac{1}{\mu}}+1\right)+\alpha\left(\mathcal{D}y^{\frac{1}{\mu}}+\mathcal{D}\right)\left(r^{\frac{\mu-1+\alpha}{\alpha}}-\delta\right)\right]^2}\geq 0$$

$$\text{since } \frac{\partial \mathcal{D}}{\partial \frac{d_1}{\gamma}} = \frac{\partial \left[\frac{(1-s\Pi_2)(\Pi_1-1)}{(1-s\Pi_1)(\Pi_2-1)}\right]}{\partial \frac{d_1}{\gamma}} = \frac{(1-s\Pi_2)(1-s)}{(\Pi_2-1)(1-s\Pi_1)^2}\frac{\partial \Pi_1}{\partial \frac{d_1}{\gamma}} \leq 0$$

$$\frac{\partial y}{\partial \frac{d_1}{\gamma}} = \frac{\partial \left[\left[\left(\frac{e^2}{e^1}\mathcal{D}\right)^{1-\alpha}\frac{\theta^2}{\theta^1}\right]^{\frac{\mu}{\mu-1}}\right]}{\partial \frac{d_1}{\gamma}} = \left[\frac{\theta^2}{\theta^1}\right]^{\frac{\mu}{\mu-1}}\frac{\mu(1-\alpha)}{\mu-1}\left[\frac{e^2}{e^1}\mathcal{D}\right]^{\frac{\mu(1-\alpha)}{\mu-1}-1}\left[\frac{-e^2}{(e^1)^2}\mathcal{D}\frac{\partial e^1}{\partial \frac{d_1}{\gamma}} + \frac{e^2}{e^1}\frac{\partial \mathcal{D}}{\partial \frac{d_1}{\gamma}}\right] \leq 0$$

$$\frac{\partial y^{\frac{1}{\mu}}}{\partial \frac{d_1}{\gamma}} = \frac{1}{\mu}y^{\frac{1}{\mu}-1}\frac{\partial y}{\partial \frac{d_1}{\gamma}} \leq 0 \text{ and } \frac{\partial \mathcal{D}y^{\frac{1}{\mu}}}{\partial \frac{d_1}{\gamma}} = y^{\frac{1}{\mu}}\frac{\partial \mathcal{D}}{\partial \frac{d_1}{\gamma}} + \mathcal{D}\frac{1}{\mu}y^{\frac{1}{\mu}-1}\frac{\partial y}{\partial \frac{d_1}{\gamma}} \leq 0$$

$$\frac{\partial N_2}{\partial \frac{d_1}{\gamma}} = \frac{\partial \left[\frac{(1-\alpha)(1-s)r}{\left[(1-\alpha)rs\left(\mathcal{D}^{-1}y^{\frac{-1}{\mu}}+1\right)+\alpha\left(y^{\frac{-1}{\mu}}+1\right)\left(r\frac{\mu-1+\alpha}{\alpha}-\delta\right)\right](\Pi_2-1)}\right]}{\partial \frac{d_1}{\gamma}}$$

$$= \frac{-(1-\alpha)(1-s)r\left[(1-\alpha)rs\frac{\partial \mathcal{D}^{-1}y^{-\frac{1}{\mu}}}{\partial \frac{d_1}{\gamma}} + \frac{\partial y^{\frac{-1}{\mu}}}{\partial \frac{d_1}{\gamma}}\alpha\left(r\frac{\mu-1+\alpha}{\alpha}-\delta\right)\right]}{(\Pi_2-1)\left[(1-\alpha)rs\left(\mathcal{D}^{-1}y^{\frac{-1}{\mu}}+1\right)+\alpha\left(y^{\frac{-1}{\mu}}+1\right)\left(r\frac{\mu-1+\alpha}{\alpha}-\delta\right)\right]^2} \leq 0$$

$$\text{since } \frac{\partial y^{\frac{-1}{\mu}}}{\partial \frac{d_1}{\gamma}} = -\frac{1}{\mu}y^{-\frac{1}{\mu}-1}\frac{\partial y}{\partial \frac{d_1}{\gamma}} \geq 0 \text{ and } \frac{\partial \mathcal{D}^{-1}y^{\frac{-1}{\mu}}}{\partial \frac{d_1}{\gamma}} = -\mathcal{D}^{-2}\frac{\partial \mathcal{D}}{\partial \frac{d_1}{\gamma}}y^{\frac{-1}{\mu}} + \mathcal{D}^{-1}\frac{y^{\frac{-1}{\mu}}}{\partial \frac{d_1}{\gamma}} \geq 0$$

References

Agell, J. and P. Lundborg. (1995). 'Theories of Pay and Unemployment: Survey Evidence from Swedish Manufacturing Firms.', *Scandinavian Journal of Economics,* vol. 97, pp. 295-307

Albrecht, James W, and Susan B Vroman. (1999). 'Unemployment Compensation Finance and Efficiency Wages.' *Journal of Labor Economics,* vol. 17, pp. 141-67.

Albrecht, James W, and Susan B Vroman. (1992). 'Dual Labor Markets, Efficiency Wages, and Search.' *Journal of Labor Economics*, vol. 10, pp. 438-61.

Alexopoulos, Michelle. (2000). 'Unemployment and the Business Cycle', Working Paper, University of Toronto.

Alexopoulos, Michelle. (2001). 'Notes on Shirking Efficiency Wage Models.' Working Paper, University of Toronto.

Alexopoulos, Michelle. (2002). 'Spillovers and Wage Differentials' Working Paper, University of Toronto.

Basu, Susanto, and John Fernald. (1994). 'Constant Returns and Small Markups in U.S. Manufacturing.' International Finance Discussion Paper 483, September, Board of Governors of the Federal Reserve System, Washington, D.C.

Basu, Susanto, and John Fernald. (1994). 'Returns to Scale in U.S. Productionpp. Estimates and Implications.' *Journal of Political Economy*, vol. 105 pp. 249-83.

Bewley, T. F. (1999). *Why Wages Don't Fall During a Recession.* Cambrigde, Mass.: Harvard University Press.

Blackburn, McKinley and David Neumark. (1992). 'Unobserved Ability, Efficiency Wages, and Interindustry Wage Differentials.' *The Quarterly Journal of Economics* vol. 107, pp. 1421-36.

Borjas, George and Valerie Ramey. (2000). 'Market Responses to Inter-Industry Wage Differentials.' National Bureau of Economic Research Working Paper #7799.

Burnside, Craig and Martin Eichenbaum. (1996). 'Factor Hoarding and the Propagation of Business-Cycle Shocks.' *American Economic Review,* vol. 86, pp. 1154-73.

Burnside, Craig, Martin Eichenbaum and Jonas Fisher. (2000). 'Fiscal Policy Shocks in an Efficiency Wage Model.' National Bureau of Economic Research Working Paper #W7515.

Dickens, William and Lawrence Katz. (1987). 'Inter-Industry Wage Differences and Industry Characteristics.' In *Unemployment and the Structure of Labor Markets*, ed. Kevin Lang and Jonathan Leonard, New York, New York: Basil Blackwell Inc. pp. 103-40

Felices, Guillermo, (2001). 'Efficiency Wages in a New Keynesian Framework', Bank of England Working Paper.

Fehr, E., G. Kirchsteiger and A. Riedl. (1996). 'Involuntary Unemployment and N?on-compensating Differentials in an Experimental Labour Market.' *Economic Journal,* vol. 106, no. 434, pp. 106-21.

Gera, Surendra and Gilles Grenier. (1994). 'Interindustry Wage Differentials and Efficiency Wagespp. Some Canadian Evidence.' *Canadian Journal of Economics*, vol. XXVII. no.1, pp. 81-100.

Gibbons, Robert and Lawrence Katz. (1992). 'Does Unmeasured Ability Explain Inter-Industry Wage Differentials.' *The Review of Economic Studies*, vol. 59, no. 3, pp. 515-35.

Gomme, Paul. (1999). 'Shirking, Unemployment and Aggregate Fluctuations.' *International Economic Review*, vol. *40,* pp. 3-21

Groshen, Erica L and Alan Krueger. (1990). 'The Structure of Supervision and Pay in Hospitals.' *Industrial and Labor Relations Review*, vol. 43, no. 3, pp. S134-46

Hall, J. D. (1993). 'The Wage Setters Guide to Wage Rigidity.' Master's Thesis, University of Southampton.

International Customer Service Association. (2001). *Customer Service Industry Incentives, Bonuses and Employee Retention. Report to Members.*

Jaeger, David A . (1997). 'Reconciling Educational Attainment Questions in the CPS and the Census.' *Monthly Labor Review*, vol. 120, no. 8, pp. 36-40.

Katz, Lawrence. (1986). 'Efficiency Wage Theoriespp. A Partial Evaluation.' *NBER Macroeconomics Annual 1*, pp. 235-76.

Keane, M. (1993). 'Individual Heterogeneity and Inter-Industry Wage Differentials.' *Journal of Human Resources*, vol. 28, pp. 136-61.

Konings, J and P. Walsh, (1994). 'Evidence of Efficiency Wage Payments in UK Firm Level Panel Data.' *Economics Journal*, vol. 104, no. 424, pp. 542-55.

Krueger, Alan. (1991). 'Ownership, Agency, and Wages: An Examination of Franchising in the Fast Food Industry.' *Quarterly Journal of Economics*, vol. 106, no. 1, pp. 75-101.

Krueger, Alan and Lawrence Summers. (1988). 'Efficiency Wages and the Inter-Industry Wage Structure.' *Econometrica*, vol. 56, pp. 259-94.

Krueger, Alan and Lawrence Summers. (1987). 'Reflections on the Inter-Industry Wage Structure.' In *Unemployment and the Structure of Labor Markets*, ed. Kevin Lang and Jonathan Leonard, New York, New York: Basil Blackwell Inc. pp. 48-81.

Levine, David I. (1992). 'Can Wage Increases Pay for Themselves? Tests with a Production Function.' *Economic Journal*, vol. 102, pp. 1102-15.

MacLeod, Bentley and James Malconson. (1998). 'Motivations and Markets.' *American Economic Review*, vol. 88, no. 3, pp. 388-411.

Malcomson, James M. (1998). 'Individual Employment Contracts.' In *Handbook of Labor Economics,* vol. 3b, ed. Orley Ashenfelter and David Card, Amsterdam, The Netherlands: Elsevier Science B.V. pp. 2291-2372.

Murphy, K and Topel, R. (1987). 'Unemployment, Risk and Earnings' In *Unemployment and the Structure of Labor Markets*, ed. Kevin Lang and Jonathan Leonard, New York, New York: Basil Blackwell Inc. pp. 103-40.

Ravn, M. and H. Uhlig. (1997). 'On Adjusting the HP-Filter for the Frequency of Observations,' CentER Working paper.

Rebitzer, James B. (1995). 'Is There a Trade-Off between Supervision and Wages? An Empirical Test of Efficiency Wage Theory.' *Journal of Economic Behavior and Organization*, vol. 28, no. 1, pp. 107-29.

Sage. (1999). *SAGE System Administrator Salary Profile.* Report.

Shapiro, Carl and Stiglitz, Joseph E. (1984). 'Equilibrium Unemployment as a Worker Discipline Device.' *American Economic Review*, vol. 74, pp. 433-44.

Solow, Robert M. (1979). 'Another Possible Source of Wage Stickiness.' *Journal of Macroeconomics,* vol. 1, pp. 595-618.

Walsh, Frank. (1999). 'A Multi-Sector Model of Efficiency Wages.' *Journal of Labor Economics*, vol. 17, no. 2, pp. 351-76.

TABLE 1A: INTER-INDUSTRY WAGE DIFFERENTIALS FROM 1973-1981

Industry	1973	1974	1975	1976	1977	1978	1979	1980	1981
MINING	0.17570	0.18851	0.22051	0.21630	0.17902	0.20263	0.15098	0.25954	0.22241
	(0.03595)	(0.03777)	(0.03710)	(0.04100)	(0.03203)	(0.03340)	(0.03182)	(0.02734)	(0.02729)
CONSTR	0.21950	0.19967	0.17182	0.18173	0.16837	0.11916	0.13848	0.12943	0.12268
	(0.01568)	(0.01647)	(0.01812)	(0.01830)	(0.01561)	(0.01507)	(0.01490)	(0.01422)	(0.01498)
MANUF	0.04094	0.04843	0.04528	0.05285	0.06504	0.06915	0.06889	0.07295	0.07118
	(0.00651)	(0.00697)	(0.00743)	(0.00762)	(0.00696)	(0.00710)	(0.00665)	(0.00618)	(0.00656)
TRANSP	0.13952	0.12264	0.13941	0.15198	0.12011	0.13762	0.08433	0.12756	0.10857
	(0.01537)	(0.01601)	(0.01588)	(0.01630)	(0.01441)	(0.01467)	(0.01350)	(0.01203)	(0.01290)
TRADE	-0.14031	-0.11667	-0.12434	-0.11958	-0.12141	-0.12436	-0.10458	-0.10267	-0.09539
	(0.00790)	(0.00831)	(0.00808)	(0.00844)	(0.00700)	(0.00736)	(0.00695)	(0.00642)	(0.00683)
FINANCE	0.06062	0.06422	0.07053	0.02842	0.04259	0.06834	0.03968	0.02924	0.05585
	(0.01674)	(0.01732)	(0.01712)	(0.01705)	(0.01490)	(0.01544)	(0.01387)	(0.01263)	(0.01306)
SERVICE	-0.05322	-0.08191	-0.04462	-0.05976	-0.05171	-0.06188	-0.06051	-0.07807	-0.08238
	(0.00948)	(0.00973)	(0.00957)	(0.00970)	(0.00841)	(0.00851)	(0.00792)	(0.00686)	(0.00721)

TABLE 1B: INTER-INDUSTRY WAGE DIFFERENTIALS FROM 1982-1990

Industry	1982	1983	1984	1985	1986	1987	1988	1989	1990
MINING	0.26339	0.23934	0.20654	0.27052	0.17485	0.16681	0.26291	0.21627	0.19010
	(0.03006)	(0.03132)	(0.02993)	(0.03414)	(0.03348)	(0.03944)	(0.04039)	(0.03625)	(0.03866)
CONSTR	0.11125	0.11243	0.11717	0.11875	0.12067	0.12114	0.11242	0.09194	0.12739
	(0.01663)	(0.01709)	(0.01553)	(0.01543)	(0.01514)	(0.01570)	(0.01589)	(0.01607)	(0.01641)
MANUF	0.09522	0.08582	0.09115	0.09793	0.08508	0.08697	0.07440	0.08559	0.08435
	(0.00729)	(0.00764)	(0.00728)	(0.00745)	(0.00741)	(0.00746)	(0.00777)	(0.00754)	(0.00822)
TRANSP	0.18219	0.15495	0.13703	0.14397	0.16402	0.14012	0.11755	0.09343	0.10295
	(0.01362)	(0.01400)	(0.01330)	(0.01327)	(0.01323)	(0.01393)	(0.01435)	(0.01427)	(0.01378)
TRADE	-0.13073	-0.11884	-0.12377	-0.13101	-0.12994	-0.12381	-0.13317	-0.12590	-0.12888
	(0.00721)	(0.00769)	(0.00734)	(0.00747)	(0.00717)	(0.00754)	(0.00804)	(0.00778)	(0.00765)
FINANCE	0.05220	0.06170	0.06072	0.03810	0.08024	0.06118	0.09251	0.07348	0.07215
	(0.01365)	(0.01397)	(0.01292)	(0.01322)	(0.01284)	(0.01300)	(0.01340)	(0.01383)	(0.01334)
SERVICE	-0.08305	-0.06821	-0.06652	-0.06306	-0.05866	-0.05504	-0.04745	-0.05553	-0.03165
	(0.00744)	(0.00731)	(0.00702)	(0.00713)	(0.00667)	(0.00692)	(0.00715)	(0.00784)	(0.00684)

TABLE 1C: INTER-INDUSTRY WAGE DIFFERENTIALS FROM 1991-1999

Industry	1991	1992	1993	1994	1995	1996	1997	1998	1999
MINING	0.16923	0.15670	0.29574	0.18339	0.23871	0.12692	0.09028	0.14950	0.17886
	(0.04241)	(0.04032)	(0.04125)	(0.04505)	(0.04534)	(0.05153)	(0.04551)	(0.04568)	(0.05202)
CONSTR	0.11102	0.11688	0.09797	0.09038	0.06835	0.08504	0.10948	0.13262	0.07849
	(0.01717)	(0.01703)	(0.01742)	(0.01834)	(0.01749)	(0.01890)	(0.01726)	(0.01728)	(0.01747)
MANUF	0.09541	0.09797	0.09460	0.08179	0.07874	0.08631	0.07299	0.07677	0.06021
	(0.00832)	(0.00844)	(0.00865)	(0.00907)	(0.00876)	(0.00964)	(0.00963)	(0.00986)	(0.00979)
TRANSP	0.13228	0.12539	0.13199	0.11175	0.08179	0.08985	0.13219	0.06132	0.08367
	(0.01427)	(0.01459)	(0.01514)	(0.01562)	(0.01534)	(0.01652)	(0.01506)	(0.01546)	(0.01597)
TRADE	-0.13598	-0.12187	-0.13375	-0.13732	-0.12193	-0.11021	-0.11858	-0.12026	-0.10948
	(0.00775)	(0.00776)	(0.00776)	(0.00827)	(0.00816)	(0.00848)	(0.00817)	(0.00828)	(0.00816)
FINANCE	0.11002	0.09821	0.10186	0.11382	0.09183	0.08529	0.06949	0.09181	0.08933
	(0.01361)	(0.01349)	(0.01361)	(0.01482)	(0.01387)	(0.01557)	(0.01462)	(0.01478)	(0.01485)
SERVICE	-0.04623	-0.05401	-0.03878	-0.01944	-0.01938	-0.02451	-0.02185	-0.01688	-0.00712
	(0.00675)	(0.00679)	(0.00676)	(0.00695)	(0.00672)	(0.00734)	(0.00686)	(0.00692)	(0.00671)

TABLE 2: INTER-INDUSTRY WAGE DIFFERENTIALS AND THE VARIANCE OF INDUSTRY WAGES

Industry	Avg. Rank of Differential	Average Differential	Variance of Wages
MINING	1.1481	0.1998	0.0382
			(0.0037)
CONSTR	2.7778	0.1250	0.0289
			(0.0028)
MANUF	4.2963	0.0765	0.0137
			(0.0013)
TRANSP	2.5185	0.1229	0.0252
			(0.0024)
TRADE	7.0000	-0.1224	0.0083
			(0.0009)
FINANCE	4.2593	0.0705	0.0110
			(0.0008)
SERVICE	6.0000	-0.0501	0.0128
			(0.0010)

TABLE 3: CYCLICALITY OF DIFFERENTIALS

Industry	α	β	γ
MINING	0.00580	-0.00014	-0.00042
	(0.02740)	(0.01430)	(0.00179)
CONSTR	-0.01160	0.00049	0.00046
	(0.00893)	(0.00468)	(0.00059)
MANUF	0.00744	-0.00058	-0.00050
	(0.00383)	(0.00201)	(0.00025)
TRANSP	-0.00330	0.01010	0.00011
	(0.01310)	(0.00685)	(0.00086)
TRADE	0.00205	-0.00060	-0.00007
	(0.00522)	(0.00273)	(0.00034)
FINANCE	-0.00079	0.00198	0.00015
	(0.00915)	(0.00480)	(0.00060)
SERVICE	-0.00435	0.00297	0.00046
	(0.00575)	(0.00301)	(0.00038)

Table IV.

Variable	Model's steady	CPS Data			Industry:
in model	state value (%)	Mean (%)	Max.(%)	Min(%)	
$\frac{N^1}{\sum_{i=1}^{7} N^i} * 100\%$	1	1	2	1	Mining
$\frac{N^2}{\sum_{i=1}^{7} N^i} * 100\%$	7	7	8	7	Transportation
$\frac{N^3}{\sum_{i=1}^{7} N^i} * 100\%$	6	6	7	5	Construction
$\frac{N^4}{\sum_{i=1}^{7} N^i} * 100\%$	7	7	9	6	Finance, Insurance and Real Estate
$\frac{N^5}{\sum_{i=1}^{7} N^i} * 100\%$	26	26	35	19	Manufacturing
$\frac{N^6}{\sum_{i=1}^{7} N^i} * 100\%$	27	27	33	19	Service
$\frac{N^7}{\sum_{i=1}^{7} N^i} * 100\%$	26	26	27	24	Trade

Table V.

Variable	Model's steady	CPS Data			Avg. Rank of
in model	state value (%)	Mean (%)	Max.(%)	Min (%)	Differential
$\left(\frac{w^1}{\overline{w}} - 1\right) \times 100\%$	20	20	30	9	1.1
$\left(\frac{w^2}{\overline{w}} - 1\right) \times 100\%$	12	12	18	6	2.5
$\left(\frac{w^3}{\overline{w}} - 1\right) \times 100\%$	12	12	22	7	2.8
$\left(\frac{w^4}{\overline{w}} - 1\right) \times 100\%$	7	7	11	3	4.3
$\left(\frac{w^5}{\overline{w}} - 1\right) \times 100\%$	8	8	10	4	4.3
$\left(\frac{w^6}{\overline{w}} - 1\right) \times 100\%$	-5	-5	-8	1	6.0
$\left(\frac{w^7}{\overline{w}} - 1\right) \times 100\%$	-12	-12	-14	-10	7.0

where $\overline{w} = \left(\sum_{i=1}^{7} w^i N^i\right) / \left(\sum_{i=1}^{7} N^i\right)$

Figure 1. Differentials and the Unemployment rate

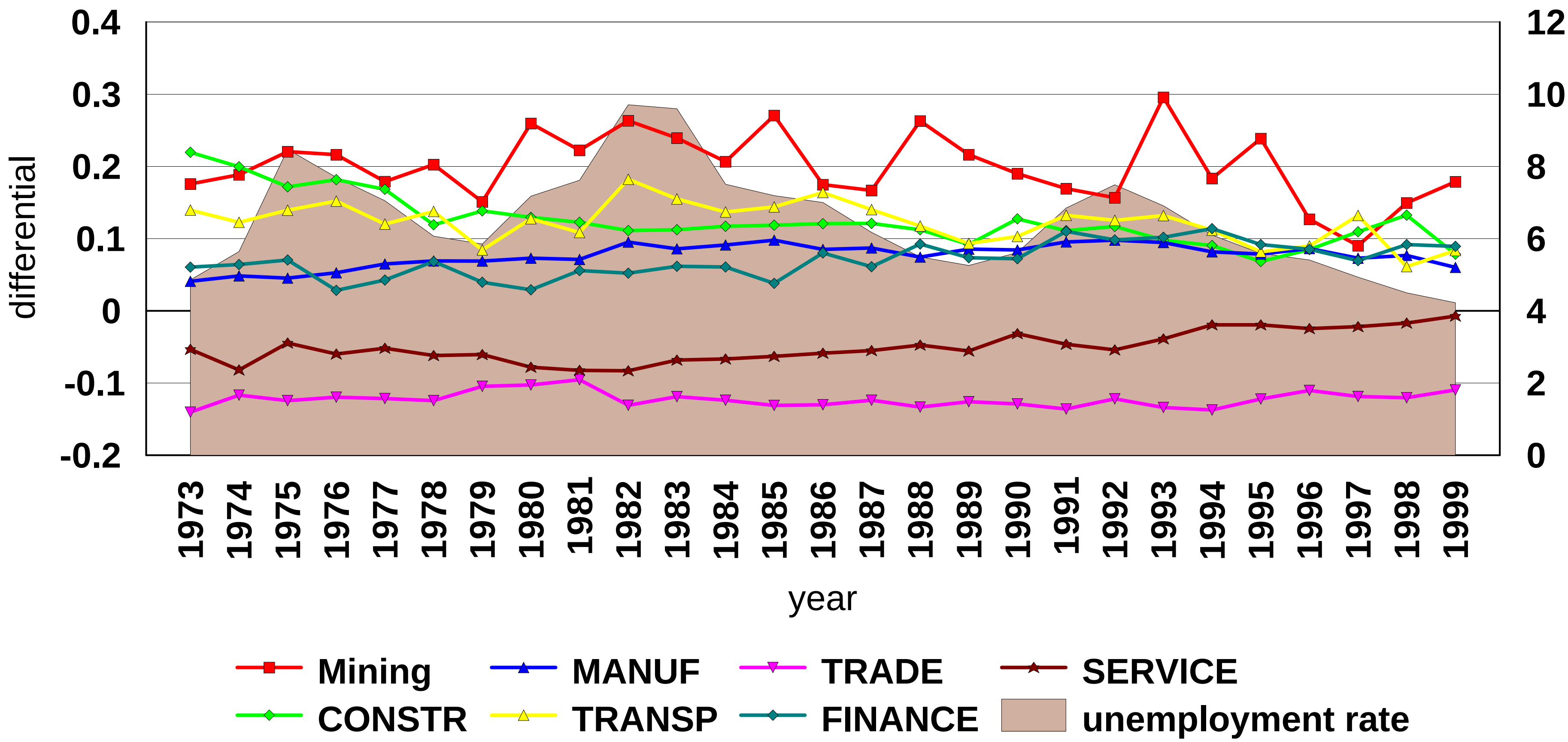

Figure 2. Impulse Responses to a Positive Technology Shock